Does the Communication of Causal Linkages Improve Effort Allocations? An Experimental Investigation Based on Melioration Theory

Anne M. Farrell
University of Illinois at Urbana-Champaign
Champaign, Illinois 61820
amf@uiuc.edu

Kathryn Kadous
Emory University
Atlanta, Georgia 30322
kkadous@emory.edu

Kristy L. Towry
Emory University
Atlanta, Georgia 30322
Kristy_Towry@bus.emory.edu

April 2007

We thank the Research Computing staff in the Goizueta Business School at Emory University (especially Lisa Harris) for development of the program used in the experiment. We thank Gary Hecht, Steve Kachelmeier, Susan Krische, Mark Peecher, Marcel van Rinsum, and doctoral students at the University of Illinois for providing helpful comments on a prior draft. Finally, we thank the Lost Dog Cafe in Arlington, VA and Zingerman's in Ann Arbor, MI for allowing us to use their creative and extensive sandwich menus in our experimental task.

Does the Communication of Causal Linkages Improve Effort Allocations?
An Experimental Investigation Based on Melioration Theory

Abstract

Proponents of strategic performance measurement systems such as The Balanced Scorecard argue that such systems can improve employee performance, especially when causal linkages among performance measures are provided. We use a multi-period production task to experimentally examine whether communication about causal linkages can improve effort allocations and payoffs. We base our investigation on melioration theory, which predicts that employees will allocate effort as though the causal linkage were not there, a behavior called melioration. Melioration theory implies that employees will over-invest in effort that increases current period pay, while sub-optimizing overall pay. Absent communication about causal linkages, we find that one-third of participants allocate effort consistent with melioration. We show that melioration is most prevalent among participants who do not come to the task with an accurate implicit theory about the causal linkage. We also show that communicating directional information about the linkage is sufficient to virtually eliminate melioration, significantly improving effort allocations and payoffs. Communicating quantified information about the linkage is no more beneficial than communicating directional information. Our results have implications for firms examining the benefits and costs of developing, validating, and quantifying causal business models, and provide important evidence about how melioration theory can be extended to a rich accounting context.

Keywords: Performance Measurement, Effort Allocation, Business Model, Causal Linkages, Melioration Theory

Data Availability: Contact the authors.

Does the Communication of Causal Linkages Improve Effort Allocations?
An Experimental Investigation Based on Melioration Theory

I. Introduction

Strategic performance measurement systems such as The Balanced Scorecard have received increased attention in the last decade. Proponents of these systems claim that they are effective tools for communicating and implementing corporate strategy (Simons 2000; Ittner and Larcker 2003). Some have argued that laying out the hypothesized directional and intertemporal linkages among performance measures (through the use of strategy maps, for example) is critical to successfully implementing a strategic performance measurement system (Kaplan and Norton 2000; Tayler 2006). While identifying directional linkages may be a relatively low-cost endeavor, validating and quantifying linkages in causal business models is costly. Therefore, it is important to understand the conditions under which those costs are likely to be rewarded with improved employee effort allocations. The purpose of this study is to examine whether communicating directional and quantified causal linkages to employees working under a goal-aligning incentive contract improves effort allocation decisions.

Judgment and decision making theory provides insight into why communication of linkages might be valuable. Specifically, research demonstrates that individuals find it difficult to identify appropriate decision strategies in multi-period settings in which one period’s choice affects the value of future outcomes. For example, an employee allocating effort to quality of production versus quantity of production over several work periods would have to consider the impact of the current period’s allocation on future payoff distributions in order to choose the pattern of effort allocations that would maximize total payoffs. However, when faced with this sort of multi-period decision problem, individuals often behave as if they are ignoring the “internality” (that is, the impact of current decisions on the utility of future decisions), and instead choose actions that maximize the payoff in the current period at the expense of the overall payoff across periods (Herrnstein 1990; Herrnstein and Prelec 1991, 1992; Mainwaring 1997). This behavior is referred to as "melioration." We expect that communicating information about causal linkages between current actions and future payoffs will make the internality more salient, and therefore will facilitate employees' choice of an appropriate effort allocation.

We conduct an experiment to examine whether melioration occurs in a production setting and, if so, whether communicating causal linkages improves employees' effort allocations. In our experiment, participants acting as production employees must choose how to allocate their effort between production quantity and quality. For our experimental firm, the quantity of output determines current period revenue (i.e., quantity is a contemporaneous indicator of financial performance), and there is an internality. Specifically, there is a causal link between the quality of products offered to customers today and product prices tomorrow (i.e., quality is a leading indicator of future financial performance). Participants work for the same firm for multiple work periods, and receive pay equal to 5% of the firm's revenue. Thus employees' long-term goals are aligned with those of the firm's owners.[1] In our setting, melioration theory predicts employees will over-allocate effort to actions that increase quantity (and current pay), and under-allocate effort to actions that increase quality (and overall pay).

We manipulate communications about causal linkages at three levels. In the *no communication* condition, we do not mention the causal linkage between current quality and future price. In the *directional linkage* condition, employees are told only that current quality impacts future prices. In the *quantified linkage* condition, employees receive full, quantified information about this link. Importantly, all of our participants receive feedback each period on

[1] We use revenue instead of profitability as our measure of firm financial performance to simplify the experimental setting. This choice is reasonable if we assume that costs vary proportionally to revenue.

quality, quantity, price, revenue, and pay, analogous to a firm that reports multiple performance measures through the use of a strategic performance measurement system.

We expect that communication of causal linkages will reduce melioration, thus increasing employees' allocation of effort to quality. Notably, prior studies have found that providing the information needed to maximize outcomes has not prevented melioration (Herrnstein et al. 1993). However, our task is different from those of prior work in a critical way. Prior experiments typically involve repeated trials of a context-free task in which participants must learn from experience whether to choose one discrete action or another, such as "left button" or "right button" (e.g., Herrnstein et al. 1993; Tunney and Shanks 2002). Learning without a context or theory is difficult under the best of circumstances (e.g., Libby 1981), and the presence of a causal link between current actions and future payoffs (i.e., an internality) in a multi-period choice task often results in individuals "learning" a poor strategy (Herrnstein et al. 1993). In production settings, employees are likely to have an implicit theory that quality today affects firm financial performance (and employee pay) tomorrow, which could potentially mitigate suboptimal behavior and facilitate learning of an appropriate strategy. Therefore, we examine whether the failure of full information to eliminate melioration in prior studies extends to a rich accounting context to which individuals are likely to bring reasonably accurate implicit theories about the internalities.

Despite our rich context, we find that, absent communication about causal linkages, over one-third of our participants meliorate. These employees over-allocate their efforts to quantity at the expense of quality, and their allocations move farther from optimal over time. While direct comparisons are difficult, the level of melioration appears to be lower than the level observed in prior context-free studies. Our analysis suggests that this difference is due to the fact that our added context allowed some participants to use implicit theories about the effect of current allocations to quality on future financial performance, and that these participants were less likely to meliorate than were those participants who did not have such theories. Thus, the suboptimal behavior resulting from melioration appears less likely to occur in rich settings such as production ones in which individuals likely have implicit theories about the effect of current choices on future payoffs than in the more abstract settings of prior melioration studies. Yet, we note that even in our production setting, suboptimal decision behavior occurs.

Results also show that communicating a directional linkage (only) between quality today and financial performance tomorrow results in a significant reduction in meliorating behavior, significantly increasing effort allocated to quality actions and overall participant earnings. Notably, communicating a quantified linkage results in virtually no increase in quality efforts above those in the directional linkage condition. This suggests that relating only directional causal linkages, such as those illustrated in a strategy map, may be sufficient to promote optimal behavior in a setting for which individuals are likely to have reasonably accurate implicit theories.

This paper contributes to the literature in several ways. From a theoretical perspective, while much of the prior accounting literature on strategic performance measurement systems suggests that hypothesized and/or validated business models facilitate employee decision-making (Eccles 1991; Copeland et al. 1996; Kaplan and Norton 1996, 2000, 2001; Young and O'Byrne 2001), there is limited evidence regarding the actual benefits of such systems in guiding employee behavior. We address this issue by providing evidence on the extent to which communicating causal linkages to employees whose goals are already aligned with those of the firm improves employee effort allocations.

We contribute to the judgment and decision making literature by extending melioration theory. In particular, we demonstrate that the meliorating behavior and suboptimal performance found in prior research using abstract, discrete-choice tasks still occur (though perhaps at a lower rate) in a richer setting in which employees have reasonably accurate implicit theories about relevant internalities. We also demonstrate that melioration can be virtually eliminated in such settings by communicating directional information about the internality.

From a practical perspective, we find that, even in a setting in which employees likely have implicit theories about causal linkages, communicating directional linkages facilitates employee decision-making as compared to communicating no information. However, we find no evidence that communicating specific, quantified information about causal linkages facilitates employee decision-making any more so than does communicating only directional information. Firms can use this information in assessing the potential benefits and costs of developing, validating, and quantifying causal business models. Further, this result suggests that simple mechanisms that communicate hypothesized, directional models (e.g., face-to-face communications, printed and electronic newsletters; see Sinickas 2006) may be effective means of implementing corporate strategy.

The remainder of this paper is organized as follows. Section II explains the relevant theories and develops our hypotheses. Sections III and IV describe the experiment and results. We summarize and conclude in Section V.

II. Literature Review and Hypothesis Development

Practitioner literature has long argued that providing information about cause-and-effect linkages among performance measures can make goal-congruent action choices more evident to employees (Eccles 1991; Copeland et al. 1996; Kaplan and Norton 1996, 2000, 2001; Young and

O'Byrne 2001). However, research has so far provided only limited insight into this issue. For example, prior studies have examined the effects of strategic performance measurement systems on performance evaluations (Lipe and Salterio 2000, 2002; Banker et al. 2004, 2006; Kaplan and Wisner 2004; Libby et al. 2004; Roberts et al. 2004; Cardinaels and Van Veen-Dirks 2005; Dilla and Steinbart 2005; Haywood and Stuart 2006; Petersen and Samuels 2006) and auditors' judgments (Salterio et al. 2005; Vera-Muñoz et al. 2006), but have not directly studied the impact on employees' effort allocation choices.

The prior research indicates that the communication of causal linkages has the potential to affect decisions. For example, in a performance evaluation setting, Banker et al. (2004) provided experimental participants with strategy narratives and pictorial business models in an effort to make them more aware of causal linkages. They report that these participants' performance evaluations relied more on strategically-linked performance measures than did the evaluations of participants who did not have the narratives or business models. Similarly, Tayler (2006) demonstrates that framing a Balanced Scorecard as a causal chain can mitigate managerial motivated reasoning. Vera-Muñoz et al. (2006) show that articulation of causal linkages can enhance auditors' evaluations of benchmark data. Prior research has also considered whether contracting on linkages is beneficial. Farrell et al. (2007) demonstrate that contracting on forward-looking measures improves employees' effort allocations and performance over merely communicating those linkages, even when employees' incentives are aligned with the firm's. Our paper extends this body of research to examine whether communication of causal linkages affects the effort allocation decisions of employees. To develop hypotheses, we turn to prior literature on melioration theory and we extend that theory to a production task.

Melioration Theory and its Applicability to Realistic Tasks

Melioration is a descriptive theory of individual decision-making in distributed choice tasks. Distributed choices are those that are spread over time, rather than made at a single point in time (Herrnstein and Prelec 1992 236). For example, the decision about how much to exercise over a lifetime is distributed, in that each day one decides whether to exercise that day. Distributed choices typically involve an internality, meaning that current period choices affect future payoffs. For example, if one exercises this week, exercising next week is less aversive and results in better health rewards. From a normative perspective, distributed choices can be made by calculating the utility of all possible distributions of choices, then selecting the distribution offering the highest total utility. However, prior research has demonstrated that individuals have difficulty attaining optimal outcomes in distributed choices.

When faced with a distributed choice with an internality, individuals tend to choose actions that maximize the payoff in the current period at the expense of the overall payoff across all periods (Herrnstein 1990; Herrnstein and Prelec 1991, 1992; Mainwaring 1997). Herrnstein and Prelec (1992) argue that this occurs because individuals make distributed choices according to the principles of value accounting and melioration. Individuals use the cognitively simplified strategy of keeping track of the average return received per unit invested in each alternative (value accounting), and then shifting their behavior toward alternatives that have provided a higher per-unit return (melioration). In short, suboptimal payoffs in a distributed choice task occur because individuals behave as though they ignore the task's internality.[2]

[2] At first blush, the behavior implied by melioration theory may seem similar to that implied by a simple model of time discounting. However, the two concepts are different in at least two important ways. First, time discounting relates simply to a *preference* for earlier consumption over later consumption, whereas melioration theory describes a *process* (i.e., the value accounting heuristic) through which decisions are made. Second, time discounting can occur within the context of a one-time (as opposed to distributed) choice problem, whereas melioration theory relates specifically to distributed choices. That is, a decision-maker engaging in time discounting may consider

For example, consider an individual choosing between hamburger and caviar each day for lunch, where the cost of the two items is identical (e.g., Herrnstein et al. 1993). The utility gained from eating hamburger is the same regardless of frequency of consumption, while that for caviar is higher when caviar is consumed less often; in other words, the utility gained from eating caviar arises in part because caviar is a rare treat. Under these conditions, when choosing what to eat, a utility-maximizing individual would consider not just the current utility gained from hamburger and caviar, but the negative effect of today's consumption of caviar on the utility gained from its future consumption. A maximizer would thus choose a sequence of hamburger and caviar consumption that equates the marginal utilities of hamburger and caviar, thus rationing caviar to preserve its value as a rare treat. Thus, for a maximizer, the long-run average utility of caviar consumption would exceed that of hamburger consumption. On the other hand, a meliorating individual would choose as though current consumption has no impact on the utility of future choices. Accordingly, a meliorator would choose in each period whichever food would lead to the highest current-period utility. This would lead to a sequence of consumption for which the long-run average utility from caviar and hamburger are equal – a point at which too much caviar is being consumed and utility is not maximized.

An interesting feature of melioration is that it is not mitigated through experience. Meliorators make decisions one period at a time, and the period-by-period outcomes of those decisions tend to reinforce suboptimal decisions, preventing learning. Returning to the example, as long as the prior-period utility from consuming caviar is higher on average than the prior-period utility from consuming hamburgers, feedback from choosing caviar will be more positive

---

his/her preferences for earlier consumption in order to maximize overall utility. A meliorator, on the other hand, explicitly maximizes utility each period rather than maximizing overall utility (across all periods). For further discussion of the distinction between time discounting and melioration, see Herrnstein and Prelec (1991).

than feedback from choosing hamburgers. Thus, feedback will reinforce the decision to choose caviar too often, and learning does not occur.

We examine a multi-period setting in which employees who are paid based on firm performance (i.e., employee and firm goals are aligned) must choose how to allocate their efforts between actions that increase the quantity of output and those that increase quality. Employees receive feedback each period on the quantity and quality of outputs, as well as on their pay and firm performance. There is a causal link in the operating environment between quality today and prices tomorrow (i.e., an internality) such that the firm's financial performance (and thus the employee's overall pay) is maximized when production quantity is maximized, subject to the constraint that quality is sufficiently high to ensure favorable price movements. Thus, in our setting, melioration theory predicts that employees will over-allocate effort to actions that increase quantity (and current pay), and under-allocate effort to actions that increase quality (and overall pay). We do not expect all employees will meliorate, for reasons provided below; however, we expect to observe a significant amount of melioration among employees receiving no communication about the causal linkage. Our first hypothesis is stated in alternative form below.

Hypothesis 1: Absent communication of information about causal linkages among performance measures, employees will meliorate.

Although the body of evidence supporting suboptimal decision-making in distributed choice tasks is impressive, encompassing both animal and human studies (e.g., Herrnstein and Prelec 1991; Herrnstein et al. 1993; Kachelmeier and Granof 1993; Antonides and Maital 2002; Tunney and Shanks 2002; Brenner and Witt 2003), there is some question as to whether we will observe melioration in our production setting. Previous tests of melioration have used context-free tasks. In such tasks, individuals cannot bring their real-world knowledge to bear and can only learn about the internality from experience with the task. While practice rounds may help individuals assess the current values of the alternatives, they are unlikely to result in accurate knowledge about the internality (Herrnstein et al. 1993). In contrast, individuals may be aware of important internalities in richer tasks (Luft 2004). In a realistic production setting, even employees who are not provided with explicit information about the causal linkage likely have an implicit theory that product quality today has implications for firm performance tomorrow. Herrnstein et al. (1993) suggest that when individuals are aware of potential internalities, maximization may be facilitated and melioration behavior may be reduced. We examine whether melioration theory extends to a rich accounting context, in which participants are likely to have reasonably accurate implicit theories.

Assuming that we find evidence of melioration when specific information about the causal linkage is absent, we next address whether communicating the causal linkage either directionally or in quantified form will reduce melioration and improve employees' effort allocations. Information about causal linkages should facilitate a more normative decision process because employees are either provided with a causal theory or they have their own implicit causal theory verified. Therefore, they would not have to rely solely on period-by-period feedback, which can be misleading in distributed tasks, in order to learn about the internality. However, communication of information necessary to maximize payoffs has not prevented participants in prior studies from meliorating. Herrnstein et al. (1993) show melioration in binary choice tasks, and examine the effect of providing information about the internality on meliorating behavior (experiment 5). They find that melioration is the prevalent behavior of participants receiving information indicating the potential existence, but not the direction, of the internality. Even information that fully describes the internality results in average decision behavior that is far from optimal: only 25% of the participants receiving full information approximate optimal behavior. We argue, though, that the failure of this information to eliminate melioration is attributable to the abstract setting in which the experiments are conducted. The lack of context likely makes it difficult for participants to use the information about the internality in developing a strategy.

In our setting, as in real-world production settings, we expect that employees come to their tasks with an implicit theory about the causal linkage (i.e., the internality). We expect them to believe that current quality affects future prices and to understand how and why this would be true. Of course, even with implicit theories about relations in the environment, maximizing long-term performance is a cognitively difficult task that involves the analysis of the intertemporal costs and benefits of different actions. By communicating to employees information that validates their implicit theories, firm owners can reduce the cognitive efforts required of employees, eliminating the need for employees to test their theories against feedback and adjust their effort allocation accordingly. Thus, we expect that melioration will be relatively easier to eliminate in our rich context because information about causal linkages will be easily incorporated into existing knowledge structures, reducing the need to rely on feedback. Moreover, we expect that more specific information about the causal linkage is likely to be more useful than less specific information in this regard. Based on this reasoning, we posit the following:

Hypothesis 2: Employees will meliorate most when no information about causal linkages is communicated, a smaller amount when directional linkages are communicated, and least when quantified linkages are communicated.

III. Research Design

Our experiment examines whether participants acting as employees exhibit behavior consistent with melioration in the absence of communication about either directional or quantified causal linkages among performance measures, and whether such communications reduce or eliminate melioration. We employ a production task in which participants must trade off the quantity and quality of goods they produce. Thus, in each work period, participants must determine how much of their limited effort to allocate to actions that increase output quantity and how much to allocate to actions that increase output quality. Participants are paid a percentage of firm revenue. Current period pay depends only on the quantity of units produced and the current period price of the good; however, future prices depend on the quality of current period production, such that very high quality this period results in higher prices in future periods and very low quality this period results in lower prices in future periods.

Participants

Seventy-seven undergraduate students enrolled in upper-level accounting courses participated in the experiment. Eighty-one percent of the participants had taken three or more economics courses, and 87% had taken three or more accounting courses; males comprised 48% of the participant pool. There were no differences across experimental conditions in the mean number of economics and accounting courses taken or in gender composition (χ^2s < 3.59, all $p >$ 0.17). Participants were paid based on their performance, as described below. Participant pay ranged from $5.98 to $43.18, with a mean of $27.19.

Experimental Design

Participants acted as employees making sandwiches to order for a virtual sandwich shop. The experiment used a 3 (between subjects) x 12 (within subjects) design. The between-subjects independent variable was the nature of communications about the causal link between product quality today and product price tomorrow *(no communication, directional linkage,* or *quantified linkage*, as described below). The within-subjects variable, *work period*, represented the number of four-minute production periods in the task. While all participants worked for 12 four-minute periods, in order to prevent end-of-task gaming, we did not tell participants the number of work periods.

All participants were informed that they would be working as a sandwich maker for the same sandwich shop for multiple work periods. Their pay was 5% of the sandwich shop's revenue generated from the sale of all sandwiches produced in each work period, and so participants' long-term goals were aligned with those of the sandwich shop.[3] Sandwich shop revenue for each period was computed as the number of saleable sandwiches made (i.e., sandwiches with fewer than three errors) times that period's selling price per sandwich. All participants received identical feedback reports at the end of each work period (as described later), and all were told that the sandwich shop had an explicit quality threshold for current period production, such that sandwiches with three or more errors as compared to a customer's order were not saleable and had to be remade.

Participants in all conditions were told that the selling price for sandwiches would be $5.00 in the first work period. The pricing function in the sandwich shop's operating environment was such that if in a given work period the average number of errors per sandwich produced was one or fewer per sandwich, the sandwich price in the next period would be 10%

---

[3] We hold incentives constant across experimental conditions to isolate the effects of our communication manipulation on melioration behavior and effort allocation. We cannot use an incentive contract that includes a measure of output quality because in our *no communication* condition, the organization does not communicate to employees that quality is a performance driver, and thus quality cannot be included in our contract across conditions. For studies that examine the effects of contracting on contemporaneous and forward-looking performance measures see, for example, Dikolli (2001) and Farrell et al. (2007).

higher; if the average was greater than one but less than two errors per sandwich, the sandwich price in the next period would remain the same; and if the average was two errors per sandwich, the sandwich price in the next period would be 10% lower.[4] This pricing function was designed such that the task strategy for maximizing total sandwich shop revenue, and thus, for maximizing the participant's pay, was to keep the average error rate at one per sandwich, and to maximize quantity given that constraint. However, a participant could always make more money in the current period by sacrificing quality and producing more sandwiches. Allocating more effort to quantity, versus quality, pushes the error rate up. Thus, a meliorator, ignoring the impact of current choices on future payoffs, would over time move increase the error rate close to the maximum saleable quality of two errors per sandwich. Note that our pricing function is symmetric (i.e., quality ranges from zero to two errors per sandwich, symmetric around the optimal strategy of one error per sandwich). This design feature is critical, because it allows us to distinguish between random errors in strategy selection (sometimes allocating too much effort to actions that increase quality and sometimes allocating too little) and the particular pattern of behavior implied by melioration theory (steadily decreasing the allocation of effort to actions that increase quality).[5]

We varied the nature of the communication about the causal relation between the quality of output in the current period and prices in the next period. Participants in the *no*

---

[4] We use average errors per sandwich of one as the threshold below which prices no longer increase to mimic diminishing (in fact, no) returns to further investments in quality beyond that threshold, consistent with prior research that finds diminishing returns to investments in customer satisfaction (Ittner and Larcker 1998) and quality (Sedatole 2003).

[5] The pricing function parameters were based on an analysis of the task performance of 43 pilot test participants. Pilot participants did not participate in the main study. To develop the pricing function, we regressed the number of sandwiches completed each period on the number of errors per sandwich, in order to quantify the tradeoff between quantity and quality. We used the resulting estimates to simulate the firm's revenue each period, comparing a high quality versus high quantity strategy. Pricing parameters were set to ensure that the pay-maximizing strategy (i.e., keeping the average error rate at one per sandwich and maximizing quantity given that constraint) resulted in higher firm revenue.

communication condition were informed only that the selling price "may rise or fall in subsequent periods"; no further information was provided about the causal mechanism determining prices (see Exhibit 1 for details of the manipulation). Note that in all conditions, participants received extensive feedback after each period (see Exhibit 2). Thus, the *no communication* condition is analogous to a firm using a strategic performance measurement system to provide various performance measures to employees, but not providing strategy maps or other information about causal linkages among these measures. This condition is also similar to conditions in most melioration studies in which participants were not provided with information about the internality and must learn it during the task using trial and error. However, unlike in prior studies, it is likely that our participants came to the task with an expectation of a causal relation between quality and price.[6]

In the *directional linkage* condition, participants were informed that "the price may rise or fall in subsequent periods, depending on the quality of sandwiches produced in prior periods." Thus, they were told that quality caused future price changes, but the exact parameters of the pricing function were not specified. This is analogous to a firm communicating information about hypothesized, directional causal linkages to employees. In the *quantified linkage* condition, participants were given full details of the pricing function as described previously. This condition is analogous to a firm communicating quantified causal linkages to employees.

Materials and Procedures

[6] Two features of the experiment increased the likelihood that participants would access their implicit theory about this internality. First, participants read the following statement during task training: "Note that your best strategy will depend on the particular sandwich shop to which you are assigned. For example, you might be successful at some shops by making as few mistakes as possible. At others, you might be more successful by making as many sandwiches as possible, with less concern for mistakes. Therefore, you should experiment with a variety of strategies." Second, each sandwich shop had an explicit quality threshold for current period production (i.e., sandwiches with three or more errors were not saleable). Importantly, neither of these features explicitly describes a causal linkage between quality today and prices tomorrow.

We conducted the experiment in a controlled laboratory environment using custom-designed software on stand-alone computers. As participants arrived, we randomly assigned them to seats at the computers and thus randomly assigned them to experimental conditions. On the first screen of the computerized task, we instructed participants to assume they made their living as sandwich makers, and as such their task was to make sandwiches ordered by a sandwich shop's customers. Customer orders would be transmitted to them via a computerized ordering system. Detailed instructions about the ordering system and the sandwich-making task followed this introduction, and then participants practiced the task during a four-minute work period. For each customer order, the computer program selected from a pool of 51 pre-programmed sandwiches. To ensure that task difficulty was the same across all experimental conditions, sandwich sequences in each work period were pre-determined and the same for each participant. A sample task screen is in Exhibit 3.

The sandwich-making task began when a customer's sandwich order appeared in the "Order" window on the computer screen. Participants used a drop-down "Menu" box to find that sandwich from among those served at the sandwich shop, and then reviewed the sandwich's ingredient list. The ingredient list disappeared when the computer mouse was moved away from the "Menu" box, although participants could review the ingredient list repeatedly. Participants assembled sandwiches by selecting ingredients from drop-down menus, and as they did so, images of the selected ingredients appeared in a production space in the center of the screen. Participants could remove ingredients from the production space by using a "Remove ingredient" button. Upon completing a sandwich, participants clicked the "Finished" button. The computer program checked the assembled sandwich against its ingredient list and added it to the production tally. The production space cleared, and the next customer order appeared in the

"Order" window. If the assembled sandwich had three or more mistakes, the sandwich was discarded. However, to prevent participants from skipping difficult sandwiches, the computer program did not advance to a new sandwich in the pre-programmed sequence; instead, the participant had to attempt the same sandwich until it met minimum quality standards.

After the instructional and practice periods, participants read details of their employment relationships with the sandwich shop and how their pay would be computed. This information included the manipulation of communications about causal linkages. In all experimental conditions, participants were told that they would work as a sandwich maker for several work periods, that they would receive their pay in cash at the end of the session, and that all sandwich shops required that any sandwich with three or more mistakes be thrown away.

Participants then took a quiz to ensure they understood how prices were determined and how revenue and pay were computed (see Exhibit 4 for quiz questions and answers). Participants could not begin the sandwich-making task until they answered all quiz questions correctly and read reinforcement explanations of the answers even if they answered questions correctly. Successful completion of the quiz and subsequent completion of the task provided assurance that participants understood their employment relationship with the sandwich shop.

Participants then worked for 12 four-minute work periods. At the end of each period, a feedback screen displayed participant pay, the total number of sandwiches made, the average errors per sandwich, a breakdown of the number of sandwiches made with zero, one, two, or three or more mistakes, and sandwich shop revenue for the work period just completed. It also displayed the per-sandwich price in the upcoming work period. Feedback was identical across experimental conditions (see Exhibit 2 for a sample feedback screen). After the twelfth work period, participants completed a post-experimental questionnaire.

IV. Results

Test of Hypothesis 1—Occurrence of Melioration

Hypothesis 1 predicts that absent specific communication about causal linkages, employees will meliorate. For our primary test of this hypothesis, we use participants' patterns of choices over time to classify them as meliorators or non-meliorators. We then examine whether the incidence of melioration is significantly greater than zero. Recall that melioration is a process through which individuals make period-by-period utility maximizing choices in a distributed choice problem rather than maximizing overall utility. In our setting, a meliorator will allocate less effort to quality than is required to maximize long-term performance. Further, experience will reinforce this suboptimal behavior, such that allocations of effort to quality will decrease over time.

Our proxy for quality efforts is *average errors per sandwich*, calculated as the simple average of the total number of errors made each period (exclusive of sandwiches thrown away) divided by the number of saleable sandwiches made that period. Lower values for *average errors per sandwich* suggest higher levels of quality efforts.[7] We classify a participant as a meliorator if *average errors per sandwich* is higher (i.e., quality efforts are lower) in the last six periods than in the first six, and if *average errors per sandwich* averages more than the optimal value of one per period over the last six periods. Such a participant is, by the second half of the task, allocating less effort to quality than is required to maximize long-term performance, and

[7] In our setting, recall that *output quality* affects future prices. In calculating *average errors per sandwich*, our measure of output quality, we exclude discarded sandwiches because they do not reach the customer, do not affect future pricing, and are therefore not an appropriate measure of future firm performance. Results are inferentially identical if we include discarded sandwiches in our calculation of *average errors per sandwich*.

over time is decreasing the allocation of effort to quality.[8] Figure 1, Panel A illustrates the trends in *average errors per sandwich* for those participants who are classified as meliorators vs. non-meliorators. This figure provides strong validation of our classification scheme. Meliorators show a strong positive trend away from the optimal of one error per sandwich and toward the maximum of two, as predicted by melioration theory, whereas non-meliorators have a consistently low error rate. Using the method of orthogonal polynomial contrasts, we find that the linear trends of meliorators and non-meliorators are significantly different ($F_{1,69} = 45.94$, $p < 0.01$).

In our *no communication* condition, 10 of 26 (38.5%) participants in this condition are classified as meliorators (Table 1, Panel A). Prior studies examined average behavior, rather than classifying individual participants as meliorators or non-meliorators. Therefore, it is difficult to statistically compare these results to those of prior studies.[9] However, casual observation would suggest that, although we observe a substantial amount of melioration, there are fewer meliorators in our setting than in previous research. For example, in describing results of a comparable condition in a binary-choice task, Herrnstein et al. (1993, experiment 5) note that most participants made nearly perfectly meliorating choices As we previously indicated, a potential explanation for this difference is that our study was conducted in a richer context that allowed participants to bring their implicit theories about internalities to bear. That is, we expected that employees would believe that quality in the current period could be related to future financial performance (and thus, to pay). To investigate this possibility, we construct a

---

[8] This is a relatively conservative measure of meliorating behavior in that if a participant's *average errors per sandwich* is trending upward (i.e., the allocation to quality is trending downward) but does not exceed the optimal level of one on average over the last six periods, the participant would be classified as a non-meliorator.

[9] Because they use abstract, binary choices, prior experiments include many (e.g., 200 or more) periods. Inferences about melioration are based on average behavior during the last half of the experimental session, by which time participants are assumed to have converged on a long-term strategy. In the contextual setting we examine, it is impractical to include many periods, and so we rely on trends indicating a convergence *toward* a long-term strategy.

proxy for the presence of an implicit theory and examine the degree to which participants who did or did not have such a theory exhibited meliorating behavior.

Our proxy for the presence of an implicit theory is constructed by calculating the *errors per sandwich* in the first period only. Participants who have a strong prior belief that quality has a positive impact on future financial performance are likely to allocate significant effort to quality at the beginning of the task. Therefore, we categorize our participants as possessing an implicit theory about the internality if they had perfect quality (*errors per sandwich* = 0) in period one. Those who had imperfect quality (*errors per sandwich* > 0) in period one are considered not to have an implicit theory.[10] In the *no communication* condition, none of the 8 participants assumed to have implicit theories were meliorators, while 10 of the 18 (55.6%) participants without implicit theories exhibit meliorating behavior. The difference in incidence of melioration is significant (Fisher's Exact Test, $p < 0.01$). Thus, it seems that unless participants have an implicit theory regarding the internality, melioration theory is descriptive of much of the behavior in our task.[11]

The existence of melioration in our setting is further borne out by an examination of the mean and distribution of *average errors per sandwich.* As noted in Table 1, Panel A, in the *no communication* condition, the mean (standard deviation of) *average errors per sandwich* is 0.94 (0.64). Recall that the optimal *average errors per sandwich* is equal to one. Further, note that if a participant were to miss that optimal by a small amount in either direction, it would be better to

[10] This is a conservative measure of those who had strong implicit theories, in that it does not allow a participant to make even one error in the first work period.

[11] An alternative measure of implicit theory yields the same inference. In the post-test, we asked participants whether the selling price each period was based on the number of mistakes made in the prior periods or on some other unknown factor (a dichotomous response). Assuming participants who indicated that future prices depend on quality had an appropriate implicit theory, 3 of 14 participants with implicit theory meliorated, while 7 of the 12 participants who did not have an appropriate implicit theory meliorated. This difference is significant (Fisher's Exact Test, $p = 0.05$). Because this question was asked after completion of the task, these results could be affected by learning.

err on the side of quality. That is, if the *errors per sandwich* in a given period is greater than one by even a tiny fraction, the price would not rise in the following period. Given this information, one might view the overall *average errors per sandwich* across work periods of 0.94 as quite close to optimal. However, an examination of the distribution of *average errors per sandwich* across all work periods and participants leads to a different conclusion.

Figure 2 (supplemented by frequency information in Table 1, Panel B) reveals that, in the *no communication* condition, *average errors per sandwich* was very infrequently close to one. Rather, the distribution is bimodal, with most work periods exhibiting either very high or very low numbers of errors. This pattern of results is consistent with the notion that the presence of an implicit theory affected participant behavior. It appears that many participants had implicit theories regarding the internality, and that these participants were likely to allocate significant effort to quality, resulting in a relatively low *average errors per sandwich* in many work periods. On the other hand, other participants did not have such an implicit theory and were more likely to succumb to meliorating behavior, resulting in a relatively high *average errors per sandwich* in many work periods.

These results support H1, which predicted meliorating behavior in the rich context of a production task. They also support our supposition that implicit theories about internalities are associated with a lower incidence of melioration. The latter idea suggests that for firms and tasks for which employees can be expected to bring reasonably accurate implicit theories to bear, communicating information about linkages may not be necessary: the combination of goal-aligning incentives, feedback on goal-relevant performance measures, and employees' specific knowledge may be sufficient to motivate high levels of goal-congruent behavior.

Test of Hypothesis 2— Reducing Melioration by Communicating Linkages

Hypothesis 2 predicts that communicating a directional linkage will decrease melioration, and that communicating a quantified linkage will decrease it even more. We test this hypothesis in two ways. First, we use the classification of meliorators and non-meliorators described above to determine whether the incidence of melioration is inversely associated with the specificity of communications about causal linkages. Recall that in the *no communication* condition, 10 of 26 (38.5%) participants were classified as meliorators. In the *directional linkage* and *quantified linkage* conditions, 2 of 25 (8.0%) and 4 of 26 (15.4%) participants, respectively, were classified as meliorators. A chi-square test confirms that the proportion of meliorators is different across the *no communication* and the *directional linkage* conditions ($\chi^2_1 = 6.57$, $p < 0.01$) as well as across the *no communication* and *quantified linkage* conditions ($\chi^2_1 = 3.52$, $p = 0.06$). However, there is no significant difference in the proportion of meliorators across the *directional linkage* and *quantified linkage* conditions ($\chi^2_1 = 0.67$, $p = 0.41$). That is, we find that communication of a directional causal linkage nearly eliminates melioration, but that communicating a quantified causal linkage provides no incremental benefit beyond that provided by communicating a directional linkage.

Our second test of Hypothesis 2 uses our quality measure, *average errors per sandwich*. The mean (standard deviation of) *average errors per sandwich* is 0.94 (0.64), 0.50 (0.42), and 0.42 (0.47) in the *no communication*, *directional linkage*, and *quantified linkage* conditions, respectively (Table 1, Panel A and Figure 1, Panel B). A Jonckheere-Terpstra test show that *average errors per sandwich* follows the predicted ordering (observed J-T statistic = 649.50, $p <$ 0.01). However, while the *average errors per sandwich* follows the expected pattern, the difference between the *directional linkage* and *quantified linkage* conditions is small. An

ANOVA (Table 1, Panel C) indicates that communication about causal linkages significantly affects *average errors per sandwich*. Consistent with results using the classification measure, pairwise comparisons reveal that the differences between the *no communication* and *directional linkage* conditions and the *no communication* and *quantified linkage* conditions are significant (both $p < 0.01$, one-tailed), but the difference between the *directional linkage* and *quantified linkage* conditions is not ($p = 0.28$, one-tailed). Both sets of results support Hypothesis 2 in that participants allocated more effort to actions that increased quality when they received communications about directional or quantified causal linkages than when they did not receive either form of communication. However, they do not support the idea that communicating more specific information is more useful than communicating less specific information. In our study, communicating quantified, as opposed to only directional, linkages had no incremental effect on participants' effort allocation choices.

Returning to the histograms of Figure 2, we can observe the effect of the two types of communication on the distribution of choices. As described earlier, in the *no communication* condition, the distribution is bimodal, with the majority of participant work periods showing either very high or very low levels of quality. In the *directional linkage* condition, we see a shift away from work periods with very low levels of quality. The Kolmogorov-Smirnov test confirms that these distributions are significantly different ($Z = 4.42$, $p < 0.01$). This shift reflects the reduction in meliorating behavior that occurs when participants receive communication about the directional causal link between quality and future financial performance. Communication of quantified causal linkages seems to create another shift toward high quality compared to the *directional linkage* condition ($Z = 1.28$, $p = 0.08$), though this shift does not involve a further reduction in the number of very low quality choices. Rather, we see more very high quality products being produced. This shift is notable, because in the *quantified linkage* condition, participants were told the exact form of the relation between current period quality and future price movements. This communication indicates a link between quality and future price movements, but also a limit to the benefits of quality. That is, there are no price-related benefits to achieving fewer average errors than one per sandwich. Thus, it is interesting that specifying this linkage actually leads to more choices of very high quality (close to zero errors per sandwich). It appears that the communication of the quantified linkage made the importance of quality particularly salient, leading to an over-investment in quality-related efforts. This suggests that it may be difficult for employees to adapt their behaviors to nonlinearities in the operating environment, even when those nonlinearities are known with certainty.

Summary of Hypothesis Tests

Overall, these results indicate that in a rich accounting task: (a) absent specific communications about causal linkages, melioration occurs (though perhaps to a lesser extent than has been shown in previous studies using abstract, discrete-choice tasks); (b) directional communications about the causal linkages, such as would be provided by a strategy map, reduce melioration to very low levels; and (c) fully quantified communications about the causal linkages are no more helpful than less specific, directional communications for reducing melioration. We also observe that melioration occurs less frequently among employees who have a reasonably accurate implicit theory about the causal linkage. An unexpected observation is that in a setting in which there are diminishing (or more specifically, no) returns to investments in actions that increase long-term performance, providing quantified information about causal linkages resulted in overinvestment in such actions.

Supplemental Analyses – Experience and Learning

Recall that one of the notable features of melioration is that this behavior is reinforced by experience, so learning is not expected to improve performance. We examine this phenomenon using the pattern of choices observed across work periods.[12] As expected, we see little effect of learning among our *no communication* participants. Learning would be indicated by a "closing in" on the optimal *average errors per sandwich* of one. Based on this notion, we calculate the *distance from 1* for each period, equal to the absolute value of [*average errors per sandwich* minus 1]. Using orthogonal polynomial contrasts, we find a significantly positive linear trend in the *distance from 1* for our *no communication* participants ($F_{1,22}$ = 5.72, p= 0.03, two-tailed). That is, despite having extensive feedback and twelve periods over which to learn, our participants moved farther from the optimal allocation over time. Even more notable, in the first work period, only 6 of 26 (23%) participants in this condition make more than the optimal of one error per sandwich, but by the final work period, 13 of 26 (50%) do so. These results highlight the difficulty of learning in distributed choice settings, and they support the notion that presence of an internality results in participants "learning" a poor strategy (Herrnstein et al. 1993).

In our *directional linkage* and *quantified linkage* conditions, the pattern is different. The linear trends in the *distance from 1* are not significantly positive in these conditions, but neither are they significantly negative (*directional linkage* $F_{1,21}$ = 0.49, p = 0.49, two-tailed; *quantified linkage* $F_{1,25}$ = 0.01, p= 0.94, two-tailed). The lack of a positive trend confirms the finding that meliorating behavior is not a significant problem in these conditions and suggest that, consistent with our theory, communication of causal linkages virtually eliminates meliorating behavior.

[12] Because the sequence of sandwiches was the same for all participants, task difficulty was held constant across experimental conditions, and thus, differences in trends across work periods are unlikely to be due to differences in task difficulty.

However, the lack of a negative trend suggests that even absent meliorating behavior, learning is difficult in this distributed choice task.

Supplemental Analysis – Total Pay

We measure overall effects of effort allocation choices on participant and firm payoffs by examining participant pay, a function of sandwich shop revenue. Our communication manipulation significantly affects *total pay* ($F_{2,74}$ = 4.58, p = 0.01; see Table 1, Panel C). Mean (standard deviation of) *total pay* is $23.63 ($6.27), $28.09 ($8.81), and $29.90 ($7.80) in the *no communication*, *directional linkage,* and *quantified linkage* conditions, respectively (see Table 1, Panel A). Follow-up contrasts reveal that the differences in *total pay* between the *no communication* and the *directional linkage* and *quantified linkage* conditions are significant (both $p < 0.04$). Thus, participants were able to extract sufficient information from the causal linkage communications to significantly increase overall payoffs to themselves and the firm. The difference in *total pay* between the *directional linkage* and *quantified linkage* conditions is not significant (p = 0.44), again indicating no evidence of additional benefits of communicating a quantified linkage beyond that provided by communicating a directional linkage.

For the average participant, total pay under perfectly-maximizing behavior (i.e., exactly one error per sandwich each period) would be $36.73. Participants earned 64.3%, 76.5%, and 81.4% of this benchmark in the *no communication, directional linkage,* and *quantified linkage* conditions, respectively. Notably, the gap between maximum and actual pay in the two conditions with communication is attributable only in small amounts to melioration (i.e., as shown in Table 1, Panel A, relatively few participants in these conditions exhibited meliorating behavior). Instead, pay is lower than possible in both of these conditions because participants over-allocated effort to quality (i.e., *average errors per sandwich* is less than one), thus making fewer sandwiches than they could have otherwise made. As noted previously, this problem is more prevalent in the *quantified linkage* condition.

We also examine *pay* across work periods, and find that *pay* exhibits a significantly positive linear trend ($F_{1,76}$ = 290.34, $p < 0.01$). An inspection of *pay* across work periods (Figure 1, Panel C) shows that a gap in *pay* between participants in the *no communication* condition and those in the *directional linkage* and *quantified linkage* conditions appears in Period 6 and widens as the task progresses. After Period 2 *no communication* condition participants never reach the pay levels that those in the other conditions do.[13]

V. Discussion and Conclusions

In this paper, we examine the effect of communicating causal linkages on meliorating behavior and on employee effort allocations. This examination is important because developing, validating, and quantifying causal linkages is costly. Before firms undertake these efforts, it is important to identify the likely benefits of doing so. Recent research indicates that despite the popularity of strategic performance measurement tools and the importance those tools ascribe to the communication of causal linkages, fewer than 30% of firms have developed causal models (also known as value-driver or strategy maps) and even fewer have validated and quantified the linkages in the models (Kaplan and Norton 2000; Ittner and Larcker 2003). Our study provides insight into potential benefits firms can expect to receive from developing and communicating these linkages.

A potential value of communicating causal linkages lies in their potential to help employees improve effort allocations in multi-period tasks, such as production tasks, in which choices made in one period influence outcomes in future periods. Consistent with theory from

[13] To confirm that this difference in trends is significant, we use the method of orthogonal polynomial contrasts, and find a significant interaction between the linear trend across periods and the *communication about causal linkage* (F = 9.95, $p < 0.01$).

judgment and decision making, we provide evidence that in the absence of communications about the linkages between performance measures, but in the presence of goal-aligning incentive contracts and goal-relevant feedback, a significant number of employees engage in meliorating behavior. That is, the complexities of maximizing performance over multiple periods causes employees to make period-to-period decisions as though they are ignoring the causal linkage. In our production setting, some employees over-invest in quantity (and pay today) while under-investing in quality (and future pay). Interestingly, other employees are fairly adept at using their implicit theories about causal linkages to make performance-maximizing effort allocation choices.

We further provide evidence that firms can enjoy benefits from developing causal business models and communicating those causal linkages among performance measures to employees, even when employees' incentives are aligned with the firm's and feedback is provided. In other words, communicating either directional or quantified causal linkages to employees significantly reduces melioration, improves employee effort allocations, and improves firm payoffs. Finally, we provide evidence that those benefits have limits, in that communication of a quantified linkage does not always improve employees' effort allocations over communication of a directional linkage. This suggests that simple mechanisms that communicate hypothesized, directional causal linkages (e.g., face-to-face communications, printed and electronic newsletters; see Sinickas 2006) may play an important decision-facilitating role in firms, and that the additional costs of validating and quantifying those linkages are not necessarily justified.

Our paper contributes to the psychology literature on melioration theory. Previous work on melioration used abstract, discrete-choice tasks. We extend the theory to consider melioration in a rich context in which participants are likely to bring implicit theories to bear. In particular, we expect that reasonably accurate implicit theories about the internality will reduce the incidence of melioration and will allow participants to make better use of any information about the internality that is provided. Consistent with this reasoning, we observe a significant amount of melioration in our rich context; however, melioration is apparently less pervasive than in previous studies that used abstract, discrete-choice tasks. We further find that participants with accurate implicit theories about the internality in the task are less likely to meliorate than are participants who do not have such theories.

The results suggest several avenues for future research. For example, we examine a setting in which the optimal allocation of efforts is not a corner solution. That is, there is a point at which increased effort toward quality yields no long-term benefits. Therefore, if an employee brings to a task an implicit theory about a lead-lag relationship (e.g., in our setting, that quality has implications for future financial performance), this theory might lead to an over-allocation of efforts to the leading indicator, whereas melioration theory would predict an under-allocation of efforts. We observed sub-optimal behavior of both sorts in our study. Future research could address mechanisms for preventing over-allocation of effort to the leading indicator.

Finally, we examine a setting in which the internality is relatively easy to understand. For example, the lag between quality and future financial performance is only one period, and so it seems that it would be relatively easy for participants to observe this internality. However, we see no evidence of learning in our experimental results. In fact, in our information-poor condition, behavior moves further from the optimal over time. A more complex lead-lag setting may provide a rich avenue for future research in feedback and cognitive theories of learning, and may enhance the value of communicating quantified linkages over directional linkages.

References

Antonides, G. and S. Maital. 2002. Effects of feedback and educational training on maximization in choice tasks: Experimental-game evidence. *Journal of Socio-Economics* 31(2): 155-165.

Banker, R., H. Chang, and M. Pizzini. 2004. The balanced scorecard: Judgmental effects of performance measures linked to strategy. *The Accounting Review* 79(1): 1-23.

Banker, R., H. Chang, and M. Pizzini. 2006. The balanced scorecard: The influence of strategy maps on decision-making. *Working Paper.* Syracuse University, Temple University, and the University of Texas at Dallas.

Brenner, T. and U. Witt. 2003. Melioration learning in games with constant and frequency-dependent pay-offs. *Journal of Economic Behavior & Organization* 50(4): 429-448.

Cardinaels, E. and P. Van Veen-Dirks. 2005. Judgmental effects of different balanced scorecard perspectives. *Working Paper*. Tilburg University.

Copeland, T., T. Koller, and J. Murrin. 1996. *Valuation: Measuring and Managing the Value of Companies*. New York, NY: John Wiley and Sons.

Dikolli, S. 2001. Agent employment horizons and the contracting demand for forward-looking performance measures. *Journal of Accounting Research* 39(3): 481-494.

Dilla, W. and P. Steinbart. 2005. Relative weighting of common and unique balanced scorecard measures by knowledgeable decision makers. *Behavioral Research in Accounting* 17: 43-53.

Eccles, R. 1991. The performance measurement manifesto. *Harvard Business Review* 69(1): 131-137.

Farrell, A., K. Kadous, and K. Towry. 2007. Contracting on contemporaneous vs. forward looking measures: An experimental investigation. *Working Paper*. University of Illinois at Urbana-Champaign and Emory University.

Haywood, M. and N. Stuart. 2006. Mental accounting and the balanced scorecard. *Working Paper*. Rider University and University of South Florida.

Herrnstein, R. 1990. Rational choice theory: Necessary but not sufficient. *American Psychologist* 45(3): 356-367.

Herrnstein, R. and D. Prelec. 1991. Melioration: A theory of distributed choice. *The Journal of Economic Perspectives* 5(3): 137-156.

Herrnstein, R. and D. Prelec. 1992. Melioration. In G. F. Loewenstein and J. Elster (Eds.), *Choice Over Time*, 235-263. New York, NY: Russell Sage Foundation.

Herrnstein, R., G. Loewenstein, D. Prelec, and W. Vaughan. 1993. Utility maximization and melioration: Internalities in individual choice. *Journal of Behavioral Decision Making* 6(3): 149-185.

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

Ittner, C. and D. Larcker. 2003. Coming up short on nonfinancial performance measurement. *Harvard Business Review* 81(11): 88-95.

Kachelmeier, S. and M. Granof. 1993. Depreciation and capital investment decisions: Experimental evidence in a governmental setting. *Journal of Accounting and Public Policy* 12(4): 291-323.

Kaplan, R. and D. Norton. 1996. *The Balanced Scorecard: Translating Strategy into Action*. Boston, MA: Harvard Business School Press.

Kaplan, R. and D. Norton. 2000. Having trouble with your strategy? Then map it. *Harvard Business Review* 78: 167-176.

Kaplan, R. and D. Norton. 2001. *The Strategy-Focused Organization: How Balanced Scorecard Companies Thrive in the New Business Environment*. Boston, MA: Harvard Business School Press.

Kaplan, S. and P. Wisner. 2004. The judgmental effects of management communications and BSC performance dimensions on performance evaluation. *Working Paper*. Arizona State University and Montana State University.

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

Libby, T., S. Salterio, and A. Webb. 2004. The balanced scorecard: The effects of assurance and process accountability on managerial judgment. *The Accounting Review* 79(4): 1075-1094.

Lipe, M. and S. Salterio. 2000. The balanced scorecard: Judgmental effects of common and unique performance measures. *The Accounting Review* 75(3): 283-298.

Lipe, M. and S. Salterio. 2002. A note on the judgmental effects of the balanced scorecard's information organization. *Accounting, Organizations and Society* 27(6): 531-540.

Luft, J. 2004. Discussion of “Managers’ commitment to the goals contained in a strategic performance measurement system”. *Contemporary Accounting Research* 21(4): 959-964.

Mainwaring, L. 1997. Maximisation and melioration as alternative forms of firm behavior. *Journal of Economic Behavior & Organization* 32(3): 395-411.

Petersen, M. and J. Samuels. 2006. Missing non-strategic targets: Are managers penalized for emphasizing important measures? *Working Paper.* Arizona State University.

Roberts, M., T. Albright, and A. Hibbets. 2004. Debiasing balanced scorecard evaluations. *Behavioral Research in Accounting* 16: 75-85.

Salterio, S., W. Knechel, and N. Kotchetova. 2005. Performance measurement systems and strategic analysis extensiveness: Auditor's usage of balanced scorecards and performance benchmarks. *Working Paper*. Queen's University, University of Florida, and University of Waterloo.

Sedatole, K. 2003. The effect of measurement alternatives on a nonfinancial quality measure's forward-looking properties. *The Accounting Review* 78(2): 555-580.

Simons, R. 2000. *Performance Measurement and Control Systems for Implementing Strategy*. Upper Saddle River, NJ: Prentice Hall.

Sinickas, A. 2006. Improving understanding of strategy. *Strategic Communication Management* 10(2): 12-13.

Tayler, W. 2006. The balanced scorecard as a strategy-evaluation tool: The effects of responsibility and causal-chain focus. *Working Paper*. Cornell University.

Tunney, R. and D. Shanks. 2002. A re-examination of melioration and rational choice. *Journal of Behavioral Decision Making* 15(4): 291.

Vera-Muñoz, S., M. Shackell, and M. Buehner. 2006 (Forthcoming). Accountants’ usage of causal business models in the presence of benchmark data: A note. *Contemporary Accounting Research.*

Young, S. and S. O'Byrne. 2001. *EVA and Value-Based Management*. New York, NY: McGraw-Hill.

Exhibit 1
Manipulation of *Communication about Causal Linkage* from the Experimental Materials

---

Now that you have experience with the sandwich-making job and the computerized ordering system, you will work as a sandwich maker for several work periods. You have signed a contractual agreement to work with <u>the same sandwich shop</u> for **all of these work periods.**

Your pay each period is based on the revenue the sandwich shop earns. Specifically, your pay will equal 5% of the total revenue you generate for the sandwich shop in each work period. **You will receive this pay in real cash at the end of today's session!**

The sandwich shop's revenue for a given period is computed as the number of saleable sandwiches you make times the selling price for that period. The selling price at this sandwich shop is $5.00 in the first work period. *The appropriate text for each experimental condition, as follows, appears here:*

No communication: However, the price may rise or fall in subsequent periods.

Directional linkage: However, the price may rise or fall in subsequent periods, depending on the quality of sandwiches produced in prior periods.

Quantified linkage: In each subsequent period, the price at this shop depends on the average number of mistakes for all sandwiches produced in the immediately-preceding period, because this affects customer demand for the sandwich shop's sandwiches. Specifically, if in a given work period there is an average of:

- one (1) or fewer mistakes per sandwich, the sandwich price in the next period will be 10% higher;
- more than one (1) but less than two (2) mistakes per sandwich, the sandwich price in the next period will remain the same;
- two (2) or more mistakes per sandwich, the sandwich price in the next period will be 10% lower.

Before the start of each work period, you will be shown the new per-sandwich price for the sandwich shop.

Finally, sandwiches with three (3) mistakes or more when compared to a customer order are thrown away and do not produce any revenue. The mistakes in sandwiches thrown away are not included in the computation of average number of mistakes. However, as long as a sandwich has fewer than three (3) mistakes, it will be sold at full price.

Remember that you will receive 5% of the revenue for all sandwiches except those that are thrown away because they have three (3) or more mistakes.

Exhibit 2
Feedback Screen for All Experimental Conditions [1]

Congratulations! You have completed round ___.

You earned $___ this period.

You completed a total of ___ sandwiches with an average of ___ mistakes per sandwich.

- You made ___ sandwiches with 0 mistakes.
- You made ___ sandwiches with 1 mistake.
- You made ___ sandwiches with 2 mistakes.
- You had to throw away ___ sandwiches, because they had 3 or more mistakes.

The sandwich shop's revenue for the period was $___.

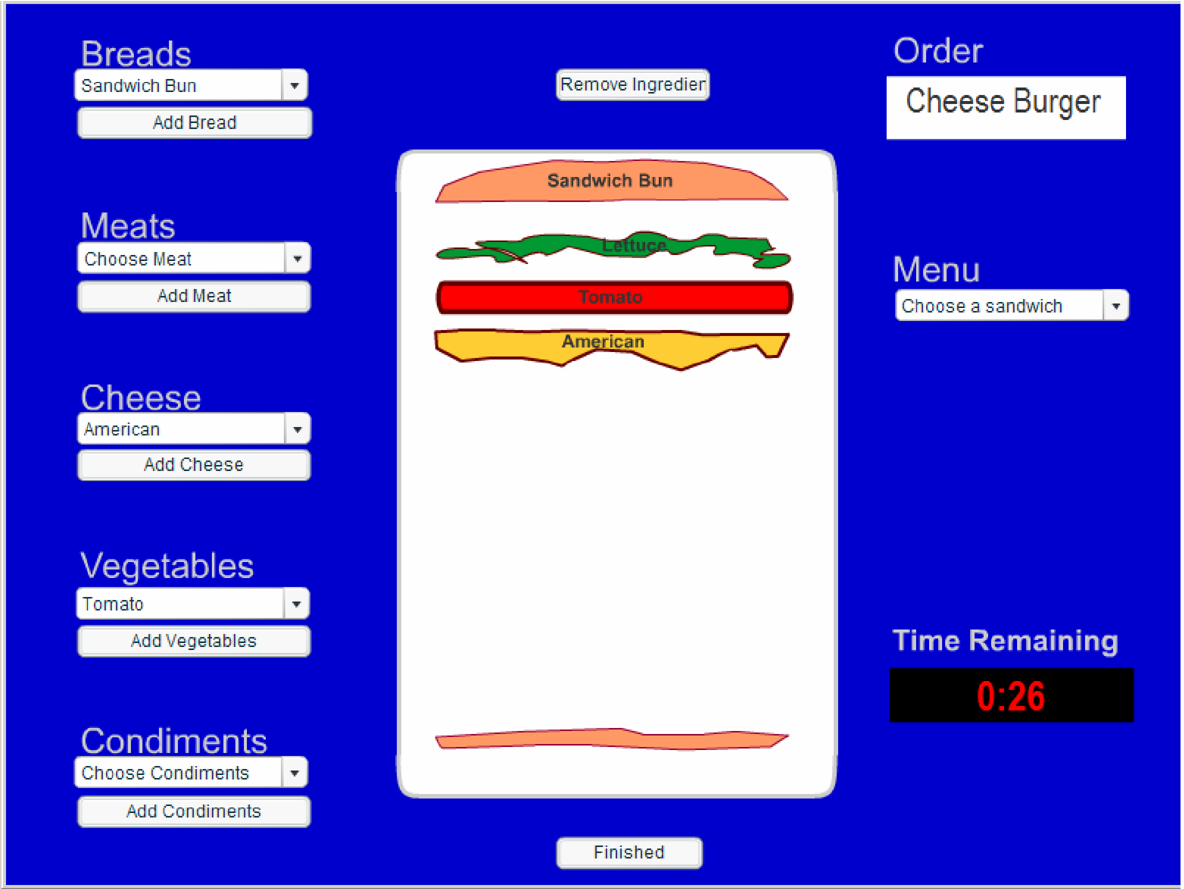
The price next period will be $___ per sandwich.

Click below when you are ready to begin the next period.

NOTE

[1] Participants viewed a feedback screen like this one at the end of each of the 12 work periods.

Exhibit 3
Sample Computer Screen from Experimental Task [1]



NOTE:

[1] Each participant, acting as sandwich maker, completed the experimental task on screens like the one shown. A customer's sandwich order (selected from a population of 51 sandwiches) appeared in the "Order" box at the top right of the screen. The participant used the "Menu" drop-down box to find the sandwich from among all those served at the sandwich shop, and then reviewed the sandwich's ingredient list (which disappeared when the computer's mouse was moved away from the "Menu" box). As the participant assembled the sandwich using the drop-down ingredient menus on the left side of the screen, images of the ingredients appeared in the production space in the center; ingredients could be removed using the "Remove ingredient" button at the top center of the screen. When the sandwich was complete, the participant clicked on the "Finished" button at the bottom center of the screen; the sandwich was automatically checked by the computer program and added to the participant's production tally, the production space was cleared, and the next customer order appeared in the "Order" box. If the last assembled sandwich had three or more mistakes, the computer program did not advance to a new sandwich in the pre-programmed sequence, but instead the participant had to attempt to make the same sandwich again until the customer order was satisfied (i.e., until that sandwich had fewer than three mistakes).

Exhibit 4
Quiz and Answers for All Experimental Conditions [1]

Please answer the following questions before going on. You may review the material on your screen to help you answer the questions. You must get these answers correct before moving on.

1. If you produce a sandwich with 5 mistakes, what happens to that sandwich?
   - a. It will be sold at a reduced price.
   - b. It will be sold at full price.
   - ⇨ c. It will be thrown away and the 5 mistakes will not be counted in the calculation of average mistakes for the period.
   - d. It will be thrown away and the 5 mistakes will be counted in the calculation of average mistakes for that period.

2. If you produce a sandwich with 2 mistakes, what happens to that sandwich?
   - a. It will be sold at a reduced price.
   - ⇨ b. It will be sold at full price.
   - c. It will be thrown away and the 2 mistakes will not be counted in the calculation of average mistakes for the period.
   - d. It will be thrown away and the 2 mistakes will be counted in the calculation of average mistakes for that period.

3. If the selling price of sandwiches is $5.00 in a period, and you make 3 saleable sandwiches, how much revenue does the sandwich shop earn in that period?
   - a. $0.75
   - b. $1.75
   - ⇨ c. $15.00
   - d. $16.50

4. If the selling price of sandwiches is $5.00 in a period, and you make 7 saleable sandwiches, how much do you earn in that period?
   - a. $0.50
   - b. $1.58
   - ⇨ c. $1.75
   - d. $35.00

NOTE

[1] Participants had to answer all quiz questions correctly before they could begin the sandwich-making task, and had to read reinforcement explanations of correct answers even if they had answered questions correctly. Arrows indicate the correct answers.

Table 1
Descriptive Statistics and Results of Hypothesis Tests

PANEL A – Means (Standard Deviations of) Dependent Variables by Experimental Condition

Dependent Variable	Communication about Causal Linkage [1]			Total
	No Communication	Directional Linkage	Quantified Linkage	
N	26	25	26	77
Number of meliorators [2]	10	2	4	16
Average errors per sandwich [3]	0.94 (0.64)	0.50 (0.42)	0.42 (0.47)	0.62 (0.56)
Total pay [4]	23.63 (6.27)	28.09 (8.81)	29.90 (7.80)	27.19 (8.04)

PANEL B – Errors per Sandwich by Condition for All Work Periods and Participants

Number (Percent) of Work Periods in which *Errors per Sandwich* [3] was:	Communication about Causal Linkage [1]		
	No Communication	Directional Linkage	Quantified Linkage
less than 1	152 (49.7%)	236 (80.0%)	264 (84.6%)
equal to 1	17 (5.5%)	26 (8.8%)	10 (3.2%)
between 1 and 2	101 (33.0%)	23 (7.8%)	34 (10.9%)
equal to 2	36 (11.8%)	10 (3.4%)	4 (1.3%)
Total	306 [5] (100%)	295 [5] (100%)	312 (100%)

Table 1 (continued)

PANEL C – Results of ANOVAs with Communication about Causal Linkage [1] as the Independent Variable

Dependent Variable	F	p (two-tailed)
Average errors per sandwich [3]	7.50	0.00
Total pay [4]	4.58	0.01

NOTES

[1] For the three levels of the independent variable *communication about causal linkage*, participants were provided the following information:

- *No communication*: the selling price "may rise or fall in subsequent periods";
- *Directional linkage:* "the price may rise or fall in subsequent periods, depending on the quality of sandwiches produced in prior periods";
- *Quantified linkage:* full details of the pricing function (if in a given work period the average number of errors was one or fewer per sandwich, the sandwich price in the next period would be 10% higher; if the average was greater than one but less than two errors per sandwich, the sandwich price in the next period would remain the same; if the average was two or more errors per sandwich, the sandwich price in the next period would be 10% lower).

[2] *Number of meliorators* is computed as a count of the number of participants who meet the following criteria: 1) *average errors per sandwich* is higher in the final six work periods than in the first six work periods, and 2) *average errors per sandwich* is greater than the optimal of one over the last six work periods.

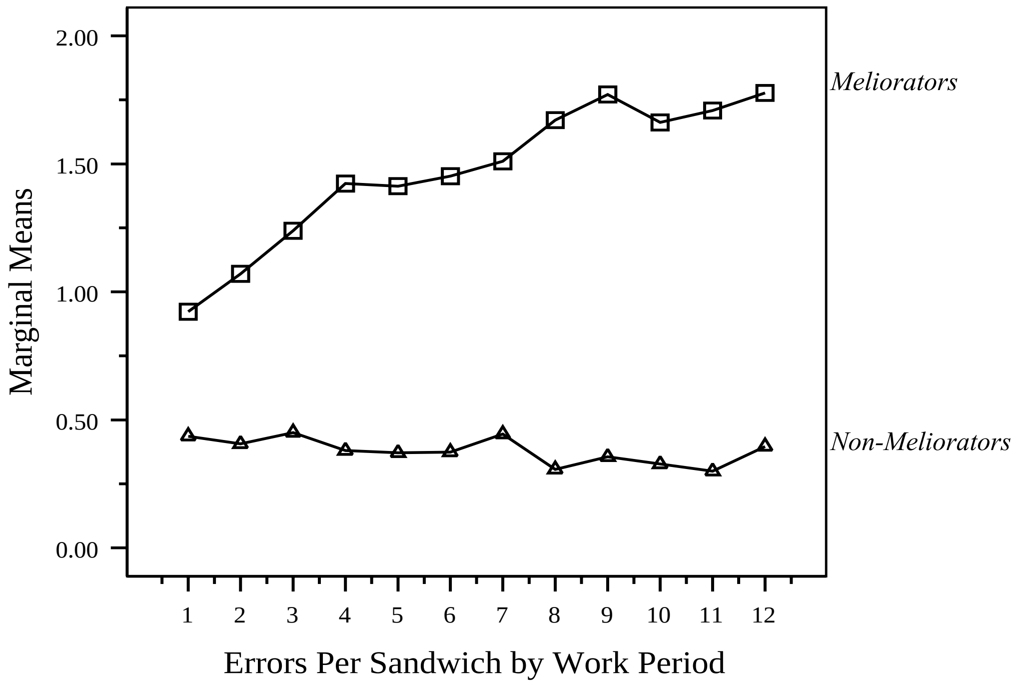
[3] *Average errors per sandwich* is computed as the simple average of the total number of errors made each period (exclusive of sandwiches thrown away) divided by the quantity of saleable sandwiches made. *Errors per sandwich* for a given work period is the total number of errors made in a given work period (exclusive of sandwiches thrown away) divided by the quantity of saleable sandwiches made that period. See Figure 2 for histograms of *errors per sandwich* across work periods.

[4] *Total pay* is the total pay earned by a participant across all work periods. Mean total sandwich shop revenue generated by participants can be computed by dividing *total pay* by 5%, the pay rate.

[5] Total does not equal the number of participants in each condition times 12 work periods because in some work periods, participants did not make any saleable sandwiches.

Figure 1
Graphs of Dependent Variables across Work Periods

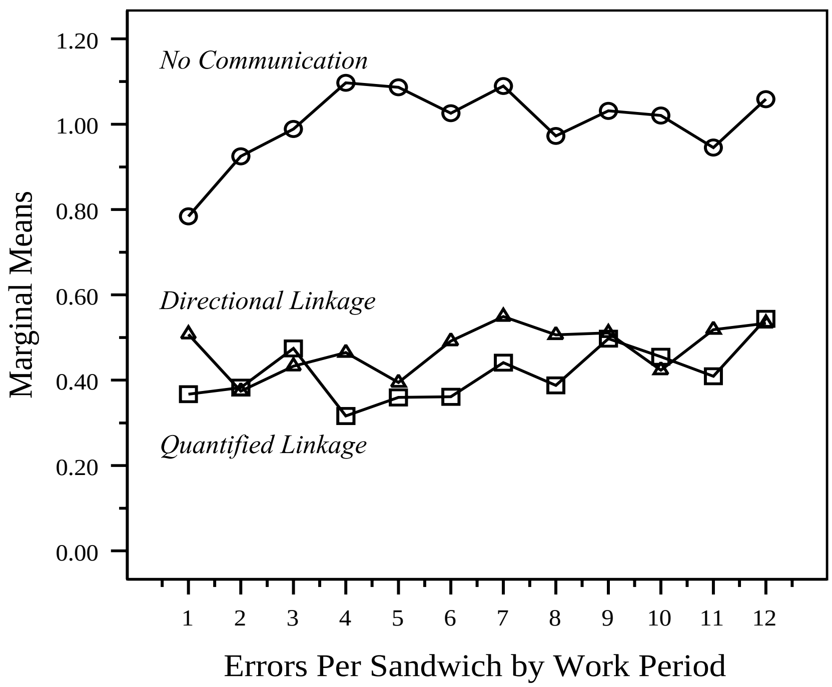
PANEL A – Errors per Sandwich [1] by Work Period for Meliorators and Non-Meliorators [2]



(continued)

Figure 1 (continued)

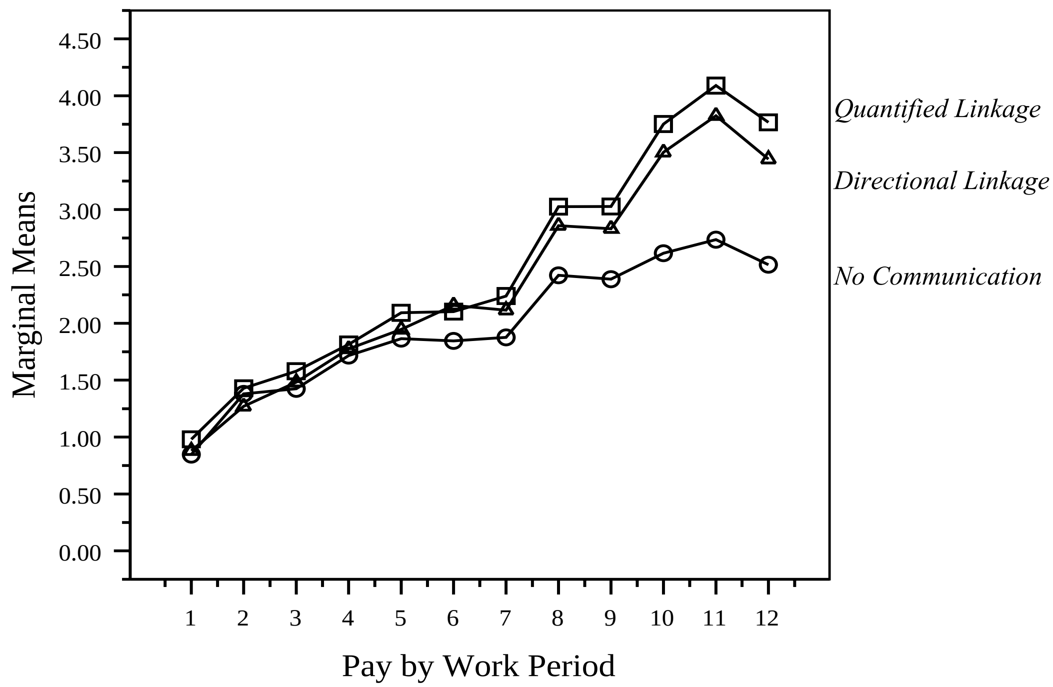
PANEL B – Errors per Sandwich [1] by Work Period and Experimental Condition [3]



(continued)

Figure 1 (continued)

PANEL C – Pay [4] by Work Period and Experimental Condition [3]



(continued)

Figure 1 (continued)

NOTES

[1] *Errors per sandwich* is computed as the total number of errors made in a work period (exclusive of sandwiches thrown away) divided by the quantity of saleable sandwiches made that period.

[2] Participants are classified as meliorators if 1) *average errors per sandwich* is higher in the final six work periods than in the first six work periods, and 2) *average errors per sandwich* is greater than the optimal of one over the last six work periods.

[3] For the three levels of the independent variable *communication about causal linkage*, participants were provided the following information:

- *No communication*: the selling price "may rise or fall in subsequent periods";
- *Directional linkage:* "the price may rise or fall in subsequent periods, depending on the quality of sandwiches produced in prior periods";
- *Quantified linkage:* full details of the pricing function (if in a given work period the average number of errors was one or fewer per sandwich, the sandwich price in the next period would be 10% higher; if the average was greater than one but less than two errors per sandwich, the sandwich price in the next period would remain the same; if the average was two or more errors per sandwich, the sandwich price in the next period would be 10% lower).

[4] *Pay* is the total pay earned by a participant in a work period. Total sandwich shop revenue generated by participants can be computed by dividing *pay* by 5%, the pay rate.

Figure 2
Frequency of Errors per Sandwich [1] From All Work Periods by Experimental Condition [2]

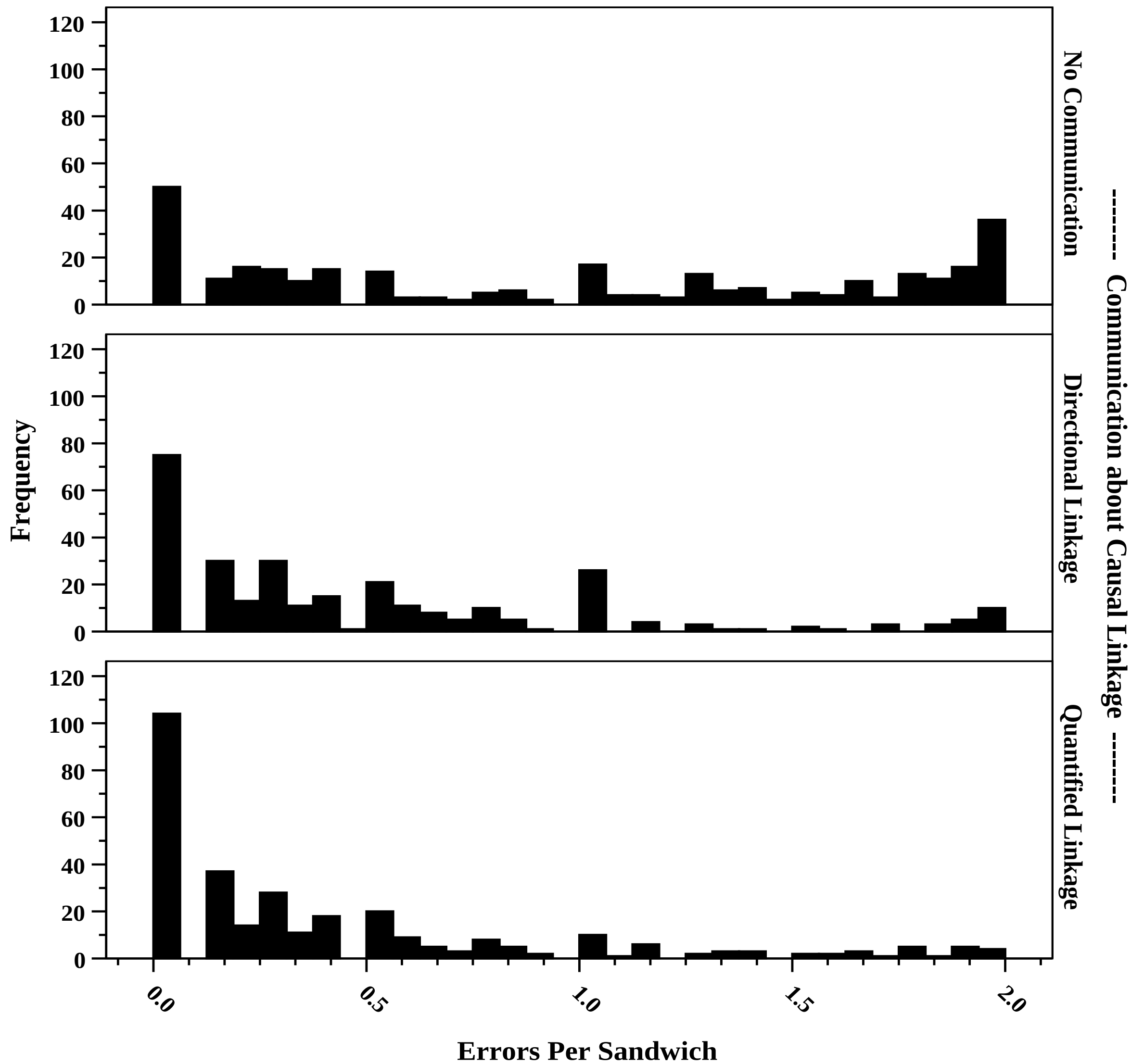


Figure 2 (continued)

NOTES

[1] *Errors per sandwich* is computed as the total number of errors made in a work period (exclusive of sandwiches thrown away) divided by the quantity of saleable sandwiches made that period. See Table 1, Panel B for counts of *errors per sandwich*.

[2] For the three levels of the independent variable *communication about causal linkage*, participants were provided the following information:

- *No communication*: the selling price "may rise or fall in subsequent periods";
- *Directional linkage:* "the price may rise or fall in subsequent periods, depending on the quality of sandwiches produced in prior periods";
- *Quantified linkage:* full details of the pricing function (if in a given work period the average number of errors was one or fewer per sandwich, the sandwich price in the next period would be 10% higher; if the average was greater than one but less than two errors per sandwich, the sandwich price in the next period would remain the same; if the average was two or more errors per sandwich, the sandwich price in the next period would be 10% lower).