Melco Foundation
Management Accounting Research
Discussion Paper Series

The Melco Foundation

Melco Management Accounting Research Discussion Paper Series

No.MDP2017-002

The Effects of Interaction among Quantity, Characteristics of Accounting Measures, and Manager's Cognitive Style on Performance Evaluations

January 2017

Koichi Hioki
Assistant Professor in Marketing,
The Graduate School of Business Administration, Kobe University

Eiichiro Suematsu*
Professor in Management Accounting,
The Graduate School of Humanities and Social Sciences, Saitama University
255 Shimo-okubo, Sakura-ku, Saitama 338-8570
e-mail: suematsu@mail.saitama-u.ac.jp

Hiroshi Miya
Professor in Management Accounting,
The Graduate School of Business Administration, Kobe University
*Corresponding author

The Melco Foundation

Nagoya, Japan

Discussion Paper Series of the *Melco Management Accounting Research*, in order to promote the study of management accounting, has published this unfinished paper onto the Web. Please obtain the permission of the author when citing this paper.

The Effects of Interaction among Quantity, Characteristics of Accounting Measures, and Manager's Cognitive Style on Performance Evaluations

Koichi Hioki,
Assistant Professor in Marketing,
The Graduate School of Business Administration, Kobe University

Eiichiro Suematsu (corresponding author),
Professor in Management Accounting,
The Graduate School of Humanities and Social Sciences, Saitama University,
e-mail: suematsu@mail.saitama-u.ac.jp

Hiroshi Miya,
Professor in Management Accounting,
The Graduate School of Business Administration, Kobe University

Abstract

This study aims to investigate appropriate numbers and kinds of accounting measures in Balanced Scorecard for managers to be utilized in their decision making. The authors apply an experimental method with 54 participants who work for a utility company in Japan. The results showed that under information overload in which many measures are handled simultaneously, managers who have high Need for Cognition can no longer use either financial or non-financial measures effectively, while they can use these measures when there is no information overload. Managers with low Need for Cognition do not use measures of customer perspective and other non-financial perspectives even when information overload does not occur. This study concluded that we need to pay careful attention to differences in managers' Need for Cognition as well as how many and what kind of measures should be provided to managers when designing multi-measures for managers in a management accounting system, such as the Balanced Scorecard.

Keywords BSC, Information Overload, Number of Measures, Characteristics of Measures, Need for Cognition

1. Introduction

How many and what kind of accounting measures should be provided to managers? This is a very interesting question in managerial accounting systems that promotes appropriate decisions by managers. We investigate how many measures and what kind of measures should be provided to managers to promote appropriate decision making and performance evaluation in the Balanced Scorecard (BSC). In addition, we investigate the relationship between managers' personal cognitive styles and measures that should be provided to them via the BSC.

In this study, first, we clarify how many measures managers can use in the BSC. Prior studies regarding agency theory and BSC suggest that providing managers with a larger number of measures leads to appropriate decision making and performance evaluation (Holmström, 1979; Holmström and Milgrom, 1991; Milgrom and Roberts, 1992, pp.228–229; Feltham and Xie, 1994; Kaplan and Norton, 1996, 2001, 2004; Malina and Selto, 2001;). Some BSC research, however, has indicated that while many measures are provided by the BSC, managers use a limited portion of financial and non-financial measures rather than all of them (Lipe and Salterio, 2000; Ittner, Larcker, and Meyer, 2003). These two inconsistent claims might suggest that decision accuracy improves as the number of measures increases; however, decision accuracy would be exacerbated when the number of measures increases beyond a certain number of measures. Psychological studies have suggested that there is a limit to the quantity of information to process (limitation of cognitive capacity) and we cannot adequately perform information processing under situations in which the quantity of information exceeds a certain quantity of information (e.g., Miller, 1956; Tuttle and Burton, 1999; Swain and Haka, 2000; Eppler and Mengis, 2004;). This situation is known as information overload, under which we do not necessarily use all provided information, as the quantity of information exceeds the limitation of cognitive capacity. Therefore, we ignore some information in order to release cognitive capacity. From this perspective, we assume that due to the limitations of cognitive capacity, managers might not be able to use all measures provided by the BSC. In this study, we examine how many measures cause information overload in the BSC.

Second, we clarify what kind of measures should be provided to managers in the BSC. In this study, we focus on the characteristics of measures. Lipe and Salterio (2000) indicated that managers tend to focus on common measures for departments when they evaluate the departments, rather than on unique measures for each department. In addition, Ittner et al. (2003) suggested that financial measures are the most frequently used, followed by measures of customer perspectives of

BSC. These studies indicated that managers have selection bias on measure usage. Furthermore, psychological studies have indicated that characteristics of information are a factor that determines the definitive quantity of information that causes information overload. We examine what characteristics of measures are used if managers cannot use all measures provided to them in the BSC.

Third, we test the effects of managers' cognitive styles. Psychological studies have indicated that people have their own personal cognitive styles, that is, personal tendencies of information processing. Some people like to consider certain problems or enjoy themselves, whereas others do not. We expect that managers who prefer to consider problems use more measures than those who do not. In this study, we examine how managers' cognitive styles influence the usage of measures in the BSC.

In this study, we investigate the relationships among the number of measures, characteristics of measures, and personal cognition style by adopting an experimental method. We contribute to the performance evaluation literature by showing that careful attention should be paid to these relationships. Various factors determine the quantity of information causing information overload, such as characteristics of information, quality of information, task and the process parameters, and personal factors. We study not only how the characteristics of measures determine the number of measures that causes information overload, but also how a manager's cognitive style determines the number.

The remainder of the paper is organized as follows. Section 2 summarizes prior research to develop hypotheses. Thereafter, we describe our experiment in Section 3 and results in Section 4. The final section discusses our conclusions.

2. Theoretical background and Hypothesis development

Many experimental studies regarding the BSC has attempted to mitigate information overload, under a condition in which managers are given 16 measures (e.g., Libby, Salterio, and Wbb, 2004; Banker, Cang, and Pizzini, 2004, 2011; Cardinals and van Veen-Dirks, 2010; Ding and Beaulieu, 2011; Humphreys and Trotman, 2011). However, prior psychological studies have mentioned that people can recognize far less than 16 measures. Therefore, managers might not be able to utilize all 16 measures, even when accounting systems attempt to mitigate information overload.

The previous psychological research has revealed that the quantity of information that causes information overload is determined by various factors. Jackson and Farzaneh (2012) classified the information overload factors into three factor types: intrinsic factors, extraneous factors, and the interaction of intrinsic and extraneous factors. Intrinsic factors directly cause information overload, and extraneous factors affect intrinsic factors and as a result, indirectly cause information overload. Jackson and Farzaneh (2012) mentioned quantity of information,

information processing capacity, and available time as intrinsic factors, and characteristics of information, quality of information, task and process parameters, and personal factors as extraneous factors. In addition, the authors mentioned source of information as intrinsic and extraneous factors. These various factors determine the quantity of information causing information overload.

We pick up characteristics of information and personal cognitive style from the information overload factors sorted by Jackson and Farzaneh (2012). We examine how the number of measures causing information overload is affected by the characteristics of each measure (four perspectives in the BSC) and by managers' cognitive styles.

Number of measures

Prior accounting research has justified that a large number of measures causes information overload. Swain and Haka (2000) showed that when many information cues (i.e., the amount of information available) of budgeting alternatives and dimensions are provided, managers do not necessarily use all of them. Ittner et al. (2003) clarified that managers subjectively neglect, add, and change the weight of a performance measure evaluation for their subordinates when they have many performance measures. In addition, Lipe and Salterio (2000) used an experimental method to show that managers provided with 16 measures did not necessarily use all measures. These studies might suppose that when the number of measures exceeds a limit of managers' cognitive capacity, they do not necessarily use all of them owing to information overload.

Prior research in psychology has clarified that approximately seven kinds of information are the limit of our cognitive capacity for simultaneous processing. (e.g., Miller, 1956). Tuttle and Burton (1999) referred to prior research regarding the information overload effect on information usage. According to Tuttle and Burton (1999), many studies have clarified that most individuals cannot use more than seven kinds of information, and even high-ability individuals cannot use all information when more than nine kinds of information are provided. Tuttle and Burton (1999) indicated the same results in their experiment that gave six or nine cues to participants to estimate stock prices. Nine cues were not necessarily used by the participants in their study to estimate stock prices.

The appropriate number of measures in the BSC was examined by Ding and Beaulieu (2011). They manipulated the information load by changing the number of measures provided for managers from 2 to 8 and then 16. Ding and Beaulieu (2011) found that mood congruency bias to evaluate the performance of subordinates can be canceled out by financial incentives when the information load is low (i.e., the number of measures is less than eight). Their result assumed that more than seven measures cause information overload, and was consistent with the results of Tuttle and Burton (1999). In this study, we expect that seven measures provided to managers in the BSC is the number causing information overload. We formalize the following hypothesis.

H1: More than seven measures cause information overload in the BSC system.

Financial versus non-financial measures

Ittner et al. (2003) mentioned that many BSC studies have assumed that each perspective has different characteristics; financial perspective is an outcome measure and an externally oriented measure, customer perspectives is a leading measure and an externally oriented measure, and internal business processes and learning and growth perspectives are leading measures and internal measures. Ittner et al. (2003) suggested that financial perspective is the most frequently used measure, followed by customer perspective. They explained that this measure usage tendency is because managers consider external measures as measures of desired outcomes. Their results are consistent with other studies indicating that it is easier to understand a direct relationship between cause and result than an indirect relationship (e.g., Luft and Shields, 2001; Farrell et al., 2007). Luft and Shields (2001), Ittner et al. (2003), and Farrell et al. (2007) suggested that measures of internal business processes and learning and growth perspectives have a more ambiguous relationship with financial perspective than do measures of customer perspective. In other words, the relationship between financial perspective and customer perspective might be more obvious than that between financial perspective and other non-financial perspectives.

Jackson and Farzaneh (2012) mentioned that ambiguity of information, as one kind of information characteristic, is an information overload-causing factor. Therefore, when information overload occurs due to a larger number of measures, we can assume that managers abandon ambiguous measures, and use only measures of financial perspective and customer perspective in order to release cognitive resources (Jackson and Farzaneh, 2012).

We formalize the following hypotheses regarding the characteristics of measure (financial , customer, internal control, and learing and growth measures) on measure usage.

H2a: Managers under the information overload condition try to use both financial and customer measures, rather than internal control measures and learning and growth measures, to reduce information load.

H2b: In the case that excess measures are provided, managers give up using financial and customer measures as well as other measures (internal processes, and learning and growth).

Personal cognitive style

Many psychological studies have indicated that there are individual differences in the strategy of information processing. Need for Cognition (NFC) is a famous example of cognitive style. NFC is the motivation to think about a certain problem or enjoy oneself. People with high NFC try more actively to process complicated information and to understand content than do those with low NFC. Thus, we employ NFC as a parameter to measure the strategy of accounting information processing by managers. We expect that managers with high NFC try actively to use

all measures provided in the BSC than do those with low NFC.

Drolet, Luce and Simonson (2009) showed that under a situation with a large quantity of information processing, people with high NFC cannot adequately carry out information processing and tend to make compromised decisions compared to a situation with a small quantity of information processing. Drolet et al. (2009) mentioned that this information processing tendency is presumably because people with high NFC try to use information actively and they expend a great deal of their cognitive capacity when provided with excessive information, and thus, they compromise on decision making. In other words, managers with high NFC might give up using measures when provided with excess amounts of measures. Thus, we formalize the following hypothesis.

H3: Managers with high NFC use more measures than do those with low NFC in the BSC system.

3. Experiment

Experimental design

We designed a 3 (the number of provided measures; 4 vs. 7 vs. 10) × 2 (cognitive styles) factorial between participants. The number of measures that cause information overload is expected to be less than nine, according to prior information overload studies. Therefore, we set 3 cases that have 4, 7, and 10 measures, and included the characteristics of measures in the three cases in order to find out the amount of measures causing information overload. Participants were divided into three groups: the first group was provided with only 4 measures belonging to the financial perspective; the second group was provided with 7 measures, of which 4 belong to the financial perspective and 1 each to the customer, internal business processes, and learning and growth perspectives; and the last group was provided with 10 measures, of which 4 belong to the financial perspective and 2 each to the other three perspectives. We employed NFC (Cacioppo and Petty, 1982) to measure managers' cognitive styles. NFC refers to the motivation to think about a certain problem or enjoy oneself.

Experimental materials

Referring to a scenario shown in Banker et al. (2011), we built a case of an American clothing sales company. Before building the scenarios, we asked two managers of a Japanese major department store if they could accept the measures used as common measures in Lipe and Salterio (2000), Banker et al. (2004), and Humphreys and Trotman (2011). Based on their suggestion, we changed two of the financial measures, as well as one of the customer measures. All other measures were the same as the common measures in Lipe and Salterio (2000), Banker et al. (2004), and Humphreys and Trotman (2011).

In all conditions, the number of measures in which actual performances are better than the

targets is the same number of measures in which the actual performances are worse than the targets (Table 1). In the four-item condition, the actual performances of two financial measures were better than the targets by 8.5%, and actual performances of the other two financial measures were worse than the targets by 8.5%. In the seven-item condition, three measures of the other perspectives were added to the measures in the four-item condition. The actual performance of one of the added three measures was better than the target by 8.5%, the second was worse than the target by 8.5%, and the last was the same as the target. In the 10-item condition, three measures of the other perspectives were added to the 7-item condition. The actual performance of one of the added three measures was better than the target by 8.5%, the second was worse than the target by 8.5%, and the last was the same as the target. Thus, if managers were to use all of the provided measures, performance evaluations by managers should not differ among the three conditions.

data

A utility company in Japan cooperated with our research; a total of 73 employees of the company participated in the experiment. The data of 19 participants were excluded from our analyses due to their low scores on recognition task (below and equal to the chance level). For the analysis, we applied data of 54 participants (10 females and 44 males, mean age =37.30, SD = 8.03, mean of managerial experience = 2.63 years, SD = 4.40).

Procedures

The experimental sessions were administered in a class. Each participant was given a booklet containing a fictional company scenario. The participants were divided into three groups; the first was provided with only 4 measures belonging to the financial perspective (named the four-item condition in Table 1), the second was provided with 7 measures, of which 4 belong to the financial perspective and 1 each to the customer perspective, internal business process perspective, and learning and growth perspective (named the 7-item condition in Table 1), and the third was provided with 10 measures, of which 4 belong to the financial perspective and 2 each to the other perspectives (named the 10-item condition in Table 1).

In the scenario, each participant was assigned as an executive officer of the company. After participants completed reading the scenario, they were presented with 4/7/10 performance measures of a business division they managed (see Table 1). Participants were asked to evaluate the performances both of the business division they managed and its director, and to score those out of 100 (0–100 points) based on the performance measures in 1 minute. Promptly after completing the questions, participants rated their levels of confidence in their evaluations using 7-point scales (1: not at all; 7: very confident). In addition, participants were asked to rate their perceptions of importance of each performance measure using 5-point scales (1: not at all; 5: very much). Participants were asked to complete to answer these questions in 2 minutes.

After completing these questions, each participant enclosed the scenario and questionnaire in an envelope. Promptly, participants were asked to perform a recognition task as a surprise task. In this task, participants were presented with 10 performance measures and were asked to rate whether each measure was better or worse than its targeted value on a three-point scale (better/equal/worse). If they thought the measures were not provided in the scenario or did not recall them, they were asked to answer "no show" or "have no idea" instead. Participants were asked to complete this task in 2 minutes (see Table 2). The recognition test was conducted because we considered that the score of the recognition test should be proxy variables indicating how much the participants gave due consideration to the measures and tried to use them.

Finally, participants answered NFC questionnaires and demographic questions. At the end of the experimental session, participants were fully debriefed and thanked for their participation.

Table 1. Targets of three conditions

	Target	Actual	Shown in condition	Perspective
Sales margins*	**46%**	**49.91%**	**4, 7 & 10item**	**financial**
Operating profit on sales*	**24%**	**26.04%**	**4, 7 & 10item**	**financial**
Sales growth*	**35%**	**32.03%**	**4, 7 & 10item**	**financial**
Inventory turnover*	**6times**	**5.49times**	**4, 7 & 10item**	**financial**
Current price relative to the price of previous year*	+5%	+5%	**10 item**	**customer**
Customer satisfaction rating*	80%	**86.80%**	**7 & 10 item**	**customer**
Time to process customer returns	4 minutes	4.34 minutes	**10 item**	**internal business processes**
Return to suppliers	6%	**6%**	**7 & 10 item**	**internal business processes**
Hours of emploee training per employee*	80 hours	86.8 hours	**10 item**	**learning & growth**
Employee satisfaction*	75%	**68.63%**	**7 & 10 item**	**learning & growth**

Note: An increase in measures with * represents improvement in performance, and those without * have the opposite meaning

Table 2. Presented measures in each condition and correct answers in recognition task

	4-item condition	7-item condition	10-item condition
Sales margins	**better**	**better**	**better**
Operating profit on sales	**better**	**better**	**better**
Sales growth	**worse**	**worse**	**worse**
Inventory turnover	**worse**	**worse**	**worse**
Current price relative to the price of previous	not shown	not shown	**equal**
Customer satisfaction rating	not shown	**better**	**better**
Time to process customer returns	not shown	not shown	**worse**
Return to suppliers	not shown	**equal**	**equal**
Hours of emploee training per employee	not shown	not shown	**better**
Employee satisfaction	not shown	**worse**	**worse**

Note: Items marked as "not shown" were not presented to participants in the evaluating task

4. Results

NFC score. We created a composite score by averaging the 15 items measuring NFC (Cronbach's $\alpha = .85$). An analysis of variance (ANOVA) test showed no differences in the NFC level in three conditions: ($F(2, 51) < 1$, ns., $M_{4\text{-item}} = 4.76$, $SD_{4\text{-item}} = 0.86$; $M_{7\text{-item}} = 4.45$, $SD_{7\text{-item}} = 0.74$; and $M_{10\text{-item}} = 4.65$, $SD_{10\text{-item}} = 0.87$).

Regression model. We translated the conditions to a dummy code named Cond, coded 0 for the 4-item condition, 1 for the 7-item condition, and −1 for the 10-item condition. To test the impact of the number of measures (4 vs. 7 vs. 10; recoded 0, 1, and −1, respectively), NFC, and their interactions, we made a hierarchical multiple regression, entering the two main effects in the first step and the interaction term in the second step. Prior to these analyses, dummy codes and NFC were centered (i.e., the mean of the variable was subtracted by each score; see Cohen and Cohen, 1983).

Evaluations of director, business division, and confidence. To test the impact of the number of measures, NFC, and these interactions on evaluations of director and business division, we did four hierarchical multiple regressions. In all regressions, there was no significant effect of the impact of the number of measures, NFC, and these interactions.

Accuracy rate in recognition task. To investigate how information was used, we calculated accuracy in recognition tasks. The correct answers are shown in Table 2. First, we calculated grand accuracy by dividing the number of correct answers by 10. Sequentially, we calculated the accuracy of the "financial," "customer," "internal business processes," and "leaning and growth" items. Means of accuracy are reported in Table 3. To test the effects of the number of

measures, we did an ANOVA on the grand accuracy rate. As anticipated, the participants in the 4item condition had significantly higher scores than those of the other conditions. However, there was no significant difference between the 7-item and 10-item conditions.

To test the interaction effects between the number of measures and NFC, we did four multiple hierarchical regressions by using accuracy rates for each measure of perspectives as dependent variables and by using the number of measures, NFC, and their interaction as predictors. The results are shown in Table 4. As expected, in all models, the effect of the number of measures was a reliable predictor of hit rate[1]. More importantly, when predicting the accuracy rate of customer perspective measures, there was an interaction effect between the number of measures and NFC. Simple slopes in which NFC predicted the accuracy rate were obtained for the three conditions. Given that the experimental condition is an ordinal variable, the values corresponding to the 4-item (control), 7-item, and 10-item were chosen (0, 1, and −1, respectively) as single values. The results revealed that during the 7-item condition, as shown in Figure 1, high NFC participants more effectively used customer relation information than low NFC participants did when they were given 7 measures (B= 10.79, t (48) = 1.62, p<.10, one-tailed). By contrast, when they were given 10 or 4 measures, there were no differences for the accuracy rates between high NFC participants and low NFC participants ($t_{4\text{-item}}$ (48) =0.86, $t_{10\text{-item}}$ (48) =1.13, ns.). This result could indicate that high NFC participants used customer relations information inefficiently when they were provided with 10 measures compared to when they were provided with 7 measures.

As shown in Table 3, regarding the total scores of the recognition test, one-factor ANOVA was performed for the number of measures. According to the result, the total score of the group provided with four measures was measurably higher than the scores of the other groups. There was no significant difference in the total scores between the groups with 7 measures and 10 measures. The results supported Hypothesis 1.

Then, in order to measure the interaction between the number of measures and NFC for each perspective, hierarchical regressions were conducted using accuracy rates for each measure of perspective as dependent variables and by using the number of measures, NFC, and their interaction as predictors. As a result, the number of measures was found to be a reliable predictor in the accuracy rate for the recognition test for all perspectives. In particular, interaction between the number of measures and NFC was observed for the customer perspective. It was shown that for participants with high NFC, in the case of 7 measures, the accuracy rate of the recognition test for measures of customer perspective was significantly high compared with the case of 10 measures. These results support Hypotheses 2a, 2b, 3.

[1] Two models (financial and internal business process) were not significant. We used these models simply as a reference.

Table 3. Mean accuracy (standard deviation)

Condition	Grand	Financial	Customer	Internal business processes	Learning and growth
4-item	94.1% (8.0)	87.5% (18.5)	95.5% (14.7)	100% (0.0)	100% (0.0)
7-item	83.5% (13.7)	72.1% (23.2)	85.3% (29.4)	88.2% (21.9)	100% (0.0)
10-item	74.7% (12.5)	86.7% (18.6)	66.7% (24.4)	66.7% (30.9)	66.7% (30.9)

Table 4. Summary of hierarchical regression analysis for variables predicting accuracy rates of four indexes (N = 54)

variable	financial	customer	internal business processes	learning and growth
	β	β	β	β
step 1				
condition	−.27 +	.30 *	.33 *	.58 **
NFC	−.06	.03	.03	.17
step 2				
condition x NFC	.10	.27 *	−.06	−.13
model significance				
corrected R^2	.03	.09	.06	.34
F	1.63 n.s.	2.83 *	2.11 n.s.	10.27 **

Note: +p < .10, *p < .05, **p < .01

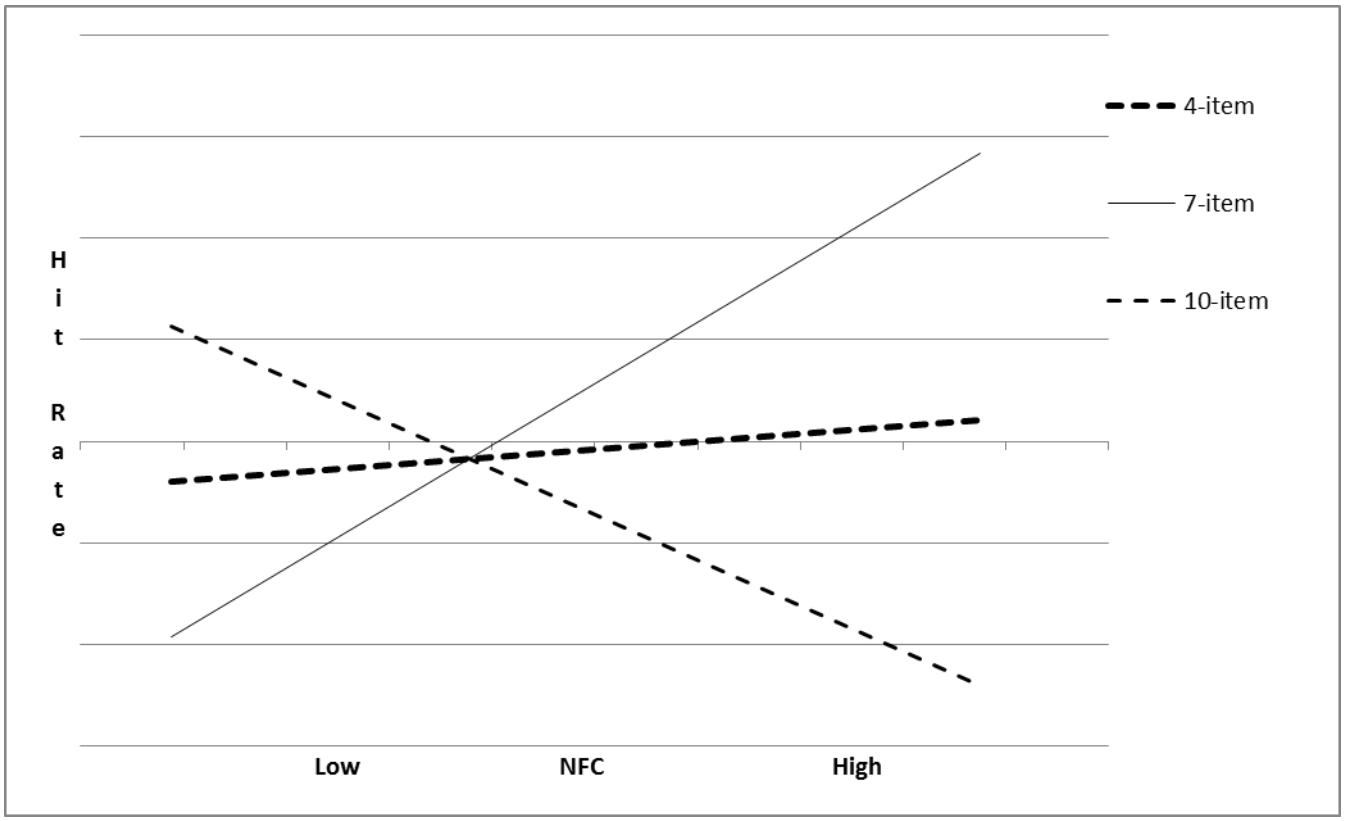

Figure 1. Interaction effects between number of items and NFC on accuracy rate of customer relations

5. Discussion and Conclusion

We examined three points in this study: first, providing managers with a lot of financial and non-financial measures causes information overload, second, managers do not necessarily use all the measures, and third, the measure usage tendency changes with either the degree of NFC or the characteristics of measures. The following three results were obtained.

First, all of the provided measures were not necessarily used on the increase in the number of measures, since the accuracy rate of the recognition test became worse as the number of measures increased. In other words, even if non-financial measures were added to financial measures, managers did not use the non-financial measures owing to information overload. Second, managers with high NFC seemed to use the non-financial measures when 7 measures are provided. In this study, the accuracy rate for measures of financial perspective is lowest in the 7-item condition and the rate for measures of customer perspective in the 7-item condition is higher than in the 10-item condition. These results suggest that managers tried to use measures of financial and customer perspectives in the 7item condition. On the contrary, in the case of the

10-item condition, the accuracy rate for measures of financial perspective was high compared with that for the 7-item condition. These results suggest that managers ceased to use measures of perspectives, except for financial perspective, in order to lower their cognition load, and their attention became focused on measures of financial perspective.

Third, concerning measures of the perspectives of internal business processes, and of learning and growth, the accuracy rate of the recognition test became worse as the number of measures increased regardless of the degree of NFC. This result indicates that smaller numbers of measures of these two perspectives are used as the number of provided measures increases.

This study indicates that more than seven measures causes information overload. However, in the case of seven measures provided, managers with high NFC try to use both financial and added non-financial measures (measures in customer perspective only). This result is the same as the findings of Ittner et al. (2003), who show that the most frequently used perspective is the financial perspective, followed by the customer perspective. On the other hand, in the case of ten measures provided, managers, even those with high NFC, can no longer use either financial or non-financial measures effectively.

The new findings of this study, which successfully build on the results of Ittner et al. (2003), are as follows. This study indicates that managers do not necessarily use all of the provided measures when an excess number of measures is provided, even though the measures belong to the customer perspective. Therefore, if more measures are provided in the company studied by Ittner et al. (2003), their findings might be different. Furthermore, this study shows the possibility that managers with low NFC do not use measures of customer perspective even when a small number of measures of customer perspective is used in addition to those of the financial perspective. In other words, we should pay careful attention to differences in NFC when we consider which measures managers use.

Our study has some limitations. First, we did not include monetary incentive in our experiment and thus, the number of measures that leads managers to use either financial or customer measures is different from that indicated by Ittner et al. (2003). Even though the company analyzed by Ittner et al. (2003) used more than 25 measures, managers still used financial and customer measures more than the other measures. Our results show that managers gave up using customer measures (non-financial measures) if the total number of measures exceeded 10. We consider the reason for the difference between our result and that of Ittner et al. (2003) to be monetary incentives. Monetary incentives can decrease information load (Ding and Beaulieu, 2011; Jackson and Farzaneh, 2012). Therefore, in the case of monetary incentives prepared, the number of measures that lead managers to focus on only financial and customer measures is larger than the case without it. If we had included monetary incentives in our experiment, we might have reached the same number of measures as Ittner et al. (2003). Consideration of monetary incentives is left as a topic for future research. Second, our participants

are employees of a single company, which might have biased the results. Using managers of multiple companies or MBA students is another direction for future study.

Acknowledgement

This work was supported by JSPS KAKENHI Grant Number 23530567 and 26380598.

Helpful comments and suggestions from seminar participants of the European Accounting Association 37th Annual Congress are gratefully acknowledged.

References

Banker, R.D., Chang, H. and Pizzini, M.J. (2004), "The balanced scorecard: Judgmental effects of performance measures linked to strategy", *The Accounting Review*, Vol.79 No.1, pp.1–23.

Banker R.D., Chang, H. and Pizzini, M. (2011), "The judgmental effects of strategy maps in balanced scorecard performance evaluations", *International Journal of Accounting Information Systems*, Vol.12, pp.259–279.

Cacioppo, J. T. and Petty, R.E. (1982), "The need for cognition", *Journal of Personality and Social Psychology*, Vol.42, pp.116–131.

Cardinals, E. and van Veen-Dirks, P.M.G. (2010), "Financial versus non-financial information: The impact of information organization and presentation in a Balanced Scorecard", *Accounting, Organizations and Society*, Vol.35, pp.565–578.

Cohen, J., and Cohen, P. (1983), *Applied multiple regression/correlation analysis for the behavioral sciences* (2nd ed.), Hillsdale, NJ: Erlbaum.

Ding, S. and Beaulieu, P. (2011), "The role of financial incentives in balanced scorecard-based performance evaluations: Correcting mood congruency biases", *Journal of Accounting Research*, Vol.49 No.5, pp.1223–1247.

Drolet, A., Luce, M.F. and Simonson, I. (2009), "When does choice reveal preference? Moderators of heuristic versus goal-based choice", *Journal of Consumer Research*, Vol.36, pp.137–147.

Eppler, M.J. and Mengis, J. (2004), "The concept of information overload: A review of literature from organization science, accounting, marketing, MIS, and related disciplines", *The*

Information Society, Vol.20, pp.325–344.

Farrell, A.M., Luft, J., and Shields, M. (2007), "Accuracy in judging the nonlinear effects of cost and profit drivers", *Contemporary Accounting Research*, Vol.24 No.4, pp.1139–1169.

Feltham, G.A. and Xie, J. (1994), "Performance measure congruity and diversity in multi-task principal-agent relations", *The Accounting Review*, Vol.69 No.3, pp.429–453.

Holmström, B. (1979), "Moral hazard and observability", *The Bell Journal of Economics*, Vol.10 No.1, pp.74–91.

Holmström, B. and Milgrom, P. (1991), "Multitask principal-agent analyses: Incentive contracts, asset ownership, and job design", *Journal of Law, Economics, and Organization*, Vol.7, pp.24–52.

Humphreys, K.A. and Trotman, K.T. (2011), "The Balanced Scorecard: The effect of strategy information on performance evaluation judgments", *Journal of Managerial Accounting Research*, Vol.23, pp.81–98.

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*, Vol.78 No.3, pp.725–758.

Jackson. T.W. and Farzaneh, P. (2012), "Theory-based model of factors affecting information overload", *International Journal of Information Management*, Vol.32, pp.523–532.

Kaplan, R.S. and Norton, D.P. (1996), *The Balanced Scorecard: Translating strategy into action*, Harvard Business School Press, Boston, MA.

Kaplan, R.S. and Norton, D.P. (2001), *The strategy focused organization: How Balanced Scorecard companies thrive in the new business environment*. Harvard Business School Press, Boston, MA.

Kaplan, R.S. and Norton, D.P. (2004), *Strategy maps: Converting intangible assets into tangible outcomes*. Harvard Business School Press, Boston, MA.

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*, Vol.79 No.4, pp.1075–1094.

Lipe, M.G. and Salterio, S.H. (2000), "The Balanced Scorecard: Judgmental effects of common and unique performance measures", *The Accounting Review*, Vo.75 No.3, pp.283–298.

Luft, J.L. and Shields, M.D. (2001), "Why does fixation persist? Experimental evidence on the judgment performance effects of expensing intangibles", *The Accounting Review*, Vol.76 No.4, pp.561–587.

Malina, M.A. and Selto, F.H. (2001), "Communicating and controlling strategy: An empirical study of the effectiveness of the Balanced Scorecard", *Journal of Managerial Accounting Research*, Vol.13, pp.47–90.

Milgrom, P. and Roberts, J. (1992), *Economics, organization and management*, Prentice Hall, Upper Saddle River, NJ.

Miller, G.A. (1956), "The magic number seven plus or minus two: Some limits on our capacity to process information", *Psychological Review*, Vol.63 No.2, pp.81–97.

Swain, M.R. and Haka, S.F. (2000), "Effects of information load on capital budgeting decisions", *Behavioral Research in Accounting*, Vol.12, pp.171–198.

Tuttle, B. and Burton, F.G. (1999), "The effect of a modest incentive on information overload in an investment analysis task", *Accounting, Organizations and Society*, Vol.24 No.8, pp.673–687.