

Determinants of the Subjective Performance
Measurement of Managerial Behavior*

NEALE G. O'CONNOR, *University of Hong Kong*

F. JOHNNY DENG, *California State University, Sacramento*

MICHAEL D. SHIELDS, *Michigan State University*

October 2006

* The authors would like to thank Ranjani Krishnan, Joan Luft, and Karen Sedatole for their helpful comments on previous versions of this paper.

Determinants of the Subjective Performance Measurement of Managerial Behavior

Abstract

Accounting research provides theory and evidence on the choice and use of objective performance measures in objective (formula) and subjective performance evaluation and reward systems. We extend the accounting research by studying determinants of the subjective measurement of managers' behavior using proprietary objective and subjective performance measurement data for 26 branches of a telecommunications firm over five years. Based on psychology research, we predict and find that objective financial and nonfinancial performance measures and immediate prior-period subjective performance measurements interact with the geographical distance between the branch and home offices to affect the current-period subjective measurement by home-office managers of branch managers' "integration" behavior. The psychology literature provides an explanation for this result, that geographical distance affects how other performance measures influence subjective performance measurement because of the measurer's cognitive limitations, in particular the measurer's limited information about the behavior being subjectively measured and his or her limited time and effort to make the subjective measurement.

Keywords: Performance measurement; Subjective performance measures; Performance appraisal; Rating theory.

Data Availability: The confidentiality agreement with the firm that provided data precludes the revelation of its identity and the dissemination of the data.

1. Introduction

Accounting research provides analysis of and evidence on costs and benefits of the choice and use of performance measures in performance evaluation and reward systems (Ittner and Larcker 1998, 2001). These performance measures are inputs to performance evaluation and reward systems that have objective (e.g., formula) or subjective mappings from the measures to the incentive (Baiman and Rajan 1995). Much of this research focuses on *objective* financial and nonfinancial performance measures, which are quantitative and verifiable. Recent research, however, has begun to provide theory-based evidence on the role that *subjective* performance measures play in (objective or subjective) performance evaluation and reward systems. Ittner, Larcker and Meyer (2003) investigate how subjective measures influence subjective overall performance evaluation ratings and annual bonuses of managers. Moers (2005) examines how subjective performance measurements bias performance ratings and bonuses in objective performance evaluation and reward systems. What is still unclear, however, is what factors influence subjective performance measurement. This is important because subjective performance measures are inherently qualitative, non-verifiable, and potentially influenced by the biases and cognitive limitations of the measurer.

We extend the accounting research by studying determinants of the subjective measurement of managers' behavior using proprietary objective and subjective performance measurement data for up to 26 (depending on the year) branch offices of a telecommunications firm over five years. The firm has an objective performance evaluation and reward system in which a branch-office manager's annual bonus is determined by four objective measures (sales, expenses, accounts receivable, customer

satisfaction) and one subjective measure of what the firm calls “integration” behavior. Integration behavior is related to a branch-office manager’s contribution to factors not captured in the objective measures such as team spirit, job attitude, and creativity.

This paper provides theory-based evidence on determinants of home-office manager’s (hereafter, supervisor [or rater]) subjective performance measurement of a branch-office manager’s (hereafter, subordinate [or ratee]) integration behavior. Psychology research predicts that as geographical distance increases, the opportunity and the time available for learning about subordinate behavior decreases, which results in more reliance on the informational cues of accessible performance measurements. Based on this research, we predict that objective financial and nonfinancial performance measures and immediate prior-period subjective performance measurements will interact with the geographical distance between the workplaces of the supervisor and subordinate to affect the current-period subjective performance measurement of the subordinate’s integration behavior. In particular, the objective performance measures are a branch’s current-period sales, expenses, accounts receivable, and customer satisfaction, and the subjective performance measure is the immediate prior-period subjective performance measurement of the subordinate’s integration behavior. We find empirical support for our predictions and conclude that cognitive limitations explain, at least in part, the influence of objective and immediate prior-period subjective performance measurements on current-period subjective performance measurement of integration behavior.

What is interesting about our prediction and result is not that subjective performance measurements are influenced by other performance measures, but that the influence of the other measures is conditional on the geographical distance between the

branch and home offices. These performance measures' influence on the subjective measurement of integration behavior would be unrelated to geographical distance *if* their influence on the subjective measurement of integration behavior derived only from their information content about that behavior. The psychology literature assumes that geographical distance affects how other measures influence a supervisor's subjective performance measurement because of the supervisor's cognitive limitations, in particular the supervisor's limited information about the subordinate's integration behavior and his or her limited time and effort to make the subjective measurement.

The remainder of this paper is organized as follows. In Section 2 we review accounting literature on performance measurement and evaluation and then psychology literature on the subjective performance measurement of work-related behavior and variables that are expected to influence subjective performance measurement as the basis to develop two interaction hypotheses. In Section 3 we describe the sample, measurement of the dependent, independent, and control variables, and the empirical models used to test the hypotheses. In Section 4 we report the results of the hypothesis testing, and in Section 5 we discuss implications, limitations, and conclusions of our paper.

2. Literature review and hypothesis development

Performance measurement and evaluation

Accounting research investigates the choice and use of performance measures to motivate and reward employees (Ittner and Larcker 1998, 2001). This research provides theory and evidence on the use of objective performance measures such as net income and economic value added that are inputs to performance evaluation and reward systems. These systems can be objective, in which formulas map measured performance to

bonuses by incentive weights on performance measures (Ittner, Larcker and Rajan 1997; Fisher et al. 2005), or subjective, in which the mapping is by judgment (Baiman and Rajan 1995; Gibbs et al. 2004; Krishnan, Luft and Shields 2005).

Some recent empirical studies extend this accounting research by examining not only objective measures, but also subjective measures in the context of subjective performance evaluation and reward systems. Subjective measures can be valuable for including in performance evaluation and reward systems because they can be informative about qualitative aspects of manager's job and behavior beyond the information provided by objective performance measures. Ittner et al. (2003) provide evidence that psychology theory may provide more valid explanations of how subjective measures influence managers' performance evaluation ratings and annual bonuses as compared to economics theory. Moers (2005) uses psychology theory to predict how subjective measurement of managers' behavior biases performance ratings that determine annual bonuses. In particular, subjective performance measurement results in performance ratings which are lenient (higher) and compressed (closer to the median rating) compared to objective performance measures.

The prior accounting research examines effects of subjective performance measurement on performance evaluation. We extend this accounting research on subjective performance measurement by providing theory and evidence on factors that influence the subjective performance measurement of managerial behavior. We use psychology theory because it has extensive analysis of (and evidence on) factors that influence the subjective measurement of qualitative aspects of work-related behavior.

Psychology theory provides several reasons why subjective performance measurement may differ from an objective performance measurement of managerial behavior.¹ Subjective performance measurement can have cognitive and social biases. Cognitive biases can be due to incomplete or biased information accessible to the supervisor for subjective processing, for example due to not observing all aspects of the subordinate's behavior, biased or incomplete recall of relevant information in memory, or incorrect classification (categorization) of a behavior as being a particular type (e.g., cooperation, innovation, motivation) or level (e.g., poor, excellent). Social costs of subjective performance measurement can be due to the social context or relationship between the supervisor and subordinate (e.g., favoritism).

In the psychology literature (see note 1), subjective measurement of performance is called "performance appraisal," the person who is doing the subjective measurement is called the "rater," the person whose behavior is being subjectively measured is called the "ratee," and the subjective measurement is called the "rating." Performance appraisal requires the rater to make a subjective judgment about how well the ratee is performing on one or more qualitative types of behavior, such as attitude, communication, interpersonal skills, leadership, and motivation. Almost all psychology research on performance appraisal experimentally investigates how a rater rates the behavior of a ratee in relation to prespecified performance criteria for various types of behavior.

Subjective performance measurement

Several psychology models have been developed concerning subjective performance measurement and factors that affect it and its accuracy. The initial research assumes that rating inaccuracy is due to unintentional errors and thus it focuses on

sources of error and how to reduce them. The first rating model, based on concepts from cognitive psychology and psychometrics, identifies sources of rating inaccuracy due to these errors, in particular actual ratee behavior, rater observation of ratee behavior, raters' recollection of past observations of ratee behavior, and characteristics of the organizational setting of the rating, the rater-ratee relation, and the measurement scale used for the rating (Wherry and Bartlett 1982). DeNisi et al. (1984) extend this model by elaborating on cognitive processes that occur when raters make ratings, including memory processes for the encoding, storage, retrieval, and integration of information about a ratee's behavior to form a judgment about the rating. This model assumes that raters minimize their limited cognitive resources when making ratings by using heuristic information processing, thus their ratings can be unintentionally biased. More recent research based on social cognition, however, assumes that rating inaccuracy is due to raters' conscious decisions arising from the social context of the rating (e.g., relation between rater and ratee) (Cleveland and Murphy 1992; Fletcher 2001; Ilgen et al. 1993).

Hypotheses development

We develop two hypotheses about how objective financial and nonfinancial performance measures and immediate prior-period subjective performance measurement are expected to interact with the geographical distance between the branch and home offices to affect the current-period subjective measurement by supervisors of subordinates' integration behavior.

Performance measures

Our firm evaluates the annual performance of its subordinates using three objective financial performance measures (sales, expenses, and accounts receivable), an

objective nonfinancial performance measure (customer satisfaction), and a subjective performance measure (integration). Integration is intended by supervisors to measure subordinates' behavior that is important to the firm's success but not directly or completely captured by the objective performance measures. Although supervisors consider the four objective performance measures to be related to different behavior by the subordinates than integration behavior which they subjectively measure, prior psychology research provides two explanations for how objective measures can be expected to influence the subjective performance measurement of integration behavior.

The first explanation assumes that supervisors in our firm believe they lack sufficient information for rating subordinates' integration behavior and/or desire additional information on which to base their ratings. In such situations, DeNisi et al. (1984) predict that raters will put more weight on knowledge that is easier to retrieve from their memories, which is consistent with contemporary research on how judgments are influenced by the accessibility of knowledge in memory (Kahneman 2003). In particular, DeNisi et al. (1984) predict that, based on research in cognitive psychology, ratings will be influenced by knowledge in raters' memory that is easily or automatically accessed and that raters believe is causally related or similar to the ratings. It is likely that raters will believe that objective performance measures are causally related or similar to behavior that is being subjectively measured. For example, when raters do not have a lot of information about rates' integration behavior from direct observation, their ratings of integration behavior are likely to be influenced by their knowledge that is readily accessible from memory and/or information in reports that they believe are somehow related (causal or similarity) to integration behavior. Because integration includes many

types of behavior, it is costly for supervisors to have timely and complete access to information about subordinates' integration behavior. Thus, raters may believe that measures of sales, expenses, accounts receivable, and customer satisfaction are informative about ratees' integration behavior.

The second explanation is based on research that predicts and finds that ratings of a particular behavior are influenced by ratings of other types of behavior that are statistically unrelated to the focal behavior and that ratings of different behaviors are more highly correlated than these behaviors are actually correlated. This halo effect is a cognitive bias in which ratings of a particular type of behavior influences (biases) ratings of unrelated types of behaviors (Balzer and Sulsky 1992; Nunnally and Bernstein 1994).²

Prior ratings

Ratings of managerial behavior can also be influenced by prior ratings of this behavior. DeNisi et al. (1984) use concepts from social cognition to argue that a rater's prior impressions about a ratee, for example as represented by prior ratings, can affect the rater's current rating of the ratee. Prior impressions can affect what knowledge is retrieved from memory (e.g., prior ratings) and/or information acquired from reports and then used to make the current rating. In particular, prior impressions can form a mental schema of the ratee, which is likely to consist of an impression of the ratee (e.g., "this manager has had excellent integration behavior"). Psychology research (e.g., Balzer 1986; Murphy et al. 1985) and more recently accounting research (e.g., Bhattacharjee and Moreno, 2004) provides evidence that prior impressions—in particular whether they are favorable or unfavorable—can unintentionally affect a rater's cognitive processing of objectively measured performance. This happens because prior impressions either induce

raters to retrieve knowledge from their memory and/or acquire information that is consistent with their prior impressions or cause the rater to change the criteria that he or she uses to judge (rate) current behavior (Baltes and Parker 2000; Martell and Willis 1993). Because prior ratings are expected to influence prior impressions and prior impressions are expected to influence current ratings, we expect that prior ratings of a particular type of behavior will influence current ratings of that behavior.

While the explanations above suggest that the raters' ratings of integration behavior will be influenced by objective performance measures of the subordinates' behavior and immediate prior-period ratings of subordinates' integration behavior, the psychology literature in the next section supports our prediction that these relationships will be stronger for subordinates who are more geographically distant from the head office.

Geographical distance

Wherry and Bartlett (1982) predict, based on psychometric theory, that the greater the geographic distance between the rater and the ratee's workplaces, the lower the probability of multiple direct experiences in which the rater has opportunity to learn about ratee behavior by observing or interacting with the ratee and thus the lower the probability of an accurate rating.³ As the number of informative experiences decreases, raters increasingly use other information to make their rating. In such situations, this other information is likely to include information about the ratee that is impounded in accessible performance measures, such as branch sales and expenses.

Related, DeNisi et al. (1984) also predict that raters who are under time pressure when making a rating will consider less new information and be more likely to rely on

their prior ratings. When applied here, this implies that to minimize the time and cognitive effort spent making a rating, it is likely that the rater will have more information about the integration behavior of subordinates who are geographically closer to the head office than those who are more geographically distant because of the cost of obtaining integration information. For example, for our firm visits by supervisors to the more distant branch offices are less frequent and on each visit there is less time for informative experiences with the subordinate, regional manager, and customers.

Based on the literature on the effects of geographic distance on ratings, we predict that the extent to which objective performance measures influence integration ratings will increase with the geographic distance between branch and home offices. Therefore, we predict the following:

Hypothesis 1. *The influence of objective performance measurement of sales, expenses, accounts receivable, customer satisfaction on a rating of integration behavior increases with geographical distance.*

Hypothesis 2. *The influence of an immediate prior-period rating of integration behavior on a current-period rating of integration behavior increases with geographical distance.*

If the influence of objective measures and immediate prior-period subjective measures on current-period subjective performance measurement of integration behavior derive only from the information content regarding integration behavior in these other measures, then the influence of these measures on the subjective measurement would be unrelated to the geographical distance between the branch and home offices. Thus, evidence in support of our hypotheses suggests that the cognitive limitations described above explain, at least in part, the influence of objective measures and immediate prior-period subjective measures on current-period subjective performance measurement of integration behavior.

3. Research Method

Research site

CTI was a for-profit firm founded in 1994 that quickly became a major producer and supplier of telecommunications equipment in China. It supplied equipment to two of the three Chinese state-owned telecommunications enterprises through over 200 regional subsidiaries of the two enterprises. In 1998 CTI had nine branch offices located in nine provinces and in 2003 it had 26 branch offices in 26 provinces (including Beijing and Shanghai). The subordinates (branch-office managers) were responsible for advertising and promoting products, managing branch employees including customer service personnel, sales and accounts, and the physical facilities of the branch.

An objective (formula) performance evaluation plan was used to provide annual bonuses to subordinates. The formula had the bonus as the dependent variable and the independent variables were four objective performance measures (each branch-office's sales, expenses, accounts receivable, customer satisfaction) and a subjective performance measure called integration. The integration performance measure was adopted for several reasons such as the need to manage controllability, encourage communication, teamwork and managerial skill development, and to discourage the non-sharing of private information and a short-term myopic focus on sales and other objective measures (see Appendix). The supervisor (CTI's general manager of sales) periodically visited each branch office unannounced to inspect the branch office for the purpose of making the integration rating. In 2001, no rating of integration behavior was made because of a change in the measurement of integration behavior (described below) and the timing of

the departure of the existing and the arrival of the new supervisor, who made these ratings of each subordinate.

Variables

Annual data were obtained for five years (1998–2000 and 2002–2003) for nine (1998), ten (1999), 17 (2000), and 26 (2002 and 2003) branch offices of CTI. Sales, accounts receivable, and expense data were collected from the accounting system at CTI's head office. Customer satisfaction and integration data were collected from the performance-management database at CTI's head office.

Integration

Integration (*INTEGRATION*) was measured by reference to CTI's rating scales for integration, which for 1998-2000 had a theoretical range from 0 to 20 points and after 2001 when a new performance evaluation plan was introduced the scale was changed to have a theoretical range from 0.8 to 1.2 points. Notwithstanding this change in scaling, the anchors for the low and high ends of both scales remained unchanged. Because of the different scales across the two time periods, we normalized the rating data to have a theoretical range from 0.80 to 1.20; the actual range was 0.89 to 1.20. The normalized measure was continuous with ratings (measured values) at intervals of a hundredth of a point.

The rating was based on judgments across two areas of integration behavior that supervisors desired subordinates to have. The first area comprised six types of subordinate behavior: (a) contribution to team spirit both within and across branch offices, (b) job attitude, (c) effort in undertaking temporary tasks, (d) working with regional managers on large customers and province-based clients, (e) working with the home

office or other branches, and (f) effort put into sales and promotion activities, on-time monthly and seasonal reporting, and keeping daily working logs. The second area related to the pro-activeness and creativity of the subordinates and their employees towards work.

Objective performance measurement

Sales (*SALES*) was the net sales revenue of products sold. Expenses (*EXPENSE*) included the total direct costs of operating the branch office, including salaries and wages, office rent, water, electricity, gas, telephone, office equipment, banquets, entertainment, and business trips. Accounts receivable (*ACCREC*) was defined as total accounts receivable at the end of a reporting period divided by mean credit sales per day. This definition meant that higher *ACCREC* represented lower performance in terms of managing credit sales.

Customer satisfaction (*CUSTOMER*) measure was obtained annually based on a two-page survey that was returned directly by customers to the head office. The survey had 20 items that were intended to measure each customer's satisfaction in eight areas: product quality and prices (5 questions), customer hotline service (3), on-site customer service (5), product supply and delivery ability (1), after-sales product maintenance service (3), customer staff training and development (2), and overall rating (1).

Responses to each of the 20 items were captured on a three-point Likert scale, anchored by 1= Dissatisfied, 2= Satisfied and 3= Very Satisfied. The responses to these 20 items were summed to form the customer satisfaction measure that ranged from 20 to 60.

Geographical distance

The distance (*DISTANCE*) between a branch sales office and the head office was measured as average flight time in hours taken from published information by airlines.

Since the flight time varied by airline, for branch offices with multiple airlines, an average time across airlines was computed. Measuring *DISTANCE* by flight time was consistent with how the supervisors and subordinates traveled between the head office and each branch. It also was consistent with studies on cognitive representations of distance, which indicated that individuals' cognitive representation of distance was influenced most by travel time (MacEachren 1980). Related, Freundsuh (1994) found that people tended to use travel time in lieu of physical distance to estimate distances between 331 kilometers and 2,080 kilometers, which was similar to the range of distances (320 kilometers to 2,766 kilometers) between the head office and branch offices in our firm.

Control variables

We included variables in the hypothesis tests to control for the potential effects of variables other than the independent variables on the dependent variable. Characteristics of a branch or its subordinate could influence the supervisor's ratings of a subordinates' integration behavior. Considering subordinates, we controlled for *TENURE*, measured as the number of months that a subordinate had worked for the firm, because subordinates with more tenure might be better known to supervisors, which could affect integration ratings. Alternatively, supervisors might focus more attention on newer subordinates to learn about them.

Considering a branch, its age (*AGE*), measured as years in operation, or size (*SIZE*), measured as number of branch employees, can influence integration ratings. For example, a supervisor could have more information about older branches because they have existed longer which provided the rater with more opportunity to learn more about

those branches and thus the subordinate's integration behavior. Alternatively, a supervisor might attend more to newer branches in order to learn about them and monitor their subordinates' integration behavior. The supervisor might know more about subordinates' integration behavior for larger branches because of their importance to the firm. Similarly, branches located in provinces with higher annual growth in sales of telecommunications terminal equipment (*PROVGWTH*) could be more important to the future success of the firm and thus the supervisor learns more about subordinates at those locations. Alternatively, the supervisor might focus more attention on branches located in lower revenue growth provinces because these branches could be risky to the firm's future and thus he or she learns more about those subordinates' integration behavior. Finally, we included a control variable to control for any effects on integration ratings due to the change in the rating system and supervisors.

Estimation models

We developed several regression models for the branch data that existed during the five years (1998–2000 and 2002–2003) with each model having 88 branch-years consisting of the nine, ten, and 17 branches that existed in 1998, 1999 and 2000, respectively, and the 26 branch offices that existed in 2002 and 2003. Each model had a dummy for the two rating systems and supervisors (1998-2000 and 2002-2003), none of which were significant ($p < 0.20$) in any of the regressions). The regression model, excluding the rating system and supervisor dummy variable, for this data was:

$$\begin{aligned}
 INTEGRATION_{it} = & B_0 + B_1 SALES_{it} + B_2 EXPENSE_{it} + B_3 ACCREC_{it} + B_4 CUSTOMER_{it} \\
 & + B_5 INTEGRATION_{it-1} + B_6 DISTANCE_{it} + B_7 TENURE_{it} + B_8 AGE_{it} \\
 & + B_9 SIZE_{it} + B_{10} PROVGWTH_{it} + B_{11} D*S + B_{12} D*E + B_{13} D*A
 \end{aligned}$$

$$+ B_{14} D * C + B_{15} D * I_{t-1} + \varepsilon_{it}$$

where, for branch i in period t :

$INTEGRATION_{it}$ = supervisor's subjective performance measurement of a subordinate's integration behavior,

$SALES_{it}$ (S) = sales in millions of Yuan,

$EXPENSE_{it}$ (E) = total branch direct expenses in millions of Yuan,

$ACCREC_{it}$ (A) = end-of-period net accounts receivable / mean credit sales per day,

$CUSTOMER_{it}$ (C) = customer satisfaction,

$DISTANCE_{it}$ (D) = distance between branch and head offices in average air travel time in hours,

$TENURE_{it}$ = months that a subordinate had been employed by CTI,

AGE_{it} = age of the branch in months,

$SIZE_{it}$ = number of employees working at a branch at the beginning of each time period, and

$PROVGWTH_{it}$ = annual percentage growth in sales of telecommunications terminal equipment in the province in which a branch was located.

Given the multi-period and multi-branch nature of our panel data, we employed Huber-White robust standard errors with the cluster command to control for the sample dependence around each branch office (Rogers 1993). This estimation procedure assumed and estimated a common component of the variance-covariance matrix for all observations from the same branch, and the standard errors were robust to heteroskedasticity and serial correlation (StataCorp 1999, 257). In order to minimize the bias created by producing interactions from variables with different scaling and ranges,

we first standardized the independent variables before multiplying them to create the interaction terms. The raw scores of the variables were used in the regression models for estimating non-interaction effects. Three influential observations were identified with *DFITS* and Cook's D-statistic and dropped from the analysis.

4. Results

Descriptive analysis

Table 1 presents descriptive statistics for the data. For these measured variables, there was considerable variance in the data as assessed by the variables' standard deviations and ranges relative to their means. The average sales per branch was 12.81 million Yuan, the average expenses per branch were 0.47 million Yuan, and the average number of day's credit sales in accounts receivables was 124.29.⁴ The average customer satisfaction was 45.44. The average distance of the branches from the home office was 1.52 hours of air travel time. The average age of the branch offices was 35.01 months, the average branch office size was 4.70 employees, and the annual sales growth in provincial sales of telecomm equipment was 32%. The average subordinate tenure was 47.15 months.

Insert Table 1 about here

Table 2 presents Pearson coefficients for the variables for the data. *INTEGRATION* had significant ($p < .15$) positive correlations with *ACCREC*, *CUSTOMER*, *AGE*, and *SIZE*. Four correlations between independent variables were significant at $p < .05$. The correlations between *SALES* with *EXPENSE*, *INTEGRATION*_{*t-1*} and *DISTANCE* were significant as was the correlation between *EXPENSE* and

*INTEGRATION*_{*t-1*}. Finally, among the control variables, the correlation between *AGE* and *TENURE* was significant at $p < .05$.

Insert Table 2 about here

Hypotheses tests

We report two regression models which varied in terms of the variables included: Model 1 had additive and interactive independent variables, and model 2 had additive and interaction independent variables as well as the control variables. Both regressions were significant ($p < 0.05$), with adjusted R^2 s of 0.34 and 0.41. Multicollinearity was not a problem as evidenced by the largest variance inflation factor in either regression model of 5.01 and the largest condition index value of 33.94, which were within the accepted limits (Belsley 1991) (Table 3). The residuals in both models were normally distributed.

Insert Table 3 about here

Hypothesis 1 predicted that the influence of objective performance measurement on *INTEGRATION* would increase with *DISTANCE*. The *SALES* by *DISTANCE* interaction was significant with the predicted positive sign in both models: models 1 ($p < .05$) and 2 ($p < .01$). The *EXPENSE* by *DISTANCE* interaction was significant ($p < .01$) with the predicted negative sign in models 1 and 2. The *ACCREC* by *DISTANCE* interaction was significant ($p < .01$) with the predicted negative sign in models 1 and 2. The *CUSTOMER* by *DISTANCE* interaction was marginally significant ($p < .10$) in both

models, but in the opposite direction to that predicted. Overall, these results provided support for hypothesis 1.

Hypothesis 2 predicted that the influence of $INTEGRATION_{t-1}$ on $INTEGRATION_t$ would increase with geographical distance. That is, $DISTANCE$ would have a positive interaction with $INTEGRATION_{t-1}$. The coefficient for $INTEGRATION_{t-1}$ was positive and significant ($p < .01$) in both models. These results provided support for hypothesis 2.

We tested the sensitivity of our results in three ways First, the hypothesis testing results were invariant to whether the four control variables were included (Model 1 vs. Model 2). Second, we performed four additional regressions, each with the additive and interaction independent variables in Model 2 and one of the control variables in Model 2. The hypothesis-testing results were robust to which control variables were included in these models. Finally, the rating system dummy variable was not significant ($p < .05$) in any of the previously presented models, and the hypothesis-testing results were robust to whether the rating dummy variable was included in Models 1 and 2.

5. Discussion and conclusion

Accounting research investigates benefits and costs of the choice and use of objective and subjective performance measures in objective (formula) and subjective performance evaluation and reward systems. Most of this research does not include subjective performance measures and the few studies that do focus on their benefits (Baiman and Rajan 1995) or on their costs arising from using subjective performance measures in subjective (Ittner et al. 2003) or objective (Moers 2005) performance evaluation and reward systems. We extend this accounting research by providing theory

and evidence on determinants of subjective performance measurement of managerial behavior.

Our prediction and result is consistent with psychology research that assumes that the effects of information (e.g., performance measures) on a subjective performance measurement are influenced by factors other than the information content of that information. In particular, we predict and find evidence that the influence of other performance measures on a subjective performance measurement depends on the geographical distance between the supervisor and subordinate's workplaces. These other performance measures have an increasing effect on the subjective measurement as geographical distance increases because of cognitive biases. These cognitive biases can be due to incomplete or biased information accessible to the supervisor, for example from not directly observing all aspects of the subordinate's behavior, biased or incomplete recall of relevant information from memory, or incorrect (categorization) of a behavior as being a particular type or level.

This research study has several limitations that warrant acknowledgement. The main limitations of this study are small sample size and data from one firm which reduces the generalizability of our results to other settings. Other limitations include omitted variables and measurement error. For example, geographic distance could be measured as the number or length of office visits, which might be better proxies for geographical distance; however, we do not have access to this data.

In spite of these potential limitations, our paper opens up a new area for accounting research on performance measurement. While we investigate cognitive limitations of the determinants of subjective performance measurement, the psychology

(see note 1) and economics (Prendergast and Topel 1993, 1996) literatures worry about potential costs of subjective performance measurement that arise from social factors such as favoritism and fairness. These social costs can affect determinants of subjective performance measurement as well as how these measurements are used, for example in performance evaluation and rewards. As accounting continues to expand its scope of performance measurement to include not only objective financial and nonfinancial performance measures but also subjective performance measures, research on the benefits and costs of the subjective performance measurement of managerial behavior such as cooperation, flexibility, and innovation is increasingly valuable.

TABLE 1
Descriptive statistics (N=85)

	Mean	Standard Deviation	Maximum	Minimum
<i>INTEGRATION</i>	1.02	0.06	1.20	0.89
<i>SALES</i>	12.81	11.91	66.86	0.01
<i>EXPENSE</i>	0.47	0.52	3.39	0.01
<i>ACCREC</i>	124.29	74.82	331.67	17.61
<i>CUSTOMER</i>	45.44	5.78	59.60	28.85
<i>DISTANCE</i>	1.52	0.85	4.17	0
<i>TENURE</i>	47.15	22.24	93.00	12.00
<i>AGE</i>	35.01	19.35	72.00	12.00
<i>SIZE</i>	4.70	1.84	9.00	1.00
<i>PROVGWTH</i>	32.00	130.00	1101.00	-98.90

Note:

Variable measurements:

- i* = 1-26 represents the individual branch offices,
- INTEGRATION* = supervisor's subjective measurement of a subordinate's integration behavior,
- SALES* = sales in millions of Yuan,
- EXPENSE* = total branch direct expenses in millions of Yuan,
- ACCREC* = end-of-period net accounts receivable / mean credit sales per day,
- CUSTOMER* = customer satisfaction,
- DISTANCE* = distance between branch and head office in air travel time in hours,
- TENURE* = months a subordinate has been employed by CTI,
- AGE* = age of the branch office in months,
- SIZE* = number of employees working at a branch at the beginning of each time period, and
- PROVGWTH* = annual percentage growth in sales of telecommunications terminal equipment in the province in which a branch was located.

TABLE 2
Pearson correlations ($N=85$)

	<i>INTEGRATION</i>	<i>SALES</i>	<i>EXPENSE</i>	<i>ACCREC</i>	<i>CUSTOMER</i>	<i>INTEGRATION_{t-1}</i>	<i>DISTANCE</i>	<i>TENURE</i>	<i>AGE</i>	<i>SIZE</i>
<i>SALES</i>	.117									
<i>EXPENSE</i>	-.108	.532‡								
<i>ACCREC</i>	.232†	-.194*	-.137							
<i>CUSTOMER</i>	.302‡	.160	-.008	.031						
<i>INTEGRATION_{t-1}</i>	.027	.097†	.204†	-.103	-.071					
<i>DISTANCE</i>	.050	-.226†	-.199*	.052	.026	-.013				
<i>TENURE</i>	.002	.181*	.393‡	.188*	.022	-.097	-.079			
<i>AGE</i>	.164	.091	.323‡	.205*	-.080	-.014	-.065	.651‡		
<i>SIZE</i>	.209†	.458‡	.249†	-.224†	.117	.213†	-.045	.080	.088	
<i>PROVGWTH</i>	-.021	-.087	-.039	.276‡	.045	-.104	-.026	-.031	-.073	-.174

Notes:

Significance levels: * $p < 0.10$, † $p < 0.05$, and ‡ $p < 0.01$ (two-tail test).

See Table 1 for how the variables were measured.

TABLE 3
OLS Regressions for *INTEGRATION*

$$\begin{aligned}
 INTEGRATION_{it} = & B_0 + B_1 SALES_{it} + B_2 EXPENSE_{it} + B_3 ACCREC_{it} + B_4 \\
 & CUSTOMER_{it} + B_5 INTEGRATION_{it-1} + B_6 DISTANCE_{it} + B_7 \\
 & TENURE_{it} + B_8 AGE_{it} + B_9 SIZE_{it} + B_{10} PROVGWTH_{it} + B_{11} D*S + B_{12} \\
 & D*E + B_{13} D*A + B_{14} D*C + B_{15} D*I_{t-1} + \varepsilon_{it}
 \end{aligned}$$

Independent Variables	Hypotheses (predicted sign)	Model 1	Model 2
<i>SALES</i>		.002 (2.65)†	.002 (2.53)†
<i>EXPENSE</i>		-.047 (-3.28)‡	-.043 (-3.72)‡
<i>ACCREC</i>		.000 (1.08)	.000 (1.56)
<i>CUSTOMER</i>		.003 (3.26)‡	.003 (3.51)‡
<i>INTEGRATION_{t-1}</i>		.299 (2.19)†	.227 (1.72)*
<i>DISTANCE</i>		-.003 (-0.59)	-.001 (-0.27)
<i>DISTANCE x SALES</i>	H1+	.020 (2.81)†	.025 (3.84)‡
<i>DISTANCE x EXPENSE</i>	H1-	-.067 (-3.17)‡	-.063 (-3.29)‡
<i>DISTANCE x ACCREC</i>	H1-	-.024 (-3.73)‡	-.022 (-2.99)‡
<i>DISTANCE x CUSTOMER</i>	H1+	-.015 (-1.87)*	-.017 (-2.22)*
<i>DISTANCE x INTEGRATION_{t-1}</i>	H2+	.023 (4.04)‡	.022 (3.70)‡
<i>TENURE</i>			-.001 (-2.28)†
<i>AGE</i>			.001 (2.13)†
<i>SIZE</i>			.005 (1.14)
<i>PROVGWTH</i>			-.004 (-1.57)
<i>Adjusted R²</i>		0.34	0.41
<i>N =</i>		85	85
<i>Maximum VIF</i>		4.92	5.01
<i>Condition Index</i>		27.66	33.94

(Continued)

(Continued)

Notes:

- ¹ Significance levels: * $p < 0.10$, † $p < 0.05$, and ‡ $p < 0.01$ (two-tail test). The relevant t -statistics in brackets were computed using Huber-White robust standard errors.
 - ² Both models were estimated with a rater system dummy variable for 1998-2000 (= 0) and 2002-2003 (=1). For both models the betas for the dummy were not significant ($p < 0.05$).
 - ³ See Table 1 for how the variables were measured.
-

Appendix

CTI incorporated a measure of integration behavior into the incentive scheme for several reasons. A face-to-face interview was conducted with the Vice President of Finance in June 2004 that lasted about one hour and a phone interview was conducted with the General Manager of Sales in October 2006 that lasted about 30 minutes. The interviews were structured by two questions, as follows, which is then followed by some quotes.

1. Why did you include the subjective measure in the incentive plan? Give some examples.

(a) Manage controllability:

“To balance the external differences among different markets (some markets were good and others are bad). In other words, taking managers' controllability into consideration when evaluating the branch office's financial performance.”
(General Manager of Sales)

(b) Encourage communication and teamwork:

“To encourage team spirit among different branches, which the sales numbers would not be able to do.”
(General Manager of Sales)

“The general manager wants an element of control over the head-office relationship with the branch-office managers and the subjective element is the mechanism used for this purpose.”
(Financial Vice President)

(c) Encourage managerial skill development:

“To encourage managers to be proactive in improving not only their sales capabilities but also their management skills.”
(General Manager of Sales)

“We want people to be able to see everything. They should not just be able to make sales, but should be able to motivate people, organize things strategically, putting down their thoughts into paper, think logically, and express themselves in a coherent fashion. Thus we observe our subordinates to see who has this potential, who could be a leader in the future. Thus the subjective measure matters to the skill development in addition to motivate sales. The purpose is to broaden the skills of the manager into a well rounded team leader.”
(Financial Vice President)

2. What behavior did senior management want to avoid by including the subjective measure? Give some examples.

(a) Not sharing private information and poor communication skills:

“When the managers didn't reveal their private information about the local markets intentionally or unintentionally.”

(General Manager of Sales)

“Even when they did reveal, their communication skills were lacking. As a result, the head office couldn't have a clear picture about the market. This, in turn, increased the difficulty of scheduling and planning.”

(General Manager of Sales)

“The subjective evaluation is also influenced by the frequency of communication. Maybe this gentleman has a better relationship to communicate with his superior. The subordinate lets his superior know everything he did which is not reflected by the incentive formula, while the other gentleman who tends to be quieter and does not know how to communicate what he did which is not reflected by the formula. This latter gentleman receives a lower subjective evaluation. He should have got more as per the formula but he failed to communicate.”

(Financial Vice President)

(b) Short-term myopic behavior:

“Some managers only paid attention to the current year sales even if doing so might damage future sales. A typical example of this is that a manager picks a huge bonus because of the sales number while he left the next year with a lot of outstanding accounts receivable.”

(General Manager of Sales)

“To avoid these types of behaviors, head office had to implement the subjective measure to complement the sales measures.”

(General Manager of Sales).

References

- Arvey, R., and K. Murphy. 1998. Performance evaluation in work settings. *Annual Review of Psychology* 49: 141–168.
- Baiman, S., and M. Rajan. 1995. The informational advantages of discretionary bonus schemes. *The Accounting Review* 70 (4): 557–569.
- Baltes, B., and C. Parker. 2000. Reducing the effects of performance expectations on behavioral ratings. *Organizational Behavior and Human Decision Process* 82 (2): 237.
- Balzer, W. 1986. Biases in the recording of performance-related information: The effects of initial impression and centrality of the appraisal task. *Organizational Behavior and Human Decision Processes* 37 (3): 329-347.
- Balzer, W., and L. Sulsky. 1992. Halo and performance appraisal research: A critical examination. *Journal of Applied Psychology* 77 (6): 975-985.
- Belsley, D. A. 1991. *Conditioning diagnostics: Collinearity and weak sata in regression*. New York: J. Wiley.
- Bhattacharjee, S., and K. Moreno. 2004. Debiasing the impact of positive and negative prior impressions on managers' performance evaluation judgments. Paper presented at the 2004 Global Management Accounting Research Symposium, Michigan State University.
- Bretz, R., G. Milkovich, and W. Read. 1992. The current state of performance appraisal research and practice: Concerns, directions, and implications. *Journal of Management* 18 (2): 321–352.

- Cleveland, J., and K. Murphy. 1992. Analyzing performance appraisal as goal-directed behavior. In *Research in Personnel and Human Resources Management*, ed. G. Ferris, and K. Rowland, Vol. 10, 121–185. Greenwich, CT: JAI Press.
- DeNisi, A. 1997. *Cognitive approach to performance appraisal: A program of research*. London: Routledge.
- DeNisi A., P. Cafferty, and M. Meglino. 1984. A cognitive view of the performance appraisal process: A model and research propositions. *Organizational Behavior and Human Performance* 33 (3): 360–396.
- Feldman, D. 1981. Beyond attribution theory: Cognitive processes in performance appraisal. *Journal of Applied Psychology* 66 (2): 127–148.
- Fisher, J., L. Maines, S. Pfeffer, and G. Sprinkle. 2005. An experimental investigation of employer discretion in employee performance evaluation and compensation. *The Accounting Review* 80 (2): 563-583.
- Fletcher C. 2001. Performance appraisal and management: The developing research agenda. *Journal of Occupational and Organizational Psychology* 74 (4): 473-487.
- Freundschuh, S. 1994. Cognitive distance at various geographic scales. *Time in geographic space: Report on the specialist meeting of research initiative 10*, edited by Egenhofer, M., and R. Golledge. eds. 1994. National Center for Geographic Information and Analysis, Report 94-9.
- Gibbons, R., and M. Waldman. 1999. A theory of wage and promotion dynamics inside firms. *The Quarterly Journal of Economics* 114 (4): 1321–1358.

- Gibbs, M., K. Merchant, W. Van der Stede., and M. Vargus. 2004. Determinants and effects of subjectivity in incentives. *The Accounting Review* 79 (2): 409–436.
- Harris, M. 1994. Rater motivation in the performance appraisal context: A theoretical framework, *Journal of Management* 20 (4), 737-756.
- Hoffman, E., K. McCabe, and V. Smith. 1996. Social distance and other regarding behavior in dictator games. *American Economic Review* 86 (3): 653–660.
- Hoyt, W. 2000. Rater Bias in Psychological Research: When Is It a Problem and What Can We Do About It? *Psychology Methods* 5 (1), 64 - 86.
- Ilgén, D., J. Barnes-Farrell, and D. McKellin. 1993. Performance appraisal process research in the 1980s: what has it contributed to appraisals in use? *Organizational Behavior and Human Decision Processes* 54 (3): 321–368.
- Ilgén, D., and J. Feldman. 1983. Performance appraisal: A process focus. In *Research in Organizational Behavior* (Volume 5), eds. Cummings, L., and B. Staw. Greenwich, CT: JAI Press.
- Ittner, C., D. Larcker, and M. Rajan. 1997. The Choice of Performance Measures in Annual Bonus Contracts. *The Accounting Review*, 72, 231-255.
- Ittner, C., and D. Larcker. 1998. Innovations in performance measurement: Trends and research implications. *Journal of Managerial Accounting Research* 10: 205–238.
- Ittner, C., and D. Larcker. 2001. Assessing empirical research in managerial accounting: A value-based management perspective. *Journal of Accounting and Economics* 32 (1-3): 349–410.

- Ittner, C., D. Larcker, and M. Meyer. 2003. Subjectivity and the weighting of performance measures: Evidence from a balanced scorecard. *The Accounting Review* 78 (3): 725–758.
- Kahneman, D. 2003. A perspective on judgment and choice: Mapping bounded rationality. *American Psychologist* 58 (9): 697–720.
- Krishnan, R., J. Luft, and M. Shields. 2005. Effects of accounting-method choices on subjective performance-measure weighting: Experimental evidence on precision and error covariance. *The Accounting Review* 80 (4): 1163–1192.
- Landy, F., and J. Farr. 1980. Performance ratings. *Psychological Bulletin* 87: 72–107.
- MacEachren, A. 1980. Travel time as the basis of cognitive distance. *Professional Geographer* 32 (1): 30-36.
- Martell, R., and C. Willis. 1993. Effects of observers' performance expectations on behavior ratings of work groups: Memory or response bias? *Organizational Behavior and Human Decision Processes* 56 (1): 91-109.
- Moers F. 2005. Discretion and bias in performance evaluation: The impact of diversity and subjectivity. *Accounting, Organizations and Society* 30 (1): 67–80.
- Murphy, K., and J. Cleveland. 1995. *Understanding performance appraisal: Social, organizational, and goal-based perspectives*. Thousand Oaks, CA: Sage.
- Murphy, K., W. Balzer, M. Lockhart, and E. Eisenman. 1985. Effects of previous performance on evaluations of present performance. *Journal of Applied Psychology* 70 (1): 72-84.

- Nunnally, J., and I. Bernstein. 1994. *Psychometric theory*. Third Edition. New York: McGraw Hill.
- Prendergast, C., and R. Topel. 1993. Discretion and bias in performance evaluation. *European Accounting Review*. 37 (2-3): 355-265.
- Prendergast, C., and R. Topel. 1996. Favoritism in organizations. *The Journal of Political Economy*. 104 (5): 958-978.
- Rogers, W. 1993. Regression standard errors in clustered samples. In *Stata technical bulletin reprints*, 13. College Station, 19-23. TX: Stata Press.
- StataCorp. 1999. *Stata statistical software: Release 6.0 user's guide*. College Station, TX: Stata Press. 23 (11):256-260.
- Wherry J., and C. Bartlett. 1982. The control of bias in ratings: A theory of rating. *Personnel Psychology* 35 (3): 521-551.

Footnotes

¹ For comprehensive reviews of this literature, see Avery and Murphy (1998), Bretz, Milkovich, and Read (1992), Murphy and Cleveland (1995), DeNisi (1997), DeNisi, Cafferty and Meglino (1984), Feldman (1981), Fletcher (2001), Harris (1994), Ilgen, Barnes-Farrell, and McKellin (1993), Ilgen and Feldman (1983), Landy and Farr (1980) and Wherry and Bartlett (1982).

² Two other types of halo effects are in the literature: (a) general impression, in which a rater's overall impression of a ratee influences how the rater evaluates a specific behavior of the ratee; and (b) inadequate discrimination, in which a rater fails to discriminate among conceptually different types of ratee behavior (Nunnally and Bernstein 1994).

³ The geographical distance construct refers to the physical distance between the rater and ratee. This is in contrast to the concept of social distance, which is defined as, "... the degree of reciprocity that subjects believe exist within a social interaction" (Hoffman, McCabe and Smith 1996).

⁴ There was one branch office that had no *ACCREC* value for 2003. For this case, *ACCREC* was reset to the mean value of the observed data for *ACCREC*. The hypothesis-testing results did not qualitatively differ depending on whether this observation was treated as missing or reset at the mean.