Can Assurers Interpret Causal Information?

Sandra C. Vera-Muñoz*
Margaret Shackell
Department of Accountancy
University of Notre Dame

Marc Buehner
School of Psychology
Cardiff University, Wales, UK

December 2003

Preliminary draft. Please do not quote without permission. Comments are welcome.

We acknowledge the helpful comments of Joanna Ho, Bill Kinney, Khim Ong, Dave Ricchiute, and helpful discussions with Lisa Sedor. The authors are indebted to the auditors who participated in the experiment, and to the Master of Science in Accountancy students who participated in the pilot study. Professor Vera-Muñoz acknowledges financial support by KPMG Peat Marwick LLP through its Faculty Fellowship program.

* Corresponding author
248 Mendoza College of Business
University of Notre Dame
Notre Dame, IN 46556-5646
E-mail: s**veramun@nd.edu**
Phone: 574-631-9041
Fax: 574-631-5255

Can Assurers Interpret Causal Information?

ABSTRACT: The accounting profession in general (e.g., AICPA 2003) and public accounting firms in particular are increasingly focusing on the organizations' business models for business measurement process audits (e.g., Bell et al. 1997). Contemporary business measurement systems incorporate a value-based management framework that uses cause-and-effect links to articulate a firm's business model as a theory of how the firm can achieve its goals by implementing concrete strategies. Thus, public accountants as auditors and advisers must develop an understanding of their clients' business models; yet, prior accounting research discredits assurers' and managers' ability to causally link financial and non-financial measures to firm value. Given the attention that all of the Big 4 audit firms place on a client's business measurement process, we address the question: Are contemporary assurers able to exploit causal relationships? In this study, accountants are prompted (or not prompted) with causal model information and with benchmark data that either support or do not support the model, and are asked to make a recommendation requested by a client as to how best to allocate a fixed budget between two non-mutually exclusive programs. We find that causal model information enhances experienced accountants' implicit covariation assessments, but only when the causal model is supported by the data accountants were asked to evaluate. Our findings suggest that a causal model can compensate for the accountants' lack of domain-specific causal knowledge by guiding their interpretation of information during their ongoing process of covariation assessments.

Keywords: *Value-based management; Cause-and-effect business model; Domain-specific causal knowledge; Implicit covariation assessments*

Data Availability: *Contact the authors*.

I. INTRODUCTION

The accounting profession in general (e.g., AICPA 2003) and public accounting firms in particular are increasingly focusing on the organizations' business models, both for business measurement process audits (e.g., Bell et al. 1997, 36), and for making strategic decisions that enhance long-term value.[1] Contemporary business measurement systems incorporate a strategic approach that emphasizes the identification, measurement, and management of key financial and non-financial drivers of shareholder value (Ittner and Larcker 2001, 350). Central to this approach is the value-based management (VBM) framework that uses cause-and-effect links to articulate a firm's business model as a *theory* of how the firm can attain its desired outcomes by implementing concrete strategies (Bell and Solomon 2002; Eccles et al. 2001; Ittner and Larcker 2001; KPMG 1999; Rucci et al. 1998; Bell et al. 1997; Kaplan and Norton 1992, 1996a; 1996b).[2]

Although value-based management presumes that cause-and-effect relationships articulate a firm's business model, prior research questions assurers' and managers' ability to use strategic business models that causally link financial and non-financial measures to firm value (e.g., Maines et al. 2002; Banker et al. 2000; Ittner and Larcker 1998a; Brancato 1995), thereby leading to the research question: Given the attention that all of the Big 4 audit firms place on a client's business measurement process (e.g., KPMG's BMP), are contemporary assurers able to exploit causal relationships?

In their role as information assurers, accountants in public accounting firms must understand their clients' business models causally linking financial and non-financial

---

[1] A business model is a knowledge-acquisition framework that guides the collection and integration of information about key attributes of an entity's full economic system, including the external forces that bear upon the entity; the interlinking activities, or business processes, that are carried out within the entity; and the entity's relationship and interactions with persons and other organizations outside of the entity (Bell et al. 1997, 2).

[2] VBM is a comprehensive framework that encompasses activity-based costing, the balanced scorecard, strategic accounting and control systems, and economic value performance measures (Ittner and Larcker 2001, 350).

performance measures. Such understanding is important for developing expectations of important historical and forward-looking financial and non-financial drivers based on alternative scenarios of business activities. Understanding a client's business model is also important to assess whether additional audit test work is needed when interrelated financial and non-financial performance measures are inconsistent, and when key financial-statement assertions are not consistent with the auditor's understanding of the organization's strategy and measures of process performance. Further, understanding a client's business model is important for analyzing interrelated performance measures (financial and non-financial) both over time and relative to those of the client's competitors (Bell et al. 1997, 36); and for evaluating the completeness and relevance of causally-linked financial and non-financial measures provided by or to management.

According to the causal-model theory of learning (Waldmann 2001, 2000, 1996; Waldmann et al. 1995), causal models assist individuals' covariation detection by providing the basis for the flexible interpretation of data, and thus, an adequate causal representation of the business's environment. The theory assumes that "learners" (e.g., information assurers), bring to bear to the learning task their abstract understanding of causal relations (e.g., domain-general causal knowledge). Then, the causal model provides domain-specific knowledge about covariation between events and outcomes to guide individuals' learning of cause-and-effect relations.[3] That is, a causal model can guide the learner, even if the learner has not had prior direct experience with similar causal situations.

Application of causal-model research results suggests that business strategy articulated as a model that causally links financial and operational drivers would make data meaningful, even

[3] Covariation research addresses how individuals detect or perceive the strength of covariation between or among events or variables (Einhorn and Hogarth 1986). In general, two or more variables are said to covary when the likelihood that their values occur in certain combinations is above chance.

when the model is not supported by the data to be judged (Karmiloff-Smith and Inhelder 1974). This study reports the results of an experiment that relies on prior cognitive psychology (Waldmann 2001, 2000, 1996; Waldman and Hagmayer 2001; Busemeyer et al. 1996, 1997; Waldmann et al. 1995) and accounting research (Ittner and Larcker 2001; 1998a, 1998b; Kaplan and Norton 1996a, 1996b) to address empirically the following question: Does articulation of a business model as a cause-and-effect theory enhance experienced accountants' evaluation of information that require implicit covariation assessments?

Guided by prior research, we sought a scenario that assumes little or no domain-specific knowledge about the covariation between the events and outcomes depicted in our hypothetical scenario. We use experienced public accountants and a complex budget allocation task to fit these requirements. In the experiment, participants make a recommendation requested by a client (an on-line grocery retailer) about how best to allocate a fixed budget between two non-mutually exclusive programs, employee training (hereafter ET) and operating process (hereafter OP) enhancement.

The participants received a set of benchmark data that show *implicitly* either a high or low correlation between ET program expenditures, intervening non-financial measures of employee turnover and complaints, and lagged earnings maximization. In addition, half of the participants are prompted with a business model that articulates a cause-and-effect theory linking higher ET program expenditures to improvements in employee- and customer-related non-financial measures, which are ultimately linked to improvements in earnings. The theory espouses the superiority of the ET program over the OP enhancement program for maximizing earnings. The remaining participants are not prompted with the causal model. To allocate resources optimally, participants must determine implicitly, from the data provided, the level of

covariation between each program and profits. We use participants' allocation of the fixed budget to the ET program as our main dependent variable.

Our main results indicate that accountants who are prompted with the causal model, and with data for which there is (implicitly) high correlation between the ET program's expenditures and reported earnings allocate significantly more resources to the ET program than do accountants who are not prompted with the causal model. However, we find no differences in the resource allocations of accountants who are prompted with the causal model, and with data for which there is (implicitly) low correlation between the ET program's expenditures and reported earnings, relative to accountants who are not prompted with the causal model. Taken together, these results are consistent with causal model theory, and provide support for the conclusion that a causal model can compensate for the accountants' lack of domain-specific causal knowledge by guiding their interpretation of the benchmark data during their ongoing process of covariation assessments.

This study departs from, and extends prior accounting research in several ways. First, except for Waller and Felix (1987, 1989), prior accounting studies make no formal attempts to examine experienced accountants' cause-and-effect covariation assessments. Waller and Felix (1987) examine accountants' explicit covariation judgments for joint (contingency) data, and Waller and Felix (1989) examine auditors' information processing of causal information when making causal judgments. However, prior research does not examine accountants' decisions based on their covariation judgments. To our knowledge, this is the first accounting study to apply causal-model and covariation research results to improve our understanding of whether contemporary assurers are able to exploit causal relationships when prompted with a causal model.

Further, our study contributes to the emerging behavioral accounting research that addresses the effects of using non-financial value drivers, in addition to financial drivers, on the learning and performance of accountants and managers (see, e.g., Ong 2003; Luft and Shields 2001). The common claim in these studies is that both practitioners and researchers have recommended supplementing accounting systems with non-financial measures that reflect key value-creating activities; that is, non-financial drivers. The current study adds to our knowledge of how a cause-and-effect model that describes both direct and indirect links between financial and non-financial drivers can enhance the quality of accountants' information assurance services.

This study also contributes to prior psychological research. In particular, our task implicitly required participants to be aware of the time-lagged nature of the profit structure associated with the ET program. Previous psychological research indicates that people find it extremely difficult to discover delayed causal relations (e.g. Shanks et al. 1989). However, from a knowledge-driven perspective (Waldmann, 1996; Einhorn & Hogarth, 1986), it should be possible to overcome contiguity biases by providing participants with a rationale for the delay. Buehner and May (2002, 2003, in press) have shown that both explicit and implicit instructions about potential delays improve the assessment of delayed contingencies in a free-operant procedure; in fact, Buehner and May's (in press) participants showed no detrimental effect of cause-effect delays at all. Similarly, Hagmayer and Waldmann (2002) have shown that causal interpretations of tabular data depend on temporal assumptions (if the data is sufficiently ambiguous). The current study extends psychological work on individuals' reasoning about time-lagged causal relations in two important ways: firstly, in contrast to all previous work we are aware of, the delayed causal relations in question are *continuous* in nature as opposed to binary (e.g. both cause and effect are realized as various amounts of money, rather than dichotomous

states). Second, the study places theoretical developments into an applied context to test whether experienced decision makers benefit from causal-model information in the way predicted by psychological research.

The remainder of this paper contains four sections. Section II presents the literature review and develops the research hypotheses. Sections III and IV, respectively, present the research design and results. Section V discusses the results and possible implications, and provides directions for future research.

II. BACKGROUND AND HYPOTHESES DEVELOPMENT

The Role of Causal Model Knowledge

Causal-model research suggests that the acquisition of causal knowledge is characterized by an interaction of data-driven and knowledge-driven processes.[4] In particular, individuals bring to bear to causal induction tasks their domain-general causal knowledge, or a knowledge base about how a cause may bring about an effect, and the timeframe the causal relation implies (Lien and Cheng 2000). However, prior research shows that this knowledge base alone is often insufficient for making accurate covariation assessments when cause and effect are separated by time and space. This is because low temporal contiguity places extra cognitive demands on memory and attention (Einhorn and Hogarth 1986). Further, delayed causal relations are harder to identify than contiguous ones because the former increase the number of potentially intervening causal factors (Buehner and May 2003, 867).

According to the causal-model theory of learning (Waldmann 2001, 2000, 1996; Waldmann and Hagmayer 2001; Busemeyer et al. 1996, 1997; Matute et al. 1996; Miller and Matute 1996; Waldmann et al. 1995; Waldmann and Holyoak 1992), causal model knowledge

[4] For example, see Wisniewski and Medin (1994); Einhorn and Hogarth (1986); Murphy and Medin (1985); Trabasso and Sperry (1985); Wright and Murphy (1984); Tversky and Kahneman (1980); Axelrod (1976).

assists covariation detection by reducing the cognitive demands on the learner. Thus, causal models that incorporate domain-specific knowledge about the direction and temporal delays of complex causal relations between causes and intervening (and often unobservable) factors, and between these factors and effects, help guide the learner, even if the learner has not had extensive prior direct experience with similar causal relations (Busemeyer et al. 1996, 1997; Waldmann et al. 1995).[5]

Owing to their experience and training, public accountants are frequently confronted with numerical data regarding an outcome variable (e.g., sales volume, earnings), along with similar data regarding several decision variables that are assumed to affect the outcome (e.g., expenditures on quality initiatives, research and development). Presumably, this experience and training would likely transfer to, and enhance accountants' covariation assessments. However, findings from early accounting studies (Kinney 1987; Waller and Felix 1987; see also Ashton 1982 and Libby 1981) do not lend support to these arguments. These studies suggest that auditors do not readily combine information into patterns in financial data because auditing procedures (e.g., analytical tests) as performed in practice do not stress the use of data in combination.[6]

Causal-model and covariation research results suggest that intuitive covariation assessments consist of an interactive process in which individuals attempt to explain objective data in light of a cause-and-effect theory and evaluate the correctness of the theory at the same

---

[5] Waldmann et al. (1995) test the predictions of causal-model theory in science-related situations where participants were expected to possess no prior direct experience with the causal relations depicted in their experiments.[5] Their results show that different causal models differentially sensitize participants to real versus spurious correlations in the learning data, consistent with the implications of the causal models. The authors conclude that the knowledge provided by the causal models guided the participants' interpretation of the data during the process of covariation assessment

[6] For instance, Waller and Felix (1987) report the results of two experiments that examine the rules by which auditors integrate information contained in a 2 x 2 contingency table when making covariation judgments. The main results show that, while auditors rely on available covariation information, they often overstate or understate the level of the objective covariation.

time. Thus, having a theory benefits tallying consistent versus inconsistent cases, looking for exceptions, and guiding encoding of the data by making it meaningful (Wisniewski and Medin 1994; Einhorn and Hogarth 1986; Murphy and Medin 1985; Trabasso and Sperry 1985; Wright and Murphy 1984; Tversky and Kahneman 1980; Axelrod 1976; Karmiloff-Smith and Inhelder 1974).

In real-life business situations, collected data may not always confirm available theories about the causal relationships that are assumed to exist between or among measures (Eccles et al. 2001, 25).[7] This points to the importance of empirically examining situations where available data either support or does not support the cause-and-effect model. Consistent with causal-model and covariation research, which suggests that articulation of business strategy as a cause-and-effect theory assists individuals' covariation assessments, even when the theory is not supported by the data to be judged (Wright and Murphy 1984; Tversky and Kahneman 1980), we propose:

H1a: Accountants who are prompted with a causal model about an employee-training program and earnings, and with data for which there is a low (implicit) correlation between the employee-training program's expenditures and reported earnings, will allocate less resources to the program than accountants who are not prompted with the model.

H1b: Accountants who are prompted with a causal model about an employee-training program and earnings, and with data for which there is a high (implicit) correlation between the employee-training program's expenditures and reported earnings, will allocate more resources to the program than accountants who are not prompted with the model.

Figure 1 shows graphically the hypothesized patterns predicted by H1a (Panel A) and H1b (Panel B).

[7] Eccles et al. (2001, 25) propose two reasons for why the data may not support the theory. First, decision-makers may be simply using the wrong business model. That is, certain measures may not have predictive value because there is no cause-and-effect relationship. In this case, decision-makers may use already collected performance information to test alternative models. Second, decision-makers may not be able to empirically verify the business model because certain measures may not properly capture the underlying phenomena that managers want to understand. In this situation, there actually is a cause-and-effect relationship, but measurement problems have obscured it. Although not explicitly mentioned in the case materials, the condition depicted in our hypothetical scenario (i.e., where the benchmark data do not support the causal model) is consistent with the first reason above. The second potential reason is beyond the scope of this study.

III. METHOD

Overview

Guided by prior research, we sought a scenario that assumes that participants possess little or no domain-specific knowledge about the causal relations depicted in our hypothetical case (discussed below). We use experienced public accountants and a complex budget allocation task to fit these requirements. Our task requires participants to provide a recommendation to a client as to how best to allocate a fixed budget between two non-mutually exclusive programs, employee training (ET) and operating process (OP) enhancement.

Participants receive a set of benchmark data (i.e., the learning input) that provide information about several peer competitors who have simultaneously implemented both programs at different levels of expenditures.[8] The benchmark data describe a set of financial and non-financial outcomes and decision variables (i.e., ET and OP expenditures) that potentially affect the outcome (i.e., the effects of each program on earnings over time). A key feature is that the data show *implicitly* either a high or low correlation between ET program expenditures in year *t*, intervening non-financial measures of employee turnover and complaints, and lagged earnings maximization in year $t + 4$. We ask participants to choose values (i.e., monetary allocations) for the decision variables that would maximize the outcome. To allocate resources optimally, participants must determine implicitly, from the benchmark data provided, the level of covariation between the outcome and each decision variable.[9] A desirable aspect of this task is

[8] Collection of benchmark data is consistent with current diagnostic business assurance practice. For instance, in their description of KPMG's business measurement process, Bell et al. (1997, 36) assert, "During the continuous improvement phase of the audit, the auditor prepares and reports process performance measures, and measures from competitors that have demonstrated consistent process advantage."

[9] Two prior studies that address individuals' implicit covariation assessments in budget allocation tasks are Hutchinson and Alba (1997) and Broniarczyk and Alba (1994a).

that it is intuitively natural (and often optimal) for allocation to be monotonically related to the level of correlation in the benchmark data.

Participants

The participants were 86 public accountants recruited from two then-Big 5 U.S. firms (n = 48), and a national Canadian firm (n = 38). The participants' average public accounting experience is 36 months (s.d. = 19.25), ranging from 12 to 126 months. The average public accounting experience of the accountants in the U.S. firms (35.40; s.d. = 17.09) is not significantly different from the average experience of accountants in the Canadian firm (36.87; s.d. = 21.88) (p = 0.727).

To rule out systematic differences in the participants' firm membership and experience (measured in months) as alternative explanations for our results, we included these variables as covariates in our planned comparison tests (discussed below). Our results show that neither of these variables affected significantly the dependent variables at conventional levels. Thus, the results reported in this study exclude these variables.

Experimental Task and Procedures

We randomly assigned each participant to one of four experimental conditions (discussed below). Part One of the case materials (experimental phase) provides a general overview of the on-line grocery industry in general, and its transition from brick-and-mortar stores to on-line ordering. The case also describes the client's operations and current financial condition.[10] The client (president of the on-line grocery delivery company) wants to know how best to allocate a fixed budget of $1.5 million between two programs, OP and ET.

[10] We developed our case materials based on Harvard Business Review and Ivey School of Business cases (e.g., Bakshi and Deighton 2000; Beckow and Huff 1998).

Participants received information containing benchmark data from six on-line grocery competitors that had invested in both the OP and ET programs, including non-financial measures for a four-year period following the investments, and excerpts from income statements for the same period. The OP program uses investment in technology, which is expected to reduce operating expenses by eliminating non-value-adding activities of the order fulfillment process, and thus, increase earnings. The ET program uses investment in employee training that is expected to increase employee satisfaction, which in turn is expected to enhance customer satisfaction, and ultimately earnings.

Upon studying the benchmark data, but prior to providing their recommendations, participants were asked to provide likelihood assessments about the relative success of each of the programs in maximizing earnings. Then they were asked to provide their resource allocation recommendations, including supporting calculations, and to explain the rationale for their recommendation in a response memorandum. To reduce task complexity, we instructed participants to ignore any time value of money and tax considerations. We also asked that they refrain from considering any other alternatives than the two described in the case. Finally, participants were asked to indicate their level of confidence in their recommendation. Part Two (post-experimental phase) includes a debriefing questionnaire, an analytical reasoning/statistics knowledge test, and a background questionnaire. Figure 2 shows graphically a summary of the experimental procedures of Parts One and Two.

Independent Variables

We used a between-subjects design with two variables manipulated at two levels each: correlation between ET program expenditures and earnings (low vs. high), and business model articulated as a cause-and-effect theory (prompted vs. not prompted). The benchmark data report

the amounts of ET and OP program expenditures that six on-line grocery competitors committed in year 0 for the subsequent four-year period, and show partial results of operations, including financial and non-financial measures. The average ET and OP program expenditures across the six on-line grocery stores are the same in the high and low correlation conditions. OP program expenditures occur at the end of year 0, whereas ET program expenditures occur throughout the four-year period. OP program expenditures have diminishing marginal returns on earnings before interest, taxes, depreciation, and amortization (EBITDA), while the effects of ET program expenditures on earnings have a three-year lag, thus reflecting economic reality (e.g., see Hilton et al. 2003, 838-839).[11] The reported non-financial measures associated with ET program expenditures are employee turnover and employee complaints, and those associated with OP program expenditures are order accuracy and orders filled per hour.

All experimental participants received information about the ET program and its related non-financial measures for an on-line bookstore company. Participants who were not prompted with the causal model received summary information about the ET program and related non-financial measures for the on-line bookstore company, as follows:

> The on-line grocery industry uses two key measures to assess employee training, namely, the employee turnover rate and employee complaints. The employee turnover rate is measured as the proportion of employees leaving the company in any year. Employee complaints are measured as the number of complaints per every 1,000 employees in any year. FG's goal is to lower the average employee turnover rate from 35 percent in 1999 to less than 20 percent by 2002, and to lower average employee complaints from 30 (per every 1,000 employees) in 1999 to less than 10 by 2002.

[11] Similar to Luft and Shields (2001), we chose to allow a difference across conditions in underlying (unobservable) revenue and expense processes, rather than in the benchmark data participants received. The difference in unobservable processes, if it had any effect, worked against finding support for the study's hypotheses. If participants tried to disentangle the revenue and expense effects, rather than simply causally relate expenditures to earnings as required by the experimental task, then the budget allocation task would be more difficult in the high correlation condition (due in part to the uncertainty about the length of the ET program's lagged effects). This fact biases against finding the predicted effects.

Participants who were prompted with the causal model received the preceding paragraph. In addition, they received another paragraph preceded by the subheading "Employee training→Customer Satisfaction→Earnings Chain." The paragraph articulated the business model as a cause-and-effect theory linking expenditures in an ET program to improvements in employee turnover, which in turn is linked to customer retention and ultimately to improvements in earnings.[12] The causal model espoused the superiority of the ET program over the OP program for maximizing earnings, as shown in the last sentence of the following excerpt from the causal model-prompted condition:

> For example, two divisions at a well-known on-line bookstore company undertook a study to quantify the links among employee turnover, customer retention, and profitability. Management estimated that a 10 percent reduction in employee turnover would reduce customer turnover by 1 percent to 3 percent and raise revenues by $50 million to $150 million. The study's conclusions are striking: even with high-end estimates for recruitment and training costs and low-end estimates for the cost of lost customers, reducing employee turnover yielded increases in profits over time that were significantly greater than the cost savings generated from programs implemented to enhance the operational efficiency of the stores.

Panels A and B of Appendix 1 show the full information provided to participants who were not prompted with the causal model vs. those who were prompted with the model, respectively.

Participants in the low correlation condition received benchmark data that do not support the causal model. Accordingly, the four-year average correlation between ET program expenditures and EBITDA is 0.20. Further, the implicit correlations for the same period range from 0.160 to 0.225, and none are significant (p-values > 0.60). In contrast, the implicit correlations between OP program expenditures and EBITDA in years 1 and 2 are positive (Pearson correlations > 0.70) and significant (p-values < 0.05). The implicit correlations in years 3 and 4 (Pearson correlations < 0.55) are not significant (p-values > 0.25). Further, the implicit

[12] The causal model for the employee training program used in our case materials is based on Harvard Business School cases (Rucci et al. 1998; Heskett et al. 1994; Schlesinger and Heskett 1991).

correlations between OP program expenditures and non-financial measures of order accuracy and orders filled per hour are positive (Pearson correlations > 0.70) and significant (p-values < 0.05).

Participants in the high correlation condition received benchmark data that support the causal model. Accordingly, the implicit correlations between ET program expenditures and lagged EBITDA are positive (Pearson correlations > 0.70) and significant in years 3 and 4 (p-values < 0.05), but they are not significant in years 1 and 2 (Pearson correlations < 0.50; p-values > 0.35). Further, the implicit correlations between ET program expenditures and three-year lagged non-financial measures of employee turnover and employee complaints are negative (Pearson correlations < -0.80) and significant ($p < 0.05$).[13] At the same time, the four-year average correlation between OP program expenditures and EBITDA is 0.24, and the implicit correlations for the same period, ranging from -0.020 to 0.490, are not significant (p-values > 0.30).

Finally, regression equations regressing EBITDA on the two programs show that, in the high correlation condition, the adjusted regression coefficients for the ET program expenditures in years 3 and 4 are above 0.90 and significant (p-values < 0.05), while none of the regression coefficients for OP program expenditures is significant at conventional levels in any of the four years. In the low correlation condition, the adjusted regression coefficients for the OP program expenditures in years 1 and 2 are above 0.90 and significant (p-values < 0.05), while none of the regression coefficients for ET program expenditures is significant at conventional levels in any of the four years. Panels A and B of Appendix 2 show a sample of benchmark data provided to participants in the low and high correlation conditions, respectively.

[13] The levels of correlations for the high condition were based on prior covariation research (Fugelsang and Thompson 2000; Hutchinson and Alba 1997; Broniarczyk and Alba 1994a; 1994b).

We conducted a pilot test with twenty-three Master of Science in Accountancy students enrolled in a graduate-level auditing course at a private Midwestern university. We randomly assigned eleven pilot subjects to the benchmark data in the low correlation-no causal model prompted condition, and twelve pilot subjects to the benchmark data in the high correlation-causal model prompted condition. All participants received the same benchmark data used in the experiment. The participants were asked to indicate their level of agreement with the following statement: "Spending on the employee training program is more effective at maximizing profits over the four-year period than spending on the operating process enhancement program," on a nine-point scale anchored on the lower end by "Strongly disagree," the higher end by "Strongly agree," and the midpoint "Neither agree nor disagree."

The mean score for participants in the low correlation-no causal model prompted condition (4.53; s.d. = 1.26) is significantly lower than the mean score for participants in the high correlation-causal model prompted condition (6.29; s.d. = 2.31) ($t = 1.941$; $p = 0.033$, one-tailed). These results indicate that the correlations were perceived in the direction consistent with the manipulation.[14]

Dependent Variables and Covariates

We use two dependent variables, namely, the participants' allocations of the fixed budget of $1.5 million to the ET program, and the participants' likelihood assessments. We constructed the second dependent variable using participants' responses to the following question: "For this question, ignore other factors that may affect the earnings of the on-line grocery companies in Exhibit 1, and concentrate only on the effects of the two programs. If earnings increase by

[14] The difference in means between the two correlation conditions may not have been as large as expected. However, this was possibly due to the implicit nature of the correlation manipulation coupled with the high task complexity (e.g., low temporal contiguity of cause-and-effects). See Hagmayer and Waldmann (2002, 1132) for similar results.

$1,000, what percentage of the increase do you think should likely be accounted for by each of the programs? Use a 0% to 100% scale to provide your likelihood assessments." We elicited participants' likelihood assessments immediately after they had studied the benchmark data but before they were asked to provide their recommendations. For our analyses, we used the participants' likelihood assessments regarding the ET program.

A large body of cognitive psychology research on learning suggests that in complex causal learning situations, such as the one portrayed in the current study, individuals learn more about structural properties of the task than they can verbalize (Berry and Broadbent 1988, 1987, 1984; FitzGerald and Broadbent 1985; Reber et al. 1980; Reber 1976).[15] This is because upon encountering a complex task, learners tend to observe all the variables unselectively and attempt to store all of the contingencies between them, including the correct factors as well as the incorrect ones. Use of the unselective mode of learning, therefore, is unlikely to be associated with accurate verbalizable knowledge (Berry and Broadbent 1988, 253-254).[16] This research suggests that, given the complexity of our experimental task, a measure of task performance (i.e., participants' allocations of the fixed budget to the ET program) should be more effective at capturing what accountants learn about the implicit covariation in the benchmark data than their response to a direct elicitation of covariation assessments. If our results are consistent with prior cognitive psychological research, then the effects of causal model should affect significantly the participants' monetary allocations to the ET program, but not their likelihood assessments. We

[15] Verbalizable knowledge refers to task-related information that is (or has been) conscious and is potentially expressable in language. In the Berry and Broadbent (1988, 1987, 1984) studies, verbalizable knowledge is assessed by written questionnaire.

[16] In contrast, the selective mode of learning is effortful and reportable. In this mode of learning a person selects a few key variables and only observes the contingencies between these key variables. This is likely to result in a form of knowledge that can be made explicit because of the relative small number of relationships involved. The distinction between these two modes has also been equated to the distinction between explicit and implicit forms of learning (Berry and Broadbent 1988, 254).

use a full MANOVA model, which allows for two or more dependent variables, as a basis for our planned comparison tests (discussed below).

Prior research suggests that analytical ability may affect decision-makers' performance (Vera-Muñoz 1998; Nelson 1993; Libby and Tan 1992; Bonner 1990; Bonner and Lewis 1990). Therefore, although analytical ability is assumed in this study to be above the minimum necessary for the experimental task, we measure the participants' analytical ability using a six-item analytical ability test. Reliability test results indicate that a total analytical reasoning score derived from the answers to the first five questions produces a reliable measure of analytical reasoning (Kuder-Richardson 20 = 0.641).[17] The participants' average score on the 5-item analytical reasoning test is 3.80 (s.d. = 1.31).

Further, because the ability to understand cause-and-effect relationships using raw data, and thus, to predict the relative success of the two investment programs in maximizing earnings might depend on statistical knowledge, we also measure the participants' statistical knowledge using a three-item statistical knowledge test.[18] The average score on the 3-item test is 1.55 (s.d. = 0.941).[19] To rule out systematic differences in the participants' analytical ability and statistical reasoning as alternative explanations for our results, we included these variables as covariates in

[17] By comparison, the reliability coefficient achieved with the same questionnaire in Vera-Muñoz (1998, 56) is 0.62. The KR20 measure is appropriate for measuring reliability when the questions are scored as either 0 or 1 for incorrect or correct answers, respectively. It is equivalent to Cronbach's alpha, which is used when individual questions can take more than two values.

[18] Anecdotal evidence also supports the argument that more statistics knowledge may help accountants' ability to understand cause-and-effect relationships. For instance, an experienced consultant asserts, "Don't change measures just because top management has forgotten Statistics 101. I recently examined a cross section of companies to determine which customer-measurement skills and practices showed convincing connection to growth. Surprise: companies whose leaders knew enough statistics to understand which measures were strategically valuable tended to grow faster" (Hochman 1999).

[19] Reliability test results indicate that a statistical reasoning score derived from the answers to two of the three questions produces a KR20 score of 0.50. The KR20 score is sensitive to test length, with longer tests receiving a higher score; thus, the low score may be due to the short length of our test.

our planned comparison tests.[20] Our results show that neither of these variables affected significantly the dependent variables at conventional levels. Thus, the results reported in this study exclude these variables.

IV. RESULTS

Validity Tests

We used post-experimental questions to elicit participants' assessments of task difficulty and familiarity across experimental conditions, because it is possible that these perceptions are related to task performance. Participants rated task difficulty on a nine-point scale anchored on the lower end by "Not difficult" and on the higher end by "Very difficult." Participants were also asked to indicate whether they had encountered any analogous information evaluation situations that helped them perform the task, using a nine-point scale anchored on the lower end by "Nothing analogous" and on the higher end by "Completely analogous."

The mean difficulty score of participants in the low correlation condition (5.84; s.d. = 1.63) is significantly lower than the mean difficulty score of participants in the high correlation condition (6.67; s.d. = 1.32) ($p = 0.018$), while the difficulty scores between causal model conditions are not significantly different ($p = 0.734$). The mean familiarity score of participants in the low correlation condition (3.74; s.d. = 1.96) is significantly lower than the mean score of participants in the high correlation condition (4.88; s.d. = 1.90) ($p = 0.016$), while the mean familiarity scores between causal model conditions are not significantly different ($p = 0.219$). We included the difficulty and familiarity variables as covariates in our planned comparison

[20] We find no correlation between analytical reasoning and experience ($p > 0.35$), nor between statistical reasoning and experience ($p > 0.76$).

tests. Our results show that neither of these variables affected significantly the dependent variables at conventional levels. Thus, the results reported in this study exclude these variables.[21]

Descriptive Statistics

Recall that prior to making their budget allocation recommendations, participants were asked to provide their likelihood assessments about the relative effectiveness of the ET and OP programs in maximizing earnings over time. Panel A of Table 1 shows descriptive statistics of the participants' mean likelihood assessments, by level of correlation between ET and earnings (denoted as $\rho(ET, \pi)$), and whether or not the participants were prompted with a causal model. The panel shows that the mean likelihood assessment score of participants in the low correlation condition is close to 50 percent across causal model conditions. Similarly, the mean likelihood assessment score of participants in the high correlation condition is almost 51 percent across causal model conditions.

Consistent with expectations, chi-square tests show that the cell proportions of participants' likelihood assessments are independent of whether or not they were prompted with a cause-and-effect model (Chi-square = 24.09; p = 0.343), and of whether they were in the low or high correlation condition (Chi-square = 21.85; p = 0.379). These results indicate that the participants' likelihood assessments were not affected by the experimental variables, and were constant across conditions.

Insert Table 1 here

Panel A of Table 1 also shows the mean budget allocations of participants to the ET and OP programs, by level of correlation between ET and earnings, and whether or not the

[21] We find no significant correlation between experience, and task difficulty and familiarity (p-values > 0.20).

participants were prompted with a causal model. The largest average budget allocation to the ET program is for accountants in the high correlation condition who were prompted with the causal model (cell 4), and the smallest average budget allocation to the ET program is for those in the low correlation condition who were prompted with the causal model (cell 2). Participants in the former experimental condition allocated almost 23 percent more resources to the ET program than did those in the latter experimental condition.[22]

Tests of Research Hypothesis

H1a predicts that accountants who are prompted with a causal model, and with data for which there is a low (implicit) correlation between an employee-training program's expenditures and reported earnings, will allocate less resources to the program than accountants who are not prompted with the model. Recall that we use two dependent variables, namely, the accountants' likelihood assessments and their monetary allocations to the ET program. Because the dependent variables are significantly and positively associated (Pearson correlation = 0.753; p = 0.000), we use a between-subjects MANOVA model as a basis for our planned comparison tests.[23] Panel B of Table 1 presents the results of planned comparisons to test our research hypotheses.

The results show that, for accountants in the low correlation condition, those who were prompted with the causal model (cell 2) did not allocate significantly less resources, on average, to the ET program than did accountants who were not prompted with the model (cell 1) (F =

[22] We asked participants to indicate their level of confidence in their resource allocation recommendations, using a 101-point scale with endpoints of 0% ("Not confident at all") and 100% ("Extremely confident"), and a midpoint of 50% ("Moderately confident"). The mean confidence scores for participants who were prompted with the causal model (57.69; s.d. = 24.43) did not differ significantly from the mean scores of those who were not prompted with the causal model (58.57; s.d. = 19.42) (p = 0.855). Also, the mean confidence scores for participants in the low correlation condition (57.21; s.d. = 24.21) did not differ significantly from the scores of those in the high correlation condition (59.10; s.d. = 19.54) (p = 0.696).

[23] All the assumptions of the MANOVA model (i.e., the null hypothesis that the observed covariance matrices of the dependent variables are equal across groups (Box's M); the variance/covariance matrix of the dependent variables are circular in form (Bartlett's test of sphericity); and the null hypothesis that the error variance of the dependent variables is equal across groups (Levene's test of equality of error variances)) are satisfied.

0.02; p = 0.439, one-tailed). This result, shown graphically in Figure 3, does not support H1a's prediction. As expected, the effect of causal model on the participants' likelihood assessments is not significant (p = 0.754, two-tailed).

H1b predicts that accountants who are prompted with a causal model, and with data for which there is a high (implicit) correlation between an employee-training program's expenditures and reported earnings, will allocate more resources to the program than accountants who are not prompted with the model. Panel B of Table 1 shows that, for accountants in the high correlation condition, those who were prompted with the causal model (cell 4) allocated significantly more resources, on average, to the ET program than did accountants who were not prompted with the causal model (cell 3) (F = 3.18; p = 0.041, one-tailed). This result, shown graphically in Figure 3, supports H1b's prediction. As expected, the effect of causal model on the participants' likelihood assessments is not significant (p = 0.823, two-tailed).

Taken together, the results indicate that when experienced public accountants are prompted with a business model articulated as a cause-and-effect theory and with data that support the model (i.e., in the high correlation condition), the model enhances accountants' evaluation of information when making decisions that require implicit cause-and-effect covariation assessments. However, prompting experienced public accountants with a causal model and with data that do not support the model (i.e., in the low correlation condition) does not affect their implicit covariation assessments.

Supplementary Analyses

We used a post-experimental question to assess how well-calibrated were the participants' ex-post perceptions of covariation between the ET program and earnings (based on the benchmark data provided), with their budget allocation recommendations to the ET program.

The post-experimental question asked participants to indicate their level of agreement with the following statement: "Spending on the employee training program is more effective at maximizing profits over the four-year period than spending on the operating process enhancement program," using a nine-point scale anchored on the lower end by "Strongly disagree," the higher end by "Strongly agree," and the midpoint "Neither agree nor disagree."

Results of correlation analyses show that the correlation between the accountants' *ex-post* perceptions of covariation and their budget allocation to the ET program is positive and significant for participants in the low (Pearson correlation = 0.543; p = 0.000) and high (Pearson correlation = 0.610; p = 0.000) correlation conditions. Further, the correlations are also positive and significant for participants who were not prompted with the causal model (Pearson correlation = 0.638; p = 0.000) and those who were prompted with the model (Pearson correlation = 0.512; p = 0.001). Taken together, these results provide evidence that the participants' self-reports of their covariation perceptions were well-calibrated, given the high level of consistency of their covariation assessments with their budget allocation recommendations.

We also compared the mean scores of the post-experimental question above (i.e., the *ex-post* perception of covariation) between-causal model conditions and between-correlation conditions. First, we find that, for participants in the low correlation condition, the mean scores are not significantly different between the causal-model prompted (4.22; s.d. = 2.00) and causal-model not prompted conditions (5.32; p = 2.88) (p = 0.182). For participants in the high correlation condition, the mean score of those who were prompted with the causal model (5.78; s.d. = 2.24) is marginally significantly higher than the mean score of those who were not prompted with the causal model (4.26; s.d. = 2.22) (p = 0.060).

Second, we find that for participants in no causal-model prompted condition, the mean scores are not significantly different between the low (5.32; s.d. = 2.88) and high (4.26; p = 2.22) correlation conditions. In contrast, for participants in the causal-model prompted condition, the mean score of those in the high correlation condition (5.78; s.d. = 2.24) is significantly higher than the mean score of those in the low correlation condition (4.22; s.d. = 2.00) (p = 0.035). Recall that participants rated the task as significantly more difficult in the high correlation condition than in the low correlation condition. However, the results reported here suggest that prompting participants with the causal model improved their covariation assessments—notwithstanding the higher perceived level of difficulty reported by participants in the high correlation condition—relative to participants who were not prompted with the causal model.

We conducted additional planned comparison tests to probe deeper into the effects of the causal model on the dependent variables for participants in the high vs. low correlation conditions.[24] Panel C of Table 1 shows that, for accountants who were not prompted with the causal model, those in the high correlation condition (cell 3) did not allocate significantly more resources, on average, to the ET program than did accountants in the low correlation condition (cell 1) (F = 0.79; p = 0.492, one-tailed). As expected, the effect of correlation on the participants' likelihood assessments is not significant (p = 0.379, two-tailed).

In contrast, for accountants who were prompted with the causal model, those in the high correlation condition (cell 4) allocated significantly more resources, on average, to the ET program than did accountants in the low correlation condition (cell 2) (F = 3.74; p = 0.030, one-tailed). As expected, the effect of correlation on the participants' likelihood assessments is not

[24] To rule out systematic differences in the participants' scores of task difficulty, task familiarity, analytical reasoning, statistical reasoning, and public accounting experience as potential explanations for our results, we included these variables as covariates in the MANOVA model used as a basis for our planned comparison tests. Neither of these variables affected significantly the dependent variables at conventional levels. Thus, the results reported here exclude these variables.

significant (p = 0.729, two-tailed). These results provide further support to the conclusion that prompting the accountants with the causal model helped them to better capture the implicit covariation in the benchmark data, relative to the accountants who were not prompted with the causal model.

V. DISCUSSION, IMPLICATIONS, AND DIRECTIONS FOR FUTURE RESEARCH

Increasingly, business measurement systems incorporate a strategic approach that emphasizes the identification, measurement, and management of key financial and non-financial drivers of shareholder value. Central to this approach is the value-based management framework, which uses cause-and-effect links to articulate a firm's business model as a theory of how the firm can achieve its desired outcomes by implementing concrete strategies. Thus, the accounting profession in general, and public accounting firms in particular are increasingly focusing on the organizations' comprehensive business models, both for business measurement process audits and for making strategic decisions that enhance long-term value. However, prior accounting research questions the ability of accountants and corporate managers to use strategic business models that causally link financial and non-financial performance drivers to firm value.

This study uses causal-model and covariation research results to address empirically: Given the attention that all of the Big 4 audit firms place on a client's business measurement process, are contemporary assurers able to exploit causal relationships? We use a highly-complex business scenario that assumes little or no domain-specific knowledge about the covariation between two programs and earnings maximization, and uses benchmark data that either support or do not support the underlying theory of the cause-and-effect relationships.

Our main results indicate that when experienced public accountants are prompted with a business model articulated as a cause-and-effect theory and with data that support the model, accountants are better able to exploit causal relationships in the data when making decisions that

require implicit cause-and-effect covariation assessments. However, prompting accountants with a causal model and with data that do not support the model does not affect their implicit covariation assessments.

Taken together, our results are consistent with the predictions of covariation research and with the causal-model theory of learning. First, covariation research shows that, due to the limited capacity of their working memory, individuals in general lack the ability to make correct attributions of covariation when simultaneous actions are present, and when causes-and-effects are separated by time and space (Shanks et al. 1989; Wasserman and Neunaber 1986). Second, consistent with the causal-model theory of learning, knowledge of the causal model allowed accountants to incorporate domain-specific knowledge about likely covariation in the benchmark data. Thus, prompting accountants with the causal model helped them overcome the inherent cognitive demands of the covariation task, but this advantage was observed only when the causal model was supported by the data.

The results of this study have implications for accountants as information assurers in public accounting firms, their clients, and accounting educators. During the business measurement phase of the business measurement process audit, auditors measure the processes and variables that have the greatest impact on the business. They also evaluate interrelated performance measures (financial and operational), both over time and relative to those of similar organizations (Bell et al. 1997, 36). This implies that public accountants must develop a better understanding of their clients' business causal models. Without such understanding accountants may be unable to develop expectations of important forward-looking financial and operational drivers, or to determine whether additional audit test work is needed when interrelated financial and operational measures are inconsistent.

Furthermore, with enhancements in information technology, many businesses have significantly more data available—both financial and operational—at a lower cost. Yet, they have not been able to appropriately link financial measures with operational measures, or to determine how these measures in combination are causally linked to long-term value creation. Thus, public accountants who understand the complex mechanisms underlying cause-and-effect relations (e.g., when intervening unobservable factors and temporal lags between causes and effects are present) may be better placed to help their clients align management processes and strategies (e.g., business planning, resource allocation), with long-term value creation (Eccles et al. 2001, 5; KPMG Consulting 1999). A better understanding of their clients' causal business models may also help accountants evaluate the appropriateness, completeness, and relevance of causally-linked financial and operational measures provided by or to management.

The CPA Vision's goal of "[m]oving the profession away from a backward looking, historical focus into a value-oriented mindset" (AICPA 1999, 6) calls for accounting educators and public accounting firms to incorporate in their curriculum and training programs use and applications of knowledge of the value-based management framework, forward-looking financial indicators, and development of causal business models. One broad question for practice and future research that arises from our findings is: How can firms facilitate accountants' evaluation of information for business measurement processes that require covariation assessments? One approach might be to develop guidance for customizing information in the form of decision aids such as cause-and-effect business models that link financial and non-financial measures (e.g., customer satisfaction, employee training and satisfaction, product quality, sales calls proposals delivered) to value creation. To ensure maximum effectiveness, the business models should show

explicitly potential intervening factors (e.g., employee training expenditures cause increases in customer satisfaction) that otherwise may be unobservable, and expected temporal lags between causes and effects. Some initial attempts have been made to develop such causal business models, such as the widely cited Rucci et al.'s (1998) employee-customer-profit chain at Sears.

Certain features of this study point to potential ideas for future research. First, in our experiment we chose to manipulate between-subjects the implicit correlation in the benchmark data (i.e., high vs. low) and provision of the causal model (i.e., prompted vs. not prompted). The model was supported by the benchmark data that show implicitly high correlation between the ET program and earnings. The value of our professional participants' time prohibited us from conducting an augmented design that would address other "theory-data" conflict scenarios. Future research could examine the robustness and generality of our findings, for example, by manipulating the correlation in the data as in the current study, but using two different causal models: one model would suggest high correlation between ET program expenditures and earnings, while the other model would suggest low correlation between ET program expenditures and earnings. Thus, one question for future research that arises from our findings is: Are experienced assurers (and managers) able to exploit causal relations from the data when they are prompted with the "right" model?

Second, while our ultimate goal was to assess whether experienced public accountants improve their decisions when prompted with a causal model, the main dependent variable used (i.e., differences in the amount of resources allocated to the ET program) is one-step removed from assessing the ultimate quality of the accountants' decisions. Our measure captures whether accountants detect differences in correlation but does not capture whether the significant differences in ET expenditures lead to significant differences in earnings. However, we argue

that the ability to correctly detect differences in covariation is the first step towards making better decisions. Future research, however, could examine the robustness and generality of our findings by using the effect of the accountants' allocation to the ET program on earnings as a measure of decision quality. Finally, future research could also address the extent to which different types of accounting experience and training (e.g., predominantly financial versus predominantly managerial) and/or different types of task (e.g., reliability assurance vs. relevance assurance) mediate accountants' and managers' implicit covariation assessments given (or not given) a causal business model.

REFERENCES

American Institute of Certified Public Accountants (AICPA). 2003. Special Committee on Enhanced Business Reporting. Available at http://www.aicpa.org/innovation.scebr.htm

American Institute of Certified Public Accountants (AICPA). 1999. Assurance services are an integral part of the CPA Vision. The CPA Letter (October): 1-6.

Ashton, R. 1982. *Human Information Processing in Accounting*. American Accounting Association Monograph No. 17. Sarasota, FL: American Accounting Association.

Axelrod, R. 1976. *The Structure of Decision: The Cognitive Maps of Political Elites*. Princeton, NJ: Princeton University Press.

Bakshi, K. and J. Deighton. 2000. Webvan: Groceries on the Internet (case #9-500-052). Harvard Business School Publishing. Boston, MA: 1-17.

Banker, R., G. Potter, and D. Srinivasan. 2000. An empirical investigation of an incentive plan that includes nonfinancial performance measures. *The Accounting Review* 75 (January): 65-92.

Beckow, D., and S. Huff. 1998. Homegrocer.com (case #9A98E019). Richard Ivey School of Business Publication. London, Ontario, Canada: 1-16.

Bell, T., and I. Solomon. 2002. *Cases in Strategic Systems Auditing*. KPMG Peat Marwick LLP. Montvale, NJ.

Bell, T., F. Marrs, I. Solomon, and H. Thomas. 1997. *Auditing Organizations Through a Strategic-Systems Lens—The KPMG Business Measurement Process*. KPMG Peat Marwick LLP. Montvale, NJ.

Berry, D., and D. Broadbent. 1988. Interactive tasks and the implicit-explicit distinction. *British Journal of Psychology* 79: 251-272.

Berry, D., and D. Broadbent. 1987. The combination of explicit and implicit learning processes in task control. *Psychological Research* 49: 7-15.

Berry, D., and D. Broadbent. 1984. On the relationship between task performance and explicit verbalizable knowledge. *Quarterly Journal of Experimental Psychology* 36A: 209-231.

Bonner, S. 1990. Experience effects in auditing: The role of task-specific knowledge. *The Accounting Review* (January): 72-92.

Bonner, S., and B. Lewis. 1990. Determinants of auditor expertise. *Journal of Accounting Research* (Supplement): 1-20.

Brancato, C.K. 1995. New Performance Measures—A Research Report. Report Number 1118-95-RR. New York, NY: The Conference Board.

Broniarczyk, S., and J. Alba. 1994a. Theory versus data in prediction and correlation tasks. *Organizational Behavior and Human Decision Processes* 57: 117-139.

Broniarczyk, S., and J. Alba. 1994b. The role of consumers' intuitions in inference making. *Journal of Consumer Research* (December): 393-407.

Buehner, M. J., and J. May. 2002. Knowledge mediates the timeframe of covariation assessment in human causal induction. *Thinking and Reasoning 8*(4): 269-295.

Buehner, M. J., and J. May. 2003. Rethinking temporal contiguity and the judgment of causality: Effects of prior knowledge, experience, and reinforcement procedure. *Quarterly Journal of Experimental Psychology Section A-Human Experimental Psychology* 56A(5): 865-890.

Buehner, M. J., and J. May. (in press). Abolishing the effect of reinforcement delay on human causal learning. *Quarterly Journal of Experimental Psychology Section B- Comparative and Physiological Psychology*.

Busemeyer, J., M. McDaniel, and E. Byun. 1997. The abstraction of intervening concepts from experience with multiple input-multiple output causal environments. *Cognitive Psychology* 32: 1-48.

Busemeyer, J., M. McDaniel, and E. Byun. 1996. The use of intervening variables in causal learning. In D.R. Shanks, K.J. Holyoak, and D.L. Medin (Eds.), *The psychology of learning and motivation*: Vol. 34. Causal Learning. San Diego: Academic Press: 357-391.

Eccles, R., R. Herz, M. Keegan, and D. Phillips. 2001. *The ValueReporting™ Revolution-Moving Beyond the Earnings Game*. New York, NY: John Wiley & Sons.

Einhorn, H., and R. Hogarth. 1986. Judging probable cause. *Psychological Bulletin* 99 (1): 3-19.

FitzGerald, P., and D. Broadbent. 1985. Memory for attended and unattended visual stimuli. *The Quarterly Journal of Experimental Psychology* 37A: 339-365.

Fugelsang, J., and V. Thompson. 2000. Strategy selection in causal reasoning: When beliefs and covariation collide. *Canadian Journal of Experimental Psychology* (March): 15-32.

Hagmayer, Y., and M. Waldmann. 2002. How temporal assumptions influence causal judgments. *Memory & Cognition* 30(7): 1128-1137.

Heskett, J., T. Jones, G. Loveman, W.E. Sasser, and L. Schlesinger. 1994. Putting the service-profit chain to work. *Harvard Business Review*: 164-174.

Hilton, R., M. Maher, and F. Selto. 2003. Using lead and lag measures to communicate, motivate, and evaluate (Chapter 20). In Hilton et al. (Editors), *Cost Management—Strategies for Business Decisions* (2nd Edition). New York: McGraw-Hill-Irwin: 830-869.

Hochman, M. 1999. Customer-satisfaction measurements: An answer to yesterday's problem? *Harvard Management Update*. Harvard Business School Publishing. Boston: MA.

Hutchinson, J.W., and J. Alba. 1997. Heuristics and biases in the "eyeballing" of data: The effects of context on intuitive correlation assessment. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 23: 591-621.

Ittner, C., D. Larcker. 1998a. Are non-financial measures leading indicators of financial performance? An analysis of customer satisfaction. *Journal of Accounting Research* 36 (Supplement): 1-35.

Ittner, C., D. Larcker. 1998b. Innovations in performance measurement: Trends and research implications. *Journal of Management 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: 349-410.

Kaplan, R., and D. Norton. 1992. The balanced scorecard—Measures that drive performance. *Harvard Business Review* 70 (1): 71-79.

Kaplan, R. and D. Norton. 1996a. *The Balanced Scorecard.* Boston, MA: Harvard Business Press.

Kaplan, R., and D. Norton. 1996b. Using the balanced scorecard as a strategic management system. *Harvard Business Review* (January-February): 75-85.

Karmiloff-Smith, A., and B. Inhelder. 1974. If you want to get ahead, get a theory. *Cognition* 3: 195-212.

Kinney, W.R. 1987. Attention-directing analytical review using accounting ratios: A case study. *Auditing: A Journal of Practice & Theory* 6 (Spring): 59-73.

KPMG Consulting. 1999. *Value-Based Management: The growing importance of shareholder value in Europe*. London: KPMG Consulting.

Libby, R. 1981. *Accounting and Human Information Processing: Theory and Applications*. Englewood Cliffs, NJ: Prentice-Hall.

Libby, R., and H. Tan. 1992. A note on modeling the determinants of audit expertise. Unpublished working paper, Cornell University.

Lien, Y., and P. Cheng. 2000. Distinguishing genuine from spurious causes: A coherence hypothesis. *Cognitive Psychology* 40 (2): 87-137.

Luft, J., and M. Shields. 2001. Why does fixation persist? Experimental evidence on the judgment performance effects of expensing intangibles. *The Accounting Review* (October): 561-587.

Maines, L., E. Bartov, P. Fairfield, E. Hirst, T. Iannaconi, R. Mallett, C. Schrand, D. Skinner, and L. Vincent. 2002. Recommendations on disclosure of nonfinancial performance measures—Commentary. *Accounting Horizons* 16 (December): 353-362.

Matute, H., F. Arcediano, and R. Miller. 1996. Test question modulates cue compeititon between causes and between effects. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 22: 182-196.

Miller, R., and H. Matute. 1996. Animal analogues of causal judgments. In D.R. Shanks, K.J. Holyoak, and D.L. Medin (Eds.), *The psychology of learning and motivation*: Vol. 34. Causal Learning. San Diego: Academic Press: 133-166.

Murphy, G., and D. Medin. 1985. The role of theories in conceptual coherence. *Psychological Review* 92: 289-316.

Nelson, M. 1993. The effects of error frequency and accounting knowledge on error diagnosis in analytical review. *The Accounting Review* (October): 804-824.

Ong, L.K. 2003. Does adding nonfinancial value drivers to a summary financial measure improve the learning and performance of managers? Working paper, University of Southern California.

Reber, A. 1976. Implicit learning of synthetic languages: The role of instructional set. *Journal of Experimental Psychology: Human Learning and Memory* 2: 88-94.

Reber, A., S. Kassin, S. Lewis, and G. Cantor. 1980. On the relationship between implicit and explicit modes of learning a complex rule structure. *Journal of Experimental Psychology: Human Learning and Memory* 6: 492-502.

Rucci, A., S. Kirn, and R. Quinn. 1998. The employee-customer-profit chain at Sears. *Harvard Business Review* 76 (1): 82-97.

Schlesinger, L., and J. Heskett. 1991. The service-driven service company. *Harvard Business Review* (September-October): 1-20.

Shanks, D. R., Pearson, S. M., & Dickinson, A. 1989. Temporal contiguity and the judgment of causality by human subjects. *Quarterly Journal of Experimental Psychology Section B-Comparative and Physiological Psychology*, 41(2): 139-159.

Trabasso, T., and L. Sperry. 1985. Causal relatedness and importance of story events. *Journal of Memory and Language* 24: 595-611.

Tversky, A., and D. Kahneman. 1980. Causal schemas in judgments under uncertainty. In M. Fishbein (Ed.), *Progress in Social Psychology*, Vol. 1 (pp. 49-72). Hillsdale, NJ: Erlbaum.

Vera-Muñoz, S. 1998. The effects of accounting knowledge and context on the omission of opportunity costs in resource allocation decisions. *The Accounting Review* 73 (January): 47-72.

Waldmann, M. 2001. Predictive versus diagnostic causal learning: Evidence from an overshadowing paradigm. *Psychonomic Bulletin & Review* 8: 600-608.

Waldmann, M. 2000. Competition among causes but not effects in predictive and diagnostic learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 28 No. 1: 53-76.

Waldmann, M. 1996. Knowledge-based causal induction. In D.R. Shanks, K.J. Holyoak, and D.L. Medin (Eds.), *The psychology of learning and motivation*: Vol. 34. Causal Learning. San Diego: Academic Press: 47-88.

Waldmann, M., and Y. Hagmayer. 2001. Estimating causal strength: the role of structural knowledge and processing effort. *Cognition* 82: 27-58.

Waldmann, M., and K. Holyoak. 1992. Predictive and diagnostic learning within causal models: Asymmetries in cue competition. *Journal of Experimental Psychology: General* 121: 222-236.

Waldmann, M., K. Holyoak, and A. Fratianne. 1995. Causal models and the acquisition of category structure. *Journal of Experimental Psychology: General* 124: 181-206.

Waller, W., and W. Felix. 1987. Auditors' covariation judgments. *The Accounting Review* (April): 275-292.

Waller, W., and W. Felix. 1989. Auditors' causal judgments: Effects of forward vs. backward inference on information processing. *Accounting, Organizations and Society* 14: 179-200.

Wasserman, E., and D. Neunaber. 1986. College students' responding to and rating of contingency relations: The role of temporal contiguity. *Journal of Experimental Analysis of Behavior* 46: 15-35.

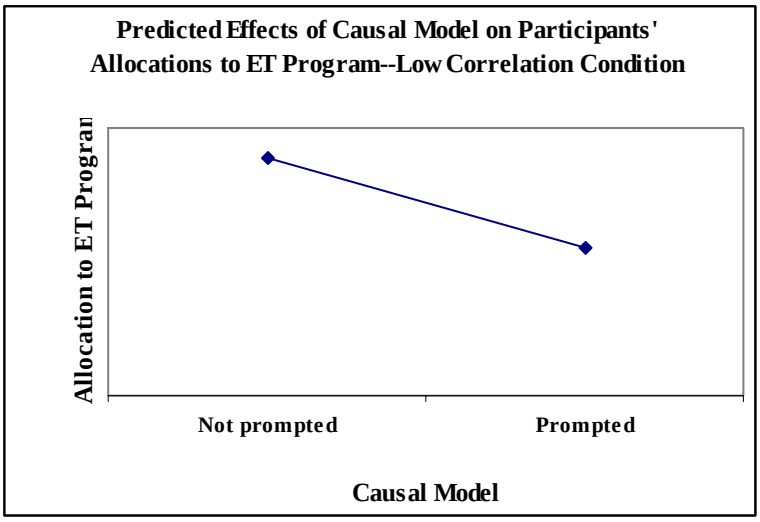
Wisniewski, E., and D. Medin. 1994. On the interaction of theory and data in concept learning. *Cognitive Science* 18: 221-281.

Wright, J., and G. Murphy. 1984. The utility of theories in intuitive statistics: The robustness of theory-based judgments. *Journal of Experimental Psychology: General* 133: 301-322.

FIGURE 1

Predicted Effects of Causal Model on Participants' Monetary Allocations to the Employee Training (ET) Program

Panel A. Predicted Effects in the Low Correlation Condition (H1a)



Panel B. Predicted Effects in the High Correlation Condition (H1b)

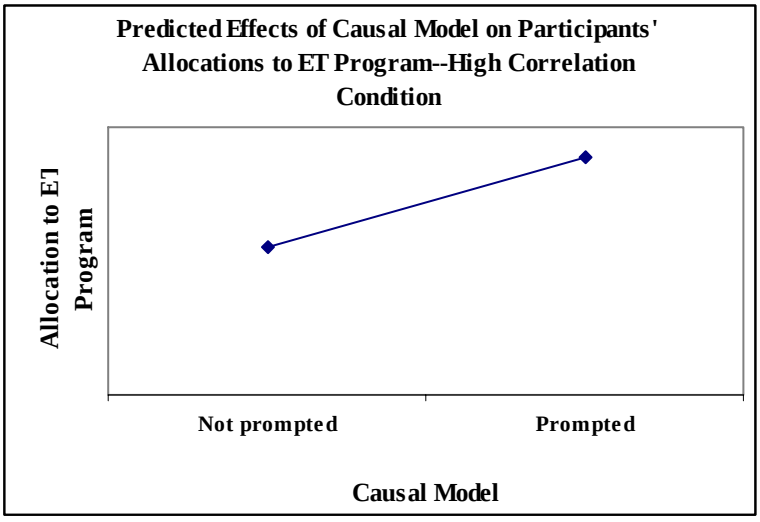
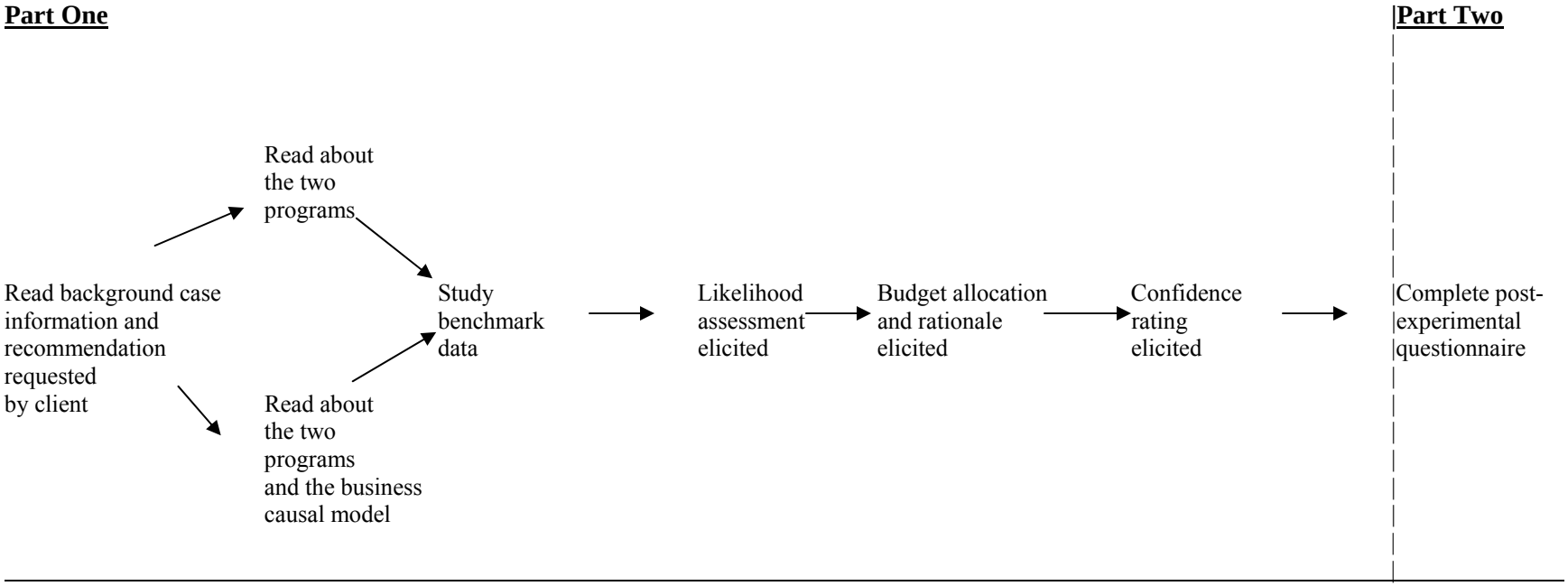


FIGURE 2
Experimental Procedures



A = No cause-and-effect business model provided
B = Cause-and-effect business model provided

FIGURE 3

Observed Effects of Causal Model on Participants' Monetary Allocations to the Employee Training (ET) Program

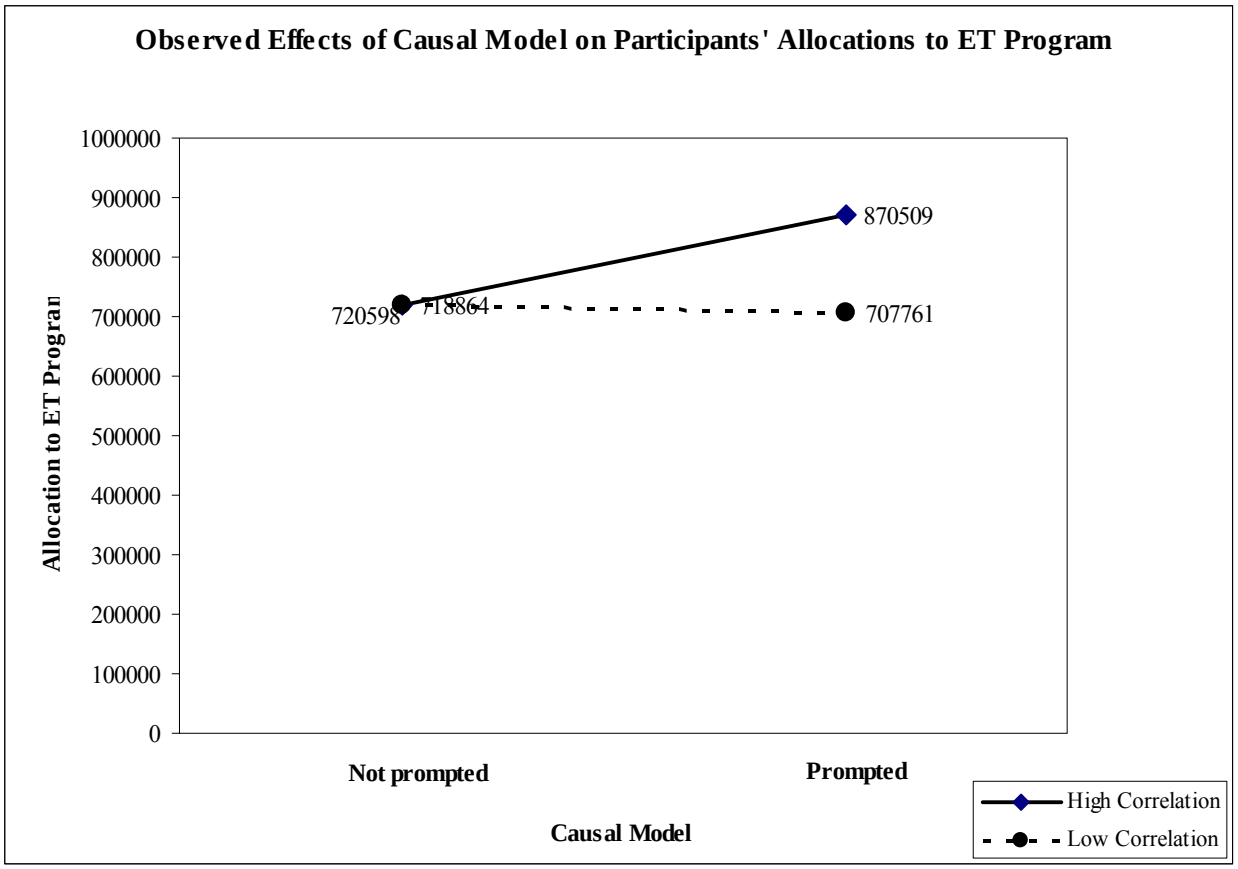


TABLE 1
Descriptive Statistics and Hypotheses Tests
(n = 86)

Panel A. Mean Likelihood Assessments and Monetary Allocations to the Employee Training Program (standard deviations)

	Cause-and-Effect Business Model[b]		
ρ(ET, π)[c]	*Not Prompted*	*Prompted*	*Total*
Low	1 [46.38% (18.19)][a] 720,598 (289,917) n = 21	2 [48.14% (20.16)] 707,761 (304,203) n = 23	[47.30% (19.49)] 713,888 (294,078) n = 44
High	3 [51.41% (17.09)] 718,864 (257,515) n = 22	4 [50.12% (17.36)] 870,509 (240,918) n = 20	[50.80% (17.02)] 791,076 (258,339) n = 42
Total	[48.95% (18.11)] 719,710 (270,523) n = 43	[49.06% (18.71)] 783,458 (285,441) n = 43	[49.01% (18.30)] 751,584 (278,294) n = 86

[a] The percentages [in brackets] in the first row of each cell on Panel A denote the mean likelihood assessments of participants in response to the question of what percentage of a fixed increase in earnings would likely be accounted for by the ET and OP programs. The percentage reported here refers to the ET program. The second and third rows of each cell denote the mean monetary allocations to the ET program and the standard deviations, respectively.

[b] Participants who were prompted with the causal model received a business model articulated as a cause-and-effect theory linking ET program expenditures to improvements in customer satisfaction and earnings. Participants who were not prompted with the causal model received only summary information about the ET program and related non-financial measures.

[c] Participants in the high correlation condition received benchmark data with a high implicit correlation between ET program expenditures and earnings (denoted as π). In this condition, the data support the causal model. Participants in the low correlation condition received benchmark data with a low implicit correlation between ET program expenditures and earnings. In this condition, the data do not support the causal model.

TABLE 1 (continued)
Descriptive Statistics and Hypotheses Tests
(n = 86)

Panel B. Planned Comparisons for Effects of Causal Model on Likelihood Assessment and Monetary Allocation to the Employee-Training Program

Planned Comparison	*Dependent Variables*	*MS*	*F*	*p*
Low correlation:				
Causal model prompted vs not prompted (H1a: cell 1 vs. cell 2)	Likelihood assessment	34.10	0.10	0.754
	Allocation to ET	1808858461.00	0.02	0.439[a]
High correlation:				
Causal model prompted vs not prompted (H1b: cell 3 vs. cell 4)	Likelihood assessment	17.41	0.05	0.823
	$Allocation to ET	240920000.00[b]	3.18	0.041[a]
Residual (full MANCOVA)	Likelihood assessment	343.44		
	$Allocation to ET	75759676281.00		

Panel C. Planned Comparisons for Effects of Correlation on Likelihood Assessment and Monetary Allocation to the Employee-Training Program

Planned Comparison	*Dependent Variables*	*MS*	*F*	*p*
Causal model not prompted:				
Low vs. High correlation (cell 1 vs. cell 3)	Likelihood assessment	271.64	0.79	0.379
	Allocation to ET	32304500.65	0.00	0.492[a]
Causal model prompted:				
Low vs. High correlation (cell 2 vs. cell 4)	Likelihood assessment	41.79	0.12	0.729
	$Allocation to ET	283350000.00[b]	3.74	0.030[a]
Residual (full MANCOVA)	Likelihood assessment	343.44		
	$Allocation to ET	75759676281.00		

[a] one-tailed
[b] 000's omitted

Appendix 1
Causal Model Manipulation

Panel A: Condition where participants were not prompted with the causal model

The on-line grocery industry uses two key measures to assess employee training, namely, the employee turnover rate and employee complaints. The employee turnover rate is measured as the proportion of employees leaving the company in any year. Employee complaints are measured as the number of complaints per every 1,000 employees in any year. FG's goal is to lower the average employee turnover rate from 35 percent in 1999 to less than 20 percent by 2002, and to lower average employee complaints from 30 (per every 1,000 employees) in 1999 to less than 10 by 2002.

Panel B: Condition where participants were prompted with the causal model

In addition to the paragraph above, participants received the following cause-and-effect model and narrative:

The Employee training → Customer Satisfaction → Earnings Chain

Proprietary data collected by Forrester Research about the economics of employee training and loyalty highlight the critical role that employees play in enhancing or diminishing customer satisfaction, and therefore, earnings. Although their research does not include data from on-line grocery companies, it provides evidence that customer satisfaction is rooted in employee training. In other words, employee training and customer satisfaction are directly correlated. This underscores the importance of focusing on the employees who actually create or deliver the things that customers value. For example, Forrester's data show that in stores that were given relatively high customer-service ratings, 54 percent of the employees turned over in a year compared with 83 percent at the poorer scoring stores. Furthermore, evidence from companies that are mounting innovative efforts to measure the full costs of employee turnover adds to the impact of these findings. For example, two divisions at a well-known on-line bookstore company undertook a study to quantify the links among employee turnover, customer retention, and profitability. Management estimated that a 10 percent reduction in employee turnover would reduce customer turnover by 1 percent to 3 percent and raise revenues by $50 million to $150 million. **The study's conclusions are striking: even with high-end estimates for recruitment and training costs and low-end estimates for the cost of lost customers, reducing employee turnover yielded increases in profits over time that were significantly greater than the cost savings generated from programs implemented to enhance the operational efficiency of the stores. [The last sentence is unbolded in the experimental instruments].**

Appendix 2
Correlation Manipulation
(Excerpt from N = 6 on-line grocers; 000's omitted)[1]

Panel A: Low Correlation Condition

1	12/31	12/31	12/31	12/31	12/31
	Year 0	Year 1	Year 2	Year 3	Year 4
Employee training[2]	68				
Operating process enhancement	154				
Order accuracy		92.7%	95.5%	96.7%	97.0%
Orders filled / hour		8.5	8.8	9.0	9.0
BizRate Score [3]		8.3	8.4	8.6	8.8
Employee Turnover		27.8%	27.4%	26.6%	26.4%
Employee Complaints[4]		24.3	23.4	23.1	23.0
Sales		17,172	24,498	36,282	49,952
Cost of Goods Sold		12,205	17,573	27,491	37,396
Gross Margin		4,967	6,925	8,791	12,556
Operating Expenses		1,719	1,928	2,489	4,494
EBITDA[5]		3,248	4,997	6,302	8,062

2	12/31	12/31	12/31	12/31	12/31
	Year 0	Year 1	Year 2	Year 3	Year 4
Employee training	253				
Operating process enhancement	110				
Order accuracy		94.8%	95.1%	96.2%	96.8%
Orders filled / hour		9.6	10.1	10.3	10.7
BizRate Score		9.0	9.1	9.2	9.5
Employee Turnover		22.8%	22.3%	21.0%	20.8%
Employee Complaints		22.0	22.4	22.4	22.3
Sales		50,322	65,533	85,747	109,836
Cost of Goods Sold		38,081	48,571	63,070	81,172
Gross Margin		12,241	16,962	22,677	28,664
Operating Expenses		8,033	12,265	14,896	16,922
EBITDA		4,208	4,697	7,781	11,742

[1] Data do not include all results from operations, just on-line grocery shopping.
[2] Spending in the employee training program occurred throughout the four-year period; spending in the operating process enhancement program occurred at the end of year 0.
[3] Based on a 10-point scale (1 = poor; 10 = excellent)
[4] Per every 1,000 employees
[5] EBITDA is Earnings before Interest, Taxes, Depreciation, and Amortization

Appendix 2-Continued
Correlation Manipulation
(Excerpt from N = 6 on-line grocers; 000's omitted)[6]

Panel B: High Correlation Condition

1	12/31 Year 0	12/31 Year 1	12/31 Year 2	12/31 Year 3	12/31 Year 4
Employee training[7]	68				
Operating process enhancement	154				
Order accuracy		92.7%	95.5%	97.2%	97.9%
Orders filled / hour		8.5	8.8	9.0	9.0
BizRate Score [8]		8.3	8.4	8.4	8.5
Employee Turnover		33.0%	31.4%	30.8%	29.9%
Employee Complaints[9]		24.3	23.4	24.2	23.1
Sales		6,299	9,086	27,045	41,372
Cost of Goods Sold		4,710	6,656	20,486	30,023
Gross Margin		1,589	2,430	6,559	11,349
Operating Expenses		928	1,259	2,591	3,590
EBITDA[10]		660	1,171	3,968	7,759

2	12/31 Year 0	12/31 Year 1	12/31 Year 2	12/31 Year 3	12/31 Year 4
Employee training	253				
Operating Process Enhancement	110				
Order accuracy		94.4%	95.2%	95.8%	96.2%
Orders filled / hour		9.6	10.4	10.8	11.1
BizRate Score		9.0	9.1	9.4	9.8
Employee Turnover		18.9%	18.9%	17.0%	16.5%
Employee Complaints		20.4	19.7	18.8	17.1
Sales		25,122	67,254	97,291	125,518
Cost of Goods Sold		19,313	50,589	73,486	95,239
Gross Margin		5,809	16,665	23,804	30,279
Operating Expenses		2,983	8,912	10,813	14,703
EBITDA		2,826	7,753	12,991	15,576

[6] Data do not include all results from operations, just on-line grocery shopping.
[7] Spending in the employee training program occurred throughout the four-year period; spending in the operating process enhancement program occurred at the end of year 0.
[8] Based on a 10-point scale (1 = poor; 10 = excellent)
[9] Per every 1,000 employees
[10] EBITDA is Earnings before Interest, Taxes, Depreciation, and Amortization