BEHAVIORAL RESEARCH IN ACCOUNTING
Vol. 24, No. 2
2012
pp. 1–20

American Accounting Association
DOI: 10.2308/bria-50151

Outcome Feedback, Incentives, and Performance: Evidence from a Relatively Complex Forecasting Task

Steve Buchheit
Texas Tech University

Derek Dalton
Clemson University

Tom Downen
Southern Illinois University

Sonja Pippin
University of Nevada, Reno

ABSTRACT: This study extends prior research by examining a fairly common sequence of business events: numeric outcome information is produced and reviewed, decisions are influenced by this information, and the process repeats (i.e., a feedback loop occurs). We find that incentivized decision makers exhibit substantial decision improvement after only one iteration of summary outcome feedback. In contrast, other between-subjects groups fail to improve performance across iterations of Luft and Shields' (2001) forecasting task. Our results suggest that financial incentives and outcome feedback are both critical to performance improvement in relatively complex iterative tasks. When either incentives *or* feedback is absent, performance suffers. While prior research has found outcome feedback relatively ineffective at improving complex task performance, our results indicate that outcome feedback and incentives complement each other to improve performance. We believe exploring the interaction of incentives and feedback offers interesting avenues for future accounting research.

Keywords: outcome feedback; repetition; incentives; performance.

Data Availability: *Study data are available from the authors upon request.*

The authors thank Beau Barnes, Wendy Buchheit, Lisa Gaynor, Yong Kim, Joleen Kremlin, Theresa Libby (editor), Ken Lorek, Jeffrey McMillan, Uday Murthy, Shane Stinson, two anonymous reviewers, participants at the 2008 AAA Annual Meeting, and workshop participants at Clemson University, Southern Illinois University Carbondale, and the University of South Florida.

Published Online: March 2012

INTRODUCTION

Many business decision scenarios are repetitive and involve feedback between decision judgments (Sprinkle 2000). In such decision-making contexts, summarized outcome feedback (OFB) is a frequently offered and/or sought-after form of feedback. In accounting settings, such feedback provides individuals with quantified outcomes that generally have intuitive directional properties (e.g., more profit is intuitively good and more cost is intuitively bad in most business settings). Relative to other forms of feedback, such as explanatory feedback (i.e., step-by-step feedback regarding why a particular result occurred), OFB is less time consuming and less costly to deliver (Bonner and Walker 1994). In a world characterized by time constraints and information overload (e.g., Epstein 2007), exploring the ability of decision makers to self-learn based on numeric outcome feedback is both (1) increasingly important and (2) naturally related to accounting information.

Feedback is a critical element of accounting environments, yet our knowledge of feedback is somewhat disjointed and limited (Bonner 2008, 233). Building upon research since Kluger and DeNisi (1996), this study explores boundary conditions for effective self-learning via low-cost summary OFB reporting. Specifically, we investigate boundaries associated with (1) task complexity and (2) decision-maker incentives. Regarding task complexity, prior accounting research demonstrates OFB can effectively facilitate self-learning in a wide variety of contexts,[1] but these studies investigate relatively simple settings due to the use of perfectly predictable outcomes, a minimal number of decision cues, and/or universally diagnostic decision cues (Bonner and Walker 1994). Recent research investigating more complex tasks has focused on auditing contexts (e.g., Leung and Trotman 2005, 2008). Representative of the typical audit process, Leung and Trotman (2005, 2008) investigate *sequential* exposure to audit evidence.[2] Consistent with prior research, Leung and Trotman (2005, 2008) find that OFB can be effective for relatively simple tasks, but OFB is generally ineffective when tasks are relatively complex. In this study, we argue that periodic summary accounting reports, a common method for obtaining OFB, offer a common yet *non-sequential* setting in which OFB might improve performance in relatively complex tasks. We present potentially surprising evidence regarding the speed at which decision makers can effectively learn from summary OFB in the profit-forecasting task developed by Luft and Shields (2001).[3]

Our study examines the joint effect of summary OFB and incentives on iterative performance improvement. We are aware of no prior study in accounting (or any other discipline) that experimentally examines the interactive effects of summary outcome feedback and incentives on decision-maker performance; however, the importance of feedback and incentives in accounting has been established (Bryant et al. 2009; Coletti et al. 2005; Drake et al. 1999; Drake et al. 2007; Libby and Thorne 2009; Sprinkle 2000, 2003). Our design allows us to directly assess whether feedback and incentives can have an interactive effect on performance. In this regard, our evidence suggests that performance incentives are critical to the effectiveness of self-learning through OFB. In Luft and Shields' (2001) forecasting task, we find that when incentives are present, summary OFB improves task performance; however, when performance incentives are absent, summary OFB does not

[1] For example, Hirst et al. (1999) find that OFB improves bankruptcy predictions, Nelson (1993) finds that OFB reduces financial statement errors, Ashton (1990) finds that OFB improves bond-rating evaluations, and Harrell (1977) finds that OFB improves performance evaluations.

[2] Regarding sequential information, obtaining audit evidence is often analogous to obtaining medical information. To illustrate, physicians, like auditors, review individual cues (e.g., the patient has a fever) and then choose whether to extend the information sequence (e.g., order an additional but costly medical test) or reach a professional conclusion (e.g., the patient has pneumonia, but not a life-threatening disease).

[3] The complexity level of the Luft and Shields (2001) task is critical to this study. While Elliott et al. (2007) explicitly categorize the task as "complex," we discuss the challenge of dichotomous classifications (i.e., complex versus not complex tasks). As discussed later, the task is, at a minimum, relatively complex.

improve performance. In total, we provide support for the positivistic contention that organizations need to provide both feedback information and incentives to enhance learning (Bonner and Sprinkle 2002). Our findings are consistent with Lee's (2007) meta-analysis regarding the need for incentives in experimental studies that utilize repetition and feedback. In short, improvement in the Luft and Shields (2001) forecasting task is dependent upon both feedback and incentives.

Extending knowledge of effective learning boundaries associated with OFB is increasingly important in flattened organizational environments where fewer managers govern larger numbers of subordinates, thus making less time available for direct mentoring or step-by-step feedback. Likewise, given the importance of feedback in facilitating learning and improving performance (Sprinkle 2000, 2003), understanding factors that influence feedback effectiveness is critical in terms of understanding how decision makers process accounting outcomes. In this regard, we make two primary contributions. First, we establish that summary OFB can be an effective mechanism for self-learning in an iterative task that has complex decision properties. Second, we respond to several dimensions of Bonner and Sprinkle's (2002) call for research regarding how feedback and incentives combine to influence performance. In particular, our findings suggest that incentives can enhance the learning and performance effects associated with summary outcome feedback. In total, our findings offer new insights regarding the joint effects of feedback and incentives on learning and performance in a complex setting.

The remainder of the paper is structured as follows. The next section summarizes relevant prior literature and develops the hypotheses. The third section discusses methodology, and the fourth section presents the experimental results. The final section concludes and outlines directions for future research.

PRIOR LITERATURE AND HYPOTHESIS DEVELOPMENT

In natural business environments, decisions can involve single iteration judgment tasks, such as merger and acquisition transactions, the incorporation of a business, or the change from LIFO to FIFO accounting. Other common business decisions, however, can involve repeated judgment tasks, such as buy/sell decisions, budgeting labor and material usage, assessing absolute and relative performance, making audit risk assessments, or forecasting future financial outcomes. When business decisions are repeated, there is frequently an opportunity for feedback between decision iterations. Indeed, the provision of feedback is one of the key elements of accounting systems (Bonner 2008, 226). According to Balzer et al. (1989, 412), "feedback is the process by which an environment returns to individuals a portion of the information in their response output necessary to compare their present strategy with a representation of an ideal strategy." Per this definition, at least one decision repetition is necessary to assess the impact of feedback (i.e., both pre- and post-feedback decisions).

Types of Feedback

Psychology research has identified several different types of feedback, including outcome feedback (OFB), cognitive feedback (CFB), task properties feedback (TPF), and explanatory feedback (EFB) (Balzer et al. 1989; Hammond et al. 1973; Kluger and DeNisi 1996; Todd and Hammond 1965).[4] The general finding is that EFB (i.e., step-by-step feedback regarding *why* a

[4] The following discussion provides a brief description of alternatives to OFB. Cognitive feedback (CFB) provides individuals with information regarding their own judgment policies (Balzer et al. 1989; Leung and Trotman 2005). Judgment policies refer to the manner in which individuals use information cues (i.e., various pieces of information) to arrive at judgments. Individuals may be unaware of their judgment policies, especially in the case of more complex tasks; thus, CFB informs decision makers of these processes. Task properties feedback provides individuals with information related to the optimal weights for decision cues in a task. Essentially, TPF instructs individuals on how to optimally use decision cues to improve judgments. Explanatory feedback (EFB) includes both OFB (i.e., the correct answer) and TPF (i.e., an explanation for why a particular answer was correct).

particular answer was correct) is superior to OFB (i.e., feedback regarding the correct answers) because EFB allows for a more simplified and structured learning process, which in turn leads to greater improvements in judgment performance (Balzer et al. 1989; Bonner and Walker 1994).

Despite the advantages of EFB, Bonner and Walker (1994) note that providing EFB may not be feasible in many real-life situations. For example, audit seniors are often too busy to provide EFB to junior auditors (Earley 2001).[5] OFB, in contrast, reduces (and often eliminates) the downstream time requirement associated with one-on-one feedback. Further, OFB may be the only type of feedback available, especially when conditions involve time constraints (Bonner and Walker 1994). Since OFB is both more feasible to provide in real-life settings and naturally related to accounting information, we focus on gaining a better understanding of OFB in this study. In particular, we focus on the ability of numeric OFB to improve performance in a relatively complex forecasting task.

Outcome Feedback and Task Complexity

Financial reports that compare projections to realized outcomes are one example of numeric OFB. Simply learning actual numeric outcomes directs decision makers' attention and encourages self-reflection, often causing modified actions in future periods. Although a major element of managerial accounting is the comparison of expected and actual performance (Dickhaut 2009), prior research has generally found limited value to OFB when decision tasks are complex.

Kluger and DeNisi's (1996) meta-analysis finds that in 38 percent of evaluated studies, OFB actually leads to declines in performance. While performance degradation after receiving OFB has little intuitive appeal, Hammond et al. (1973) note that complex probabilistic tasks often contain erroneous information that can be a liability to the learner. For example, in a probabilistic task, decision makers can be misled by one-off OFB observations. Accounting researchers (e.g., Hirst et al. 1999; Libby 1981) also speculate that the ineffectiveness of OFB in prior psychology studies may exist because psychology research tends to use abstract tasks. In more natural decision environments in which participants have knowledge about the relative importance of the decision cues and the relationships between decision cues and outcomes, OFB may lead to improved judgments (Hirst et al. 1999; Libby 1981). Consistent with this assertion, prior accounting research finds that OFB can indeed improve performance, with improvement potentially limited to relatively simple tasks (Bonner and Walker 1994).

Recent feedback research in accounting has largely been in the audit domain. Notably, Leung and Trotman (2005, 2008) compare the effectiveness of various types of feedback in a relatively complex audit setting characterized by a combination of diagnostic and non-diagnostic decision cues. They find that OFB is relatively ineffective in complex configural tasks that require joint consideration of decision cues.[6] While Leung and Trotman (2005, 2008) effectively respond to calls to jointly consider task complexity and feedback (e.g., Hirst et al. 1999; Kluger and DeNisi 1996), the sequential nature of evidence processing in audit settings differs from the periodic summary OFB received by decision makers in many nonaudit settings.

---

[5] In some settings, EFB can be automated through expert systems (i.e., the system provides the decision maker with both outcome feedback and an explanation); however, the complexity and subjectivity inherent in many accounting and auditing decisions suggests that professional judgment (i.e., using accounting OFB reports to forecast) will not be replaced by expert systems anytime soon.

[6] The following is an illustration of configural versus non-configural tasks: If audit procedures A and B provide comfort for different audit risks, assessing the effectiveness of each procedure is considered a non-configural (i.e., linear) task. In contrast, if A and B address the same audit risk, assessing the effectiveness of each procedure is considered a configural task. That is, the marginal audit comfort provided by procedure B decreases if procedure A has been effectively performed.

For example, periodic summary financial reports provide OFB that frequently contains a matrix of information such as sales or cost information provided by product line across multiple locations. Algebraically, such information can be represented by linear models; however, decision makers tend to use simplifying heuristics rather than executing linear rules as the number of decision cues increases, suggesting summary reports are not viewed as "linear tasks" by decision makers.[7] The accuracy of such heuristics, relative to linear models, can be quite high depending on a decision maker's ability to match an appropriate heuristic to the decision environment (Hogarth and Karelaia 2007).

We contend that summary OFB can improve performance in relatively complicated prediction tasks involving summary financial information because summary OFB provides a signal for decision makers to improve the match between heuristic choice and the decision environment. Unlike sequential feedback, which would be unnatural in the current forecasting task,[8] summary OFB provides decision makers with a clear opportunity to see broad patterns rather than continually changing strategies (based on sequential feedback) in an effort to improve future decisions. Our contention is consistent with Bonner's (2008, 230) assertion that receiving infrequent summary OFB may be superior to receiving frequent sequential feedback in a complex task environment.

Formally, we predict:

H1: In total, decision makers who receive summary outcome feedback will improve performance more than decision makers who do not receive summary outcome feedback in a relatively complex prediction task.

Incentives, Repetition, and Task Complexity

Intuitively, incentives improve performance because "incentives increase effort and that increased effort leads to improvements in performance" (Bonner and Sprinkle 2002, 310). At a more nuanced level, performance incentives speak to the perceived complexity of an experimental task. To explain, if an individual has insufficient skill to complete a task (i.e., the task is too complex), then financial incentives will not improve performance (Bonner and Sprinkle 2002). Similarly, if a task is too simple, then incentives are not needed (i.e., more effort will not improve performance).[9]

Furthermore, for incentives to influence the effort exerted on a task, the benefits from the incentives must outweigh the cost of completing the task in a high-quality manner (Bonner and Sprinkle 2002, 306). Indeed, incentives need to be adequately large to effectively incentivize participants; otherwise, rewards will be ineffective or even counter-effective if deemed too small (Heyman and Ariely 2004; Gneezy and Rustichini 2000). The performance-contingent monetary incentives in this study ($11 to $25) are slightly larger than the average payouts in Luft and Shields (2001) and in Hodge et al. (2010). Thus, we posit that financial incentives will promote learning

[7] Hogarth and Karelaia (2007, 734) suggest that more than three cue prediction tasks result in heuristic decision making.

[8] This seems to be true for many tasks involving moderately sophisticated information users and matrix-style financial reports. For example, an analyst would view the current balance sheet as a "summary of inter-related parts" rather than a series of stand-alone sequential pieces.

[9] A large body of economics research also investigates incentive effects in experiments (see Camerer and Hogarth [1999] and Lee [2007] for reviews). General findings are consistent, albeit with slightly different terminology. Specifically, both accounting and economics researchers find that if a task is not overly simplistic, then performance-based incentives should incite cognitive effort that positively influences task performance (particularly for judgment tasks and routine tasks, as defined by Camerer and Hogarth [1999]). Improved performance is derived, in part, from a decrease in the overall variance of participant performance (Kahneman and Tversky 1979; Smith and Walker 1993; Camerer and Hogarth 1999; Wilcox 1993).

because decision makers will exert more mental effort, which in turn will increase performance in this relatively complex task.

Formally, we predict:

H2: Incentivized decision makers will outperform non-incentivized decision makers in a relatively complex prediction task.

While H1 predicts that, in total, OFB will improve performance in a relatively complex task,[10] and H2, perhaps not surprisingly, predicts that incentivized decision makers will outperform non-incentivized decision makers in a relatively complex task, Bonner and Sprinkle (2002, 329) note that "it is unclear whether feedback has additive or interactive effects with monetary incentives." To explain, if feedback and incentives have an additive effect on performance, then feedback and incentives will each improve performance, but the interaction of feedback and incentives will be insignificant. If, however, OFB and incentives positively complement each other, then incentivized participants with OFB will improve performance by more than the combination of each main effect.[11]

Although an economic meta-analysis finds that the positive effect of financial incentives becomes more prevalent when a task is faced repeatedly (Lee 2007, 630), this collective result does not directly speak to the interactive effect of feedback and incentives. In fact, Lee (2007) explicitly notes this issue requires further exploration. We respond to the call for research by examining the interaction between summary OFB and incentives.

While prior psychology research offers limited insight into the interactive effects of monetary incentives and OFB,[12] recent accounting studies suggest that incentives and OFB may complement each other. Earley (2001, 2003) finds that self-explanations can enhance the ability of OFB to improve performance. Specifically, when auditors are required to self-explain (i.e., explain to themselves the underlying rationale of the task) after receiving OFB, the effect of OFB on performance in a valuation exercise is enhanced (Earley 2001, 2003). As such, Earley (2001, 2003) shows that in a relatively difficult OFB task, decision makers can acquire more in-depth knowledge contained in OFB reports by exerting additional mental effort (via self-explanations), thus increasing performance. Similarly, performance incentives should induce decision makers to apply additional mental effort to OFB, thereby increasing the likelihood that they will develop the innovative strategies required to perform well in complex judgment tasks (Sprinkle 2000, 299). In short, to the extent that incentives increase mental effort, incentives should enhance the effectiveness of OFB.

Similarly, Sprinkle (2000) finds that incentivized participants are more likely to request and use OFB relative to non-incentivized participants, indicating that monetary incentives can encourage individuals to obtain and use OFB, thus enhancing learning and performance. Sprinkle (2000),

[10] We note that H1 does not predict that OFB will improve performance for both incentivized and non-incentivized participants; H1 simply predicts that, in aggregate, across both incentivized and non-incentivized participants, OFB will improve performance. In the results section, we assess the separate effects of OFB on performance for both incentivized and non-incentivized participants.

[11] A negative interaction is also possible; however, there is limited theoretical reason to expect such an interaction.

[12] For example, Sipowicz et al. (1962) manipulate feedback and incentives, but fail to interact the two variables. Wiener (1969) and Arkes et al. (1986) interact feedback and incentives, but find their incentive level manipulations are too weak to improve performance. More recently, Hogarth et al. (1991) find that feedback and incentives can interact; however, Hogarth et al.'s (1991) operationalization of feedback consists of penalizing participant-based errors (a concept they refer to as exactingness). Importantly, Hogarth et al. (1991) do not provide participants with correct answers (i.e., OFB). As such, their findings do not directly address the interaction between OFB and incentives.

however, does not manipulate the provision of OFB; thus, an interaction between OFB and incentives was not directly assessed. In this study, we directly assess the interaction of OFB and incentives in a relatively complex task environment.

In essence, we expect that incentives will increase the effort applied to OFB, thus enhancing the learning process and improving performance. While the mere provision of OFB can enable learning to occur, if participants are not motivated to use the OFB to learn the properties of the task, the effectiveness of OFB will likely be attenuated. However, if adequate incentives are provided, we predict that participants will exert a sufficient level of effort to use the OFB and incrementally improve performance.

Formally, we predict:

H3: Incentivized decision makers who receive summary outcome feedback will improve performance more than non-incentivized decision makers who receive summary outcome feedback in a relatively complex prediction task.

METHOD

Task

To investigate our predictions, we replicate and extend the task developed by Luft and Shields (2001; hereafter L&S)[13] for three reasons. First, the L&S task includes complex components but is also intuitive. In particular, the L&S task is based on a real-life profit-forecasting task that avoids the simplistic concerns outlined by Bonner and Walker (1994) because (1) the L&S task does not contain perfectly predictable outcomes, (2) the L&S task contains four decision cues across 20 observations, and (3) the decision cues are not universally diagnostic.[14] While prior feedback studies include some of these complex components, we note that the L&S task is relatively unique in that there is a lagged relationship between the decision cues and profit outcomes, adding an extra level of difficulty to the task investigated in this study. A second benefit of investigating the L&S forecasting task is a clear opportunity for improvement (as implied by the experimental results of L&S). Finally, we are able to extend the original L&S design by one period in order to include repetition-based OFB.[15]

In the L&S setting, research participants were informed that a multi-plant manufacturing company had recently implemented a quality improvement program. The company wanted to determine what effect, if any, discretionary quality improvement spending would have on the gross profits of each plant. Each plant produced the same product using the same technology and participants were told that quality improvement effects should be comparable across plants.

Participants were provided with learning data that included historic quality improvement spending (four consecutive quarters) and actual profit from the most recently completed quarter. Each participant viewed results from 20 company-owned manufacturing plants.[16] After studying the learning data, participants received quality improvement spending data for 20 additional (but very similar) plants within the same company. Participants were asked to predict gross profit for each plant (as shown in Appendix A). All of the experimental materials used in the learning phase

[13] We thank Joan Luft for providing the L&S experimental instrument.

[14] As such, the L&S task is a likely candidate for heuristic decision making (see Hogarth and Karelaia 2007).

[15] L&S's primary research interest was accounting fixation (i.e., differential attention to a cosmetic label change of the quarterly quality improvement expenditures) rather than repetition and feedback.

[16] We illustrate the experimental task using our second-iteration decision data (Appendix A). Data from the learning phase and first iteration (i.e., the L&S replication) are discussed fully in pages 570–573 of L&S.

and in the first task iteration (comprising the replication) are identical to the materials used by L&S.[17]

We add a second task iteration in order to investigate the effects of OFB and incentives. As described below, we use a between-participants design to manipulate financial incentives and the presence of OFB. In order to illustrate the similarity of data provided to participants, Appendix B shows the correlation matrices for each of the three data sets utilized in this study. As in the L&S data sets, actual gross profit (i.e., the realization of participants' predictions) is most closely correlated with the quality improvement spending in the earliest quarter presented (three quarters prior to the most recently completed quarter). The three-quarter lagged quality improvement spending and current gross profit have a correlation of 0.90 in the learning data, 0.95 in the first task iteration, and 0.91 in the second task iteration (see Appendix B). In all three data sets, the three-quarter lagged quality improvement spending is significantly associated with realized profit ($p < 0.05$), but correlations between profit and quality improvement spending in the other quarters (as well as the correlations among the separate quality improvement spending amounts) are not statistically different from zero. Therefore, participants who learn the lagged relationship between quality improvement expenditures and current period profits should outperform participants who do not learn the lagged relationship.

Experimental Design and Procedure

We investigate the effects of OFB and incentives in a $2 \times 2 \times 2$ repeated-measures (iterative) design. The first two factors, OFB and financial incentives, are manipulated between participants (i.e., present or absent), and the third factor, iteration, is a within-participants variable referring to the two-iteration design used in this study. Thus, our experiment includes four distinct treatment groups, namely: (1) a no OFB with no incentives group (*NOFB/NI*); (2) an OFB with no incentives group (*OFB/NI*); (3) a no OFB with incentives group (*NOFB/I*); and (4) an OFB with incentives group (*OFB/I*).

One hundred sixty-five volunteer upper-division accounting and Master of Accounting students completed the pen-and-paper task, which was administered in four separate sessions late in the academic semester at the conclusion of a regularly scheduled class meeting.[18] In each session, the instructor clarified that participation was voluntary and would take no more than 45 minutes. A teaching assistant administered the task using pre-printed participant-specific forms, which were used to randomize the OFB manipulation and increase administrative efficiency.

In the discussion that follows, we outline the chronological order of task administration. Before participants received any of the experimental materials, we administered the between-participants financial incentives manipulation. Financial incentives (absent versus present) were varied by session. When financial incentives were present, the L&S compensation method was used. Specifically, participants were told the following:

> Prediction accuracy will be measured by subtracting actual gross profits from your prediction for each plant, squaring this difference, and summing the resulting numbers across the 40 plants. Squaring the prediction error means that *big* errors contribute proportionally more than small errors do to the total performance measure, so try not to make big errors.

[17] L&S investigate fixation by including a between-participants labeling manipulation (the quality improvement expenditures were described as either *expense* or *investment*). We also include this manipulation in a portion of our study, but fail to detect significant fixation effects.
[18] Twenty-two students excused themselves from participation (17 from the no incentives condition).

FIGURE 1
Sequence of Experimental Task

Instructions and Financial Incentives	→	**Background Information and Learning Data**	→	**First Iteration Profit Predictions**[a]	→	**Provision of Outcome Feedback**	→	**Second Iteration Profit Predictions**[a]	→	**Final Experimental Questions**
Participants are given task instructions and financial incentives are manipulated between participants (i.e. present or absent).		Participants are provided with background information and learning data regarding the task.		Participants complete twenty profit predictions for the first iteration and then turn in their experimental materials.		The provision of outcome feedback is manipulated between participants (i.e., present or absent).		Participants complete twenty profit predictions for the second iteration.		Participants answer several post-experimental questions.

[a] The accuracy of the profit predictions serves as the dependent variable. We measure accuracy as the *mean absolute prediction error* (APE), which is, for each participant, the average of the absolute difference between the predicted profit and the realized profit for the 20 plants that were evaluated in each iteration.

> Pay for this task will range from \$11 to \$25. The smaller your total (squared) prediction error, the higher your pay will be.

Average payments were slightly higher than those used in L&S and in Hodge et al. (2010), who also investigate a prediction-accuracy task using a similar compensation method.[19] In the session without financial incentives, references to compensation were removed and participants were simply encouraged to "do their best."

Next, participants completed the experimental materials in two parts. In the first part, participants received background information about the task, responded to some questions, received learning data (described above), and then completed the first-iteration profit prediction (i.e., the L&S replication). At the conclusion of the L&S replication (i.e., the first iteration), participants turned in their profit predictions together with the first part of their experimental materials.[20]

After submitting materials from the first part to the teaching assistant, participants received an additional set of experimental materials. At this point, the OFB manipulation was administered (present versus absent). Participants receiving OFB were provided the actual profit outcomes for each plant from the first task iteration. Note that from an information economics perspective, having outcomes from the first prediction task adds minimal statistical value relative to the "learning data" (which were available to all participants); however, participants could actively modify strategies by comparing their initial predictions to realized outcomes. In contrast, participants in the no outcome feedback group (*NOFB*) did not receive OFB after the first prediction task. All participants then completed the second prediction task (shown in Appendix A). After completing the profit predictions in the second iteration, participants responded to some final questions and turned in their materials. Figure 1 outlines the sequence of the experimental task.

To help evaluate the perceived complexity of the L&S task, a subset of participants (representing both the feedback condition and the no-feedback condition) also completed a brief assessment of the complexity of the task after all other portions of the task had been completed and

[19] Commenting on the Hodge at al. (2010) study, Luft (2010, 137) notes that if additional effort reduces prediction error, this type of scheme offers a meaningful incentive. As such, Luft (2010) defends the "meaningfulness" of the incentives used by Hodge et al. (2010). By manipulating the use of financial incentives, we provide complementary evidence regarding the "meaningfulness" of incentives in a similar experimental setting.

[20] Using pre-printed forms, participants were instructed to copy their profit predictions for use on the second task iteration.

TABLE 1

Primary Analysis of Results

Panel A: Repeated-Measures ANOVA

	SS	df	MS	F-statistic	p
Between-Participants					
OFB_i	7.86	1	7.86	2.14	0.145
$INCENTIVES_i$ (H2)	73.41	1	73.41	19.99	< 0.001
$OFB_i \times INCENTIVES_i$	5.96	1	5.96	1.62	0.204
Error	591.21	161	3.67		
Within-Participants					
$ITERATION_t$	4.91	1	4.91	5.48	0.020
$ITERATION_t \times OFB_i$ (H1)	6.61	1	6.61	7.39	0.007
$ITERATION_t \times INCENTIVES_i$	4.04	1	4.04	4.52	0.035
$ITERATION_t \times OFB_i \times INCENTIVES_i$	1.61	1	1.61	1.80	0.182
Error	144.03	161	0.90		

Dependent Variable: *Average Absolute Prediction Errors* (*APE*s) [predicted profit minus realized profit].

Variable Definitions:
$OFB_i = 0$ if outcome feedback was not provided, and 1 if outcome feedback was provided, for participant *i*;
$INCENTIVES_i = 0$ if financial incentives were not provided, and 1 if financial incentives were provided, for participant *i*; and
$ITERATION_t = 0$ during the first task iteration, and 1 during the second task iteration.

Panel B: Planned Contrasts

	F-statistic	p
OFB/I improvement > *OFB/NI* improvement (H3)	5.74	0.019

Variable Definitions:
OFB/I = Outcome feedback with incentives group; and
OFB/NI = Outcome feedback with no incentives group.

submitted. The results of that assessment are discussed in more detail later, but in brief these results support the categorization of the L&S experimental task as relatively complex (i.e., difficult but not excessively complex).[21]

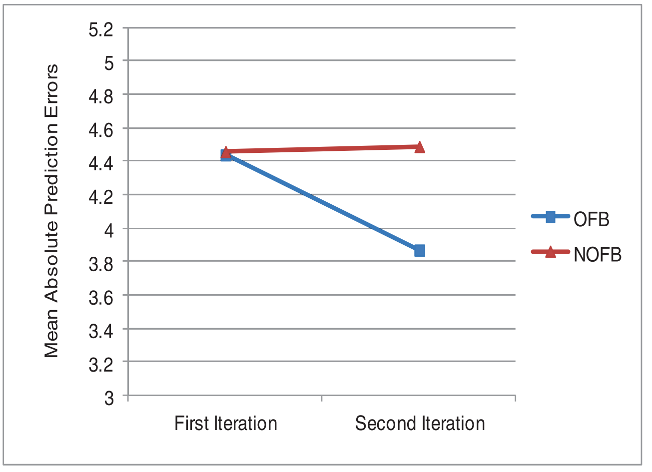
RESULTS

Table 1 presents results from the repeated-measures analysis of variance (ANOVA) used to test our hypotheses. *OFB* and *INCENTIVES* are between-participants measures, and *ITERATION* is a within-participants repeated measure (referring to the two-iteration design structure).

H1 predicts that performance will improve incrementally more when summary OFB is present. As in the L&S study, the average absolute prediction error (APE), measured as the absolute difference between forecasted and realized profit (summed across all 20 plants for each individual participant), serves as the dependent measure. The variable *ITERATION* is significant ($p = 0.02$), indicating an overall reduction of average APEs when repeating the task. Of particular importance

[21] Our participants rated the L&S task 8.04 on an 11-point scale (0 = really easy, 10 = really hard).

FIGURE 2
Interaction Effects between Task Iteration and Outcome Feedback



Variable Definitions:
OFB = outcome feedback group;
NOFB = no outcome feedback group; and
Mean Absolute Prediction Error = for each participant, the average of the absolute difference between the predicted profit and the realized profit for the 20 plants that were evaluated in each iteration, with amounts reported in millions of dollars.

to this study, the *ITERATION* × *OFB* interaction is highly significant ($p = 0.007$), indicating that participants who receive OFB are able to improve decision performance significantly more (as measured by a larger reduction in APEs) than those who do not receive OFB, providing support for H1.[22] While this result assesses the overall impact of OFB across all participants (i.e., both incentivized and non-incentivized participants), we assess the separate, differential effects of OFB on performance for incentivized and non-incentivized participants later in this section. Figure 2 graphically demonstrates the interaction of *ITERATION* and *OFB*.

H2 predicts that incentivized participants will outperform non-incentivized participants. Table 1 shows that the variable *INCENTIVES* is significant ($p < 0.001$), indicating that across both iterations, incentivized participants performed better. In untabulated t-tests, we find that incentives are positively associated with performance in both the first and second prediction iterations ($p = 0.002$ and 0.001, respectively).[23] The significant *ITERATION* × *INCENTIVES* interaction indicates

[22] Because participants could not view OFB until the first task iteration was complete, the *ITERATION* × *OFB* interaction (and not the *OFB* variable) is an appropriate test of H1. At first glance, the insignificant *OFB* main effect ($p = 0.145$) may appear surprising; however, this result is, in fact, reasonable. To explain, given that participants could not use OFB until the second task iteration, it is not surprising that first-iteration performance did not differ between participants who received versus did not receive OFB ($p = 0.931$). The main effect for *OFB* includes such first-iteration *noise* leading to an insignificant ($p = 0.145$) overall *OFB* effect.

[23] The lower first-iteration performance for incentivized participants is consistent with other accounting research (e.g., Bonner and Sprinkle 2002). For purposes of this study, we discuss performance improvement in terms of improvement from the first iteration to the second iteration. As such, the lower first-iteration performance for incentivized participants does not speak to the issue of performance improvement. It does, however, create a lower baseline for which performance is assessed over multiple iterations.

TABLE 2

Segregated Analysis of Results Based on Provision of Incentives

Panel A: Repeated-Measures ANOVA (No Incentives Condition)

	SS	df	MS	F-statistic	p
Between-Participants					
OFB_i	0.06	1	0.06	0.01	0.913
Error	360.13	73	4.93		
Within-Participants					
$ITERATION_t$	0.02	1	0.02	0.03	0.865
$ITERATION_t \times OFB$[a]	0.77	1	0.77	1.20	0.278
Error	47.13	73	0.65		

Panel B: Repeated-Measures ANOVA (Incentives Condition)

	SS	df	MS	F-statistic	p
Between-Participants					
OFB_i	15.27	1	15.27	5.81	0.018
Error	231.08	88	2.63		
Within-Participants					
$ITERATION_t$	9.91	1	9.91	9.00	0.004
$ITERATION_t \times OFB$[a]	8.18	1	8.18	7.43	0.008
Error	96.91	88	1.10		

Dependent Variable: *Mean Absolute Prediction Errors* (*APE*s) [predicted profit minus realized profit].
[a] H3 is tested by comparing the significance of the interaction term between Panel A (no incentives condition) and Panel B (incentives condition).

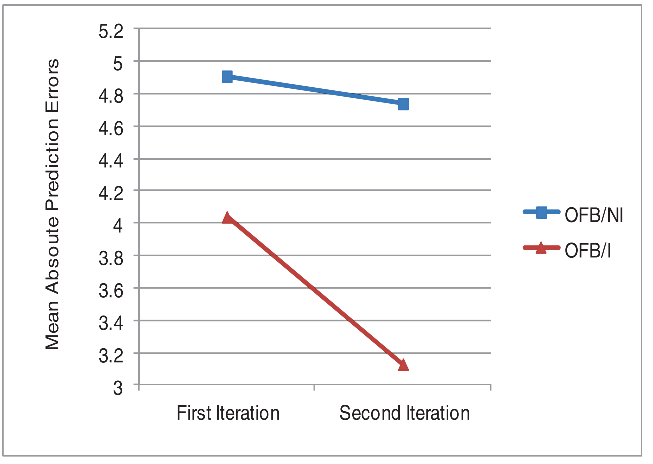
Variable Definitions:
$OFB_i = 0$ if outcome feedback was not provided, and 1 if outcome feedback was provided, for participant *i*; and $ITERATION_t = 0$ during the first task iteration, and 1 during the second task iteration.

that the effect of incentives on performance increases in the second iteration. As discussed later, this result is driven by participants who received both OFB and incentives.

H3 predicts that incentivized decision makers who receive summary OFB will improve performance more than non-incentivized decision makers who receive summary OFB. To test this hypothesis, we use a planned contrast to compare the performance improvement of the incentivized participants who received OFB to the non-incentivized participants who received OFB. As shown in Table 1, Panel B, summary OFB leads to incrementally greater performance improvement when incentives are present compared to when incentives are absent ($p = 0.019$). This finding supports H3.

To further explore H3, we separately evaluate the effect of summary OFB on both the group that received financial incentives and the group that did not receive financial incentives. Table 2, Panel A (Panel B) presents the repeated-measures ANOVA results for non-incentivized (incentivized) participants. In Panel A (non-incentivized group), the *ITERATION* $\times$ *OFB* interaction is not significant ($p = 0.278$); however, in Panel B (incentivized group), the *ITERATION* $\times$ *OFB* interaction is highly significant ($p = 0.008$). These results indicate that OFB significantly improved performance only when incentives were provided.

FIGURE 3
Interaction Effects between Task Iteration and Incentives (with Feedback Provided)



Variable Definitions:
OFB/NI = Outcome feedback with no incentives group;
OFB/I = Outcome feedback with incentives group; and
Mean Absolute Prediction Error = For each participant, the average of the absolute difference between the predicted profit and the realized profit for the 20 plants that were evaluated in each iteration, with amounts reported in millions of dollars.

Figure 3 illustrates Table 2's numeric results. The performance improvement associated with OFB largely depends on the presence of incentives. With performance incentives, improvement is dramatic; without performance incentives, improvement is attenuated to insignificance. We note that this difference occurs in spite of relatively high first-iteration APEs in the non-incentivized group (which, all else equal, should make second-iteration improvement easier). These results highlight the importance that incentives play in optimizing the use of OFB in tasks that require substantial mental effort.

To provide additional support for our primary findings, Table 3 reports absolute prediction errors (APEs) across the four treatment groups. The first-iteration APE differences *within* the feedback manipulation groups (4.80 million versus 4.91 million in the non-incentivized group and 4.20 million versus 4.04 million in the incentivized group) are not attributable to experimental manipulation because research participants had not yet received OFB when making their first-iteration forecasts. Because these randomized differences impact the interpretation of full-model results, we provide pairwise analyses below.

As shown in Panel A of Table 3, OFB without financial incentives (*OFB/NI*) is not associated with significant performance improvement ($p = 0.341$), suggesting that non-incentivized participants failed to learn from the OFB and improve second-iteration performance. Likewise, financial incentives without OFB (*NOFB/I*) is not associated with significant improvement ($p = 0.794$). This result is not surprising because incentivized participants who merely repeat a task (without receiving OFB) are not provided additional decision-relevant information (i.e., feedback) to enable performance improvement. Only the combination of summary OFB and financial incentives (*OFB/I*) leads to performance improvement in the L&S forecasting task ($p < 0.01$).

TABLE 3

Analysis of *Mean Absolute Prediction Errors* (*APE*s) by Experimental Condition

Panel A: *Mean Absolute Prediction Errors* (*APE*s) across Iterations, by Experimental Condition

Conditions	n	First Task Iteration *APE* (in Millions)	Second Task Iteration *APE* (in Millions)	Change	p
NOFB/NI	30	4.80	4.92	–0.12	0.530
OFB/NI	45	4.91	4.74	0.17	0.341
NOFB/I	38	4.20	4.15	0.05	0.794
OFB/I	52	4.04	3.13	0.91	<0.01

Variable Definition:
Mean Absolute Prediction Error (*APE*) = For each participant, the average of the absolute difference between the predicted profit and the realized profit for the 20 plants that were evaluated in each iteration, with amounts reported in millions of dollars.

Panel B: Pairwise Comparisons Using the Tukey HSD Test on Mean Change in *APE* across Iterations

	Change in *APE* (Column − Row)			
Conditions	***NOFB/NI***	***OFB/NI***	***NOFB/I***	***OFB/I***
NOFB/NI		0.29	0.17	1.03**
OFB/NI			–0.07	0.74*
NOFB/I				0.86*
OFB/I				

*, ** Significant at p = 0.05 and p = 0.01, respectively.

Variable Definitions:
NOFB/NI = No outcome feedback with no incentives group;
OFB/NI = Outcome feedback with no incentives group;
NOFB/I = No outcome feedback with incentives group; and
OFB/I = Outcome feedback with incentives group.

Table 3, Panel B presents Tukey HSD pairwise comparisons[24] that assess whether there are significant differences in performance improvement across the four treatment groups. Again, the treatment group that received both OFB and financial incentives (*OFB/I*) improved performance incrementally more than each of the other three treatment groups. Consistent with prior analysis, when OFB and incentives are both provided, performance significantly improves; however, when either is absent, performance improvement is attenuated.

Supplemental Analysis

Given that OFB helps incentivized decision makers improve performance in the L&S forecasting task (see prior section), a largely unanswered question is: "How, exactly, is this happening?" We conjecture that the additional mental effort devoted to interpreting relatively

[24] The Tukey HSD test controls for potential Type 1 errors associated with multiple comparisons.

TABLE 4

Supplemental Test of Perceived Complexity

Panel A: Repeated-Measures ANOVA for Incentivized Participants (Dependent Variable: *Mean Absolute Prediction Error*; n = 114)

	SS	df	MS	F-statistic	p
Between-Participants					
OFB_i (H3)	17.29	1	17.29	6.54	0.012
Error	298.95	113	2.65		
Within-Participants					
$ITERATION_t$	8.57	1	8.57	9.60	< 0.01
$ITERATION_t \times OFB_i$	8.07	1	8.07	9.04	< 0.01
Error	100.89	113	0.89		

Variable Definitions:
Mean Absolute Prediction Error (*APE*) = For each participant, the average of the absolute difference between the predicted profit and the realized profit for the 20 plants that were evaluated in each iteration, with amounts reported in millions of dollars;
$OFB_i = 0$ if outcome feedback was not provided, and 1 if outcome feedback was provided, for participant *i*; and
$ITERATION_t = 0$ during the first task iteration, and 1 during the second task iteration.

Panel B: Perceived Complexity

	No *OFB* Condition n = 57	*OFB* Condition n = 57	t-statistic	p
Average Perceived Complexity[a]	8.63	7.46	3.42	< 0.01

[a] Perceived complexity is measured by the following question: "How difficult was this task?" (0 = really easy, 10 = really difficult).

complex OFB makes the forecasting task less complex, thereby improving performance. To investigate this conjecture, we administered our instrument to an additional 114 upper-division accounting and Master of Accounting students.[25] All participants were incentivized; however, 57 participants received OFB (*OFB*/*I*) and 57 participants received no OFB (*NOFB*/*I*). Our expectations were: (1) we would replicate our original findings; and (2) participants with OFB would find the forecasting task less complex, relative to participants without OFB.

Regarding consistency with prior results, Panel A of Table 4 presents repeated-measures ANOVA results for these additional (and universally incentivized) participants. Consistent with our main experiment (see Table 2, Panel B), the *ITERATION* × *OFB* interaction is significant ($p < 0.01$). Also consistent with our main findings, incentivized participants who received OFB improved performance ($p < 0.01$), but incentivized participants who did not receive OFB did not improve performance ($p = 0.39$).

After completing both iterations of the forecasting task, participants were asked to respond to the following question: "How difficult was this task?" (0 = really easy and 10 = really difficult).

[25] A slightly modified instrument was used. Specifically, we removed several supplemental questions contained in the original L&S study and added questions about task difficulty for both the forecasting task and a comparison task (described in footnote 26).

The mean response of 8.04 indicates that, overall, participants perceived the task to be relatively complex.[26] Directly investigating our conjecture, Panel B of Table 4 shows that participants who received OFB perceived the task to be less complex (7.46) than participants who did not receive OFB (8.63). This significant between-groups difference ($p < 0.01$) in perceived task complexity corresponds with the OFB group outperforming the no OFB group. In addition, our results are consistent with the mental processing behavior described in Bonner and Sprinkle (2002).

SUMMARY AND DISCUSSION

We investigate the interactive effect of outcome feedback and incentives on decision performance in the profit-forecasting task developed by Luft and Shields (2001). Although prior research has found minimal support for OFB's ability to improve complex task performance, we document that rapid performance improvement can occur in a relatively complex setting.[27] In our study, performance improvement is contingent on the presence of both OFB and incentives. In our setting, we do not observe performance improvement when either OFB or performance incentives are absent, suggesting that both factors are needed to enhance learning in relatively complex environments. Consistent with Bonner's (2008) conjecture, we provide evidence that OFB can reduce perceived task complexity, which, in turn, enables performance improvement in this study.[28]

Our results are subject to a number of limitations. First, we do not test the extent to which performance improvement might be affected by other matrix-style accounting reports. Future researchers might consider the interactive effects of feedback and incentives both within and between audited financial statements. For example, investigating decision cues contained in a series of balance sheets and/or income statements (rather than the quality improvement decision cues employed by the L&S forecasting task) appears to be an interesting avenue for future research. Second, we did not manipulate task complexity. Therefore, while we find that OFB and incentives serve as complements in this relatively complex setting, a complementary relationship should not be expected in some settings. For example, if incentives are already provided, the addition of OFB likely will not improve performance in overly simplistic tasks. Likewise, in excessively complex tasks, OFB and incentives likely cannot improve performance. Future accounting research might explore the interactive effects of OFB and incentives along various ranges of task complexity.[29] Third, we investigate performance improvement over only two decision iterations. Thus, our results do not speak to the speed at which OFB might affect performance over additional task iterations.

[26] In order to provide a benchmark measure of how this subject group generally perceives experimental tasks, all 114 participants also completed the Holt and Laury (2002) risk aversion task (a common experimental economics task). In contrast to the 8.04 difficulty rating for the L&S forecasting task, the average difficulty rating for the Holt and Laury (2002) task was much lower (2.10), thereby making the L&S task significantly more complex ($p < 0.001$) in the perception of a common participant group.

[27] In untabulated results, we also test for *upper bound* performance limits by making the diagnostic decision cue (the three-quarters lagged quality improvement expenditure) more salient by boldfacing that decision cue. Participants with OFB and performance incentives were not statistically different from participants who also received a salient diagnostic decision cue, suggesting that the *OFB/I* group approached the upper limits of their performance ability after only one iteration of feedback in the L&S task.

[28] We hypothesize that incentives increase effort, which in turn reduces perceived task complexity, thus improving performance. We measure and analyze perceived task complexity and performance; however, given that prior research indicates incentives increase effort (e.g., Bonner and Sprinkle 2002), we infer (rather than directly assess) that incentives increased mental effort in our task. Future researchers might consider gathering evidence such as participants' perceived effort to provide additional insight into our findings.

[29] The L&S task is one in which simple heuristics such as one-reason decision making (Gigerenzer and Gaissmaier 2011) appear to be effective in terms of productively utilizing OFB information. Although one-reason decision making is a logical heuristic in the L&S task, Hogarth and Karelaia (2007) find that as noise increases and information redundancy increases, the effectiveness of simple prediction heuristics decreases. Future research could investigate how quickly less diagnostic information is processed by decision makers.

As a final note of caution when interpreting our results, we do not investigate whether incentives interact with other types of feedback (such as CFB, TPF, and EFB).[30] Thus, future research could investigate how alternative types of feedback are affected by performance incentives. It is reasonable to presume that explanatory feedback would limit the amount of mental effort needed to improve performance; as such, incentives may not incrementally improve performance in the presence of EFB. In an increasingly time-constrained world, rigorously establishing the conditions under which OFB can improve performance seems to be a natural area for future accounting research.

In terms of research contribution, we believe our study is the first to directly manipulate summary OFB and incentives; our findings highlight the important role that incentives can play in enhancing the effectiveness of OFB. When incentives were provided, individuals exerted sufficient effort to learn from OFB and improve subsequent performance in this study. In contrast, when incentives were absent, individuals did not exert sufficient effort to make the provision of OFB effective. These findings highlight the importance of Sprinkle's (2000) call for behavioral researchers to consider both feedback and incentives when designing an experiment. Further, to the extent OFB enables the effort to be more productive in assisting the learning process, organizations should provide adequate incentives to ensure that individuals exert sufficient effort to learn from OFB and improve performance in relatively complex tasks.

REFERENCES

Arkes, H. R., R. M. Dawes, and C. Christensen. 1986. Factors influencing the use of a decision rule in a probabilistic task. *Organizational Behavior and Human Decision Processes* 37 (1): 91–110.

Ashton, R. H. 1990. Pressure and performance in accounting decision settings: Paradoxical effects of incentives, feedback, and justification. *Journal of Accounting Research* 28 (Supplement): 148–180.

Balzer, W. K., M. E. Doherty, and R. O'Connor Jr. 1989. Effects of cognitive feedback on performance. *Psychological Bulletin* 106 (November): 410–433.

Bonner, S. E. 2008. *Judgment and Decision Making in Accounting*. Upper Saddle River, NJ: Prentice Hall.

Bonner, S. E., and G. B. Sprinkle. 2002. The effects of monetary incentives on effort and task performance: Theories, evidence, and a framework for research. *Accounting, Organizations and Society* 27 (4): 303–345.

Bonner, S. E., and P. L. Walker. 1994. The effects of instruction and experience on the acquisition of auditing knowledge. *The Accounting Review* 69 (1): 157–178.

Bryant, S., U. Murthy, and P. Wheeler. 2009. The effects of cognitive style and feedback type on performance in an internal control task. *Behavioral Research in Accounting* 21 (1): 37–58.

Camerer, C. F., and R. M. Hogarth. 1999. The effect of financial incentives in experiments: A review and capital-labor production framework. *Journal of Risk and Uncertainty* 19 (1): 7–42.

Coletti, A. L., K. L. Sedatole, and K. L. Towry. 2005. The effect of control systems on trust and cooperation in collaborative environments. *The Accounting Review* 80 (2): 477–500.

Dickhaut, J. 2009. The brain as the original accounting institution. *The Accounting Review* 84 (6): 1703–1712.

Drake, A. R., J. Wong, S. F. Haka, and S. P. Ravenscroft. 1999. Cost system and incentive structure effects on innovation, efficiency, and profitability in teams. *The Accounting Review* 74 (3): 323–346.

Drake, A. R., J. Wong, and S. B. Salter. 2007. Empowerment, motivation, and performance: Examining the impact of feedback and incentives on nonmanagement employees. *Behavioral Research in Accounting* 19 (1): 71–89.

[30] Although prior accounting research has investigated the relative effectiveness of different types of feedback (e.g., CFB, TPF) across different experimental settings (Leung and Trotman 2005, 2008; Hirst et al. 1999; Hirst and Luckett 1992), these studies did not manipulate performance-contingent monetary incentives.

Earley, C. E. 2001. Knowledge acquisition in auditing: Training novice auditors to recognize cue relationships in real estate valuation. *The Accounting Review* 76 (1): 81–97.

Earley, C. E. 2003. A note on self-explanation as a training tool for novice auditors: The effects of outcome feedback timing and level of reasoning on performance. *Behavioral Research in Accounting* 15 (1): 111–123.

Elliott, W. B., F. D. Hodge, J. J. Kennedy, and M. Pronk. 2007. Are M.B.A. students a good proxy for nonprofessional investors? *The Accounting Review* 82 (1): 139–168.

Epstein, B. J. 2007. Information overload can threaten sound decision making. *CPA Journal* 77 (3): 6–10.

Gigerenzer, G., and W. Gaissmaier. 2011. Heuristic decision making. *Annual Review of Psychology* 62: 451–482.

Gneezy, U., and A. Rustichini. 2000. Pay enough or don't pay at all. *The Quarterly Journal of Economics* 115 (3): 791–810.

Hammond, K. R., D. A. Summers, and D. H. Deane. 1973. Negative effects of outcome-feedback in multiple-cue probability learning. *Organizational Behavior and Human Performance* 9 (1): 30–34.

Harrell, A. M. 1977. The decision-making behavior of Air Force officers and the management control process. *The Accounting Review* 52 (4): 833–841.

Heyman, J., and D. Ariely. 2004. Effort for payment: A tale of two markets. *Psychological Science* 15 (11): 787–793.

Hirst, D., L. Koonce, and J. Miller. 1999. The joint effect of management's prior forecast accuracy and the form of its financial forecast on investor judgment. *Journal of Accounting Research* 37 (Supplement): 101–124.

Hirst, M. K., and P. F. Luckett. 1992. The relative effectiveness of different types of feedback in performance evaluation. *Behavioral Research in Accounting* 4 (1): 1–22.

Hodge, F., P. Hopkins, and D. Wood. 2010. The effects of financial statement information proximity and feedback on cash flow forecasts. *Contemporary Accounting Research* 27 (1): 101–133.

Hogarth, R. M., B. J. Gibbs, C. R. M. McKenzie, and M. A. Marquis. 1991. Learning from feedback: Exactingness and incentives. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 17 (4): 734–752.

Hogarth, R. M., and N. Karelaia. 2007. Heuristic and linear models of judgment: Matching rules and environments. *Psychological Review* 114 (3): 733–758.

Holt, C. A., and S. Laury. 2002. Risk aversion and incentive effects. *American Economic Review* 92 (5): 1644–1655.

Kahneman, D., and A. Tversky. 1979. Prospect theory: An analysis of decision under risk. *Econometrica* 47 (2): 263–291.

Kluger, A. N., and A. DeNisi. 1996. The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin* 119 (2): 254–285.

Lee, J. 2007. Repetition and financial incentives in economic experiments. *Journal of Economic Surveys* 21 (3): 628–681.

Leung, P. W., and K. T. Trotman. 2005. The effects of feedback type on auditor judgment performance for configural and non-configural tasks. *Accounting, Organizations and Society* 30 (6): 537–553.

Leung, P. W., and K. T. Trotman. 2008. Effect of different types of feedback on the level of auditors' configural information processing. *Accounting and Finance* 48 (2): 301–318.

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

Libby, T., and L. Thorne. 2009. The influence of incentive structure on group performance in assembly lines and teams. *Behavioral Research in Accounting* 21 (2): 57–72.

Luft, J. 2010. Discussion of: The effects of financial statement information proximity and feedback on cash flow forecasts. *Contemporary Accounting Research* 27 (1): 135–142.

Luft, J., and M. D. Shields. 2001. Why does fixation persist? Experimental evidence on the judgment performance effects of expensing intangibles. *The Accounting Review* 76 (4): 561–587.

Nelson, M. W. 1993. The effects of error frequency and accounting knowledge on error diagnosis in analytical review. *The Accounting Review* 68 (4): 804–824.

Sipowicz, R. R., J. R. Ware, and R. A. Baker. 1962. The effects of reward and knowledge of results on the performance of a simple vigilance task. *Journal of Experimental Psychology* 64 (1): 58–61.

Smith, V. L., and J. M. Walker. 1993. Monetary rewards and decision cost in experimental economics. *Economic Inquiry* 31 (2): 245–261.

Sprinkle, G. B. 2000. The effect of incentive contracts on learning and performance. *The Accounting Review* 75 (3): 299–326.

Sprinkle, G. B. 2003. Perspectives on experimental research in managerial accounting. *Accounting, Organizations and Society* 28 (2-3): 287–318.

Todd, F. J., and K. R. Hammond. 1965. Differential feedback in two multiple-cue probability learning tasks. *Behavioral Science* 10 (4): 429–435.

Wiener, E. L. 1969. Money and the monitor. *Perceptual and Motor Skills* 29 (2): 627–634.

Wilcox, N. T. 1993. Lottery choice: Incentives, complexity, and decision time. *Economic Journal* 103 (November): 1397–1417.

APPENDIX A

Second Task Iteration Data Sheet[a]

Plant	Actual Quality Improvement Expenditure: 3 Quarters Ago	Actual Quality Improvement Expenditure: 2 Quarters Ago	Actual Quality Improvement Expenditure: 1 Quarter Ago	Actual Quality Improvement Expenditure: Quarter Just Completed	Your Predicted Gross Profit for the Quarter Just Completed
101	$2.393 M	$1.614 M	$1.039 M	$1.899 M	
102	$0.855 M	$1.047 M	$1.084 M	$1.392 M	
103	$1.602 M	$0.618 M	$2.451 M	$1.116 M	
104	$1.641 M	$0.237 M	$1.558 M	$0.623 M	
105	$1.604 M	$2.537 M	$2.232 M	$1.498 M	
106	$1.266 M	$2.270 M	$0.979 M	$0.205 M	
107	$1.070 M	$1.732 M	$2.115 M	$0.757 M	
108	$0.704 M	$1.579 M	$1.675 M	$1.636 M	
109	$1.313 M	$1.966 M	$0.763 M	$0.945 M	
110	$1.342 M	$0.270 M	$1.094 M	$0.719 M	
111	$2.683 M	$0.606 M	$0.613 M	$1.548 M	
112	$1.004 M	$1.700 M	$0.723 M	$0.827 M	
113	$2.209 M	$1.641 M	$0.972 M	$0.644 M	
114	$2.054 M	$1.652 M	$0.789 M	$1.039 M	
115	$1.632 M	$2.087 M	$0.875 M	$1.325 M	
116	$0.668 M	$1.836 M	$1.463 M	$1.128 M	
117	$0.879 M	$1.702 M	$0.660 M	$0.997 M	
118	$2.164 M	$1.352 M	$1.111 M	$0.859 M	
119	$0.814 M	$0.805 M	$0.655 M	$2.206 M	
120	$2.221 M	$2.002 M	$1.446 M	$2.101 M	

[a] The terms "expense" or "investment" replaced the word "expenditure" in some sessions, but had no effect on reported results. See footnote 17 for additional discussion.

APPENDIX B

Correlations between Reported Data Elements by Iteration

		Quality Spending			
	Gross Profit$_t$	t	$t-1$	$t-2$	$t-3$
Learning Data (from Luft and Shields 2001)					
Gross Profit$_t$	1.00	−0.27	−0.02	−0.17	0.90*
Quality Spending$_t$		1.00	−0.23	−0.12	−0.18
Quality Spending$_{t-1}$			1.00	0.30	0.01
Quality Spending$_{t-2}$				1.00	−0.12
Quality Spending$_{t-3}$					1.00
First Task Iteration Data (from Luft and Shields 2001)					
Gross Profit$_t$	1.00	−0.36	0.17	0.07	0.95*
Quality Spending$_t$		1.00	−0.33	0.09	−0.28
Quality Spending$_{t-1}$			1.00	−0.03	0.13
Quality Spending$_{t-2}$				1.00	0.04
Quality Spending$_{t-3}$					1.00
Second Task Iteration Data (new to the current study)					
Gross Profit$_t$	1.00	0.05	0.13	−0.02	0.91*
Quality Spending$_t$		1.00	−0.00	0.01	0.11
Quality Spending$_{t-1}$			1.00	0.03	−0.11
Quality Spending$_{t-2}$				1.00	−0.07
Quality Spending$_{t-3}$					1.00

* Correlation differs significantly from zero ($p < 0.05$).