Analyzing the Variance Investigation Decision: The Effects of Outcomes, Mental Accounting, and Framing

Marlys Gascho Lipe
University of Colorado

SYNOPSIS AND INTRODUCTION: In evaluating the variance investigation decision made by a manager, information ex ante or ex post to the decision may have an impact. Although normative models generally consider only ex ante data, research on hindsight or outcome bias has shown that performance evaluations are often affected by decision outcomes. This outcome effect can be explained through the following cognitive links. First, the outcome of the investigation will have an impact on the perceived benefits of the investigation. Second, as posited by decision research on mental accounting, investigation expenditures matched with perceived benefits are framed as costs while those without perceived benefits are framed as losses. Third, evaluators with a cost frame provide higher performance ratings than those with a loss frame.

A series of experiments with students and members of the Institute of Management Accountants demonstrates that managers making variance investigation decisions were evaluated more favorably when investigations revealed problems in the system. This is the well-known outcome effect. Further, the investigation outcome affected the perceived benefits from the investigation and, as predicted by mental accounting, expenditures with perceived benefits were framed as costs while those without perceived benefits were framed as losses. Finally, these frames affected performance evaluation as predicted. Thus, the outcome effect on performance evaluation can now be understood within a framework that includes the cognitive impact of mental accounting and framing.

My thanks to Richard Hellevold and to participants at the accounting workshops at the University of Oklahoma, Oklahoma State University, Pennsylvania State University, and the University of Texas at Arlington for helpful suggestions on an earlier draft. Steve Selterio and Andy Terner provided research assistance and the Ann Arbor chapter of the IMA provided experimental subjects for which I am grateful.

Accepted March 1993.
Surveys report that 85 percent of large U.S. manufacturing companies use standard costing systems (Cress and Pettijohn 1985; Gaumnitz and Kollaritsch 1988). Standard costing is particularly useful for managerial control purposes since variances from standards can be calculated for management use (Horngren and Foster 1987). At that point, managers must decide whether to collect more information regarding the performance reflected in the variances. Mathematical models have been formulated to analyze this variance investigation decision (e.g., see Kaplan 1969; Jacobs and Marshall 1984). Although these mathematical models provide normative or optimal prescriptions for the variance investigation decision, managers may not follow the recommendations of the models if their decisions are judged using criteria unrelated to the models (Magee 1976). Research on human information processing and decision making has shown that judgments are affected by outcomes (or hindsight) and mental accounting (a form of decision framing). If the managerial evaluations are affected by these phenomena, managerial actions will be effected by their potential impact.

This study investigates the effects of outcomes and framing on the evaluations of managers responsible for the variance investigation decision. Specifically, it develops a cognitive mechanism for the impact of outcomes on such evaluations using prior results and ideas from judgment and decision making research. Although the impact of hindsight and outcomes has been demonstrated in prior work in accounting (Hellholeid 1988; Brown and Solomon 1987), a satisfactory cognitive explanation for the effect remains elusive (but, see Brown and Solomon 1991). Thus, the contribution of the present study is in the development and testing of this cognitive mechanism.

**Key Words:** Variance investigation, Outcome effects, Hindsight, Mental accounting, Framing.

**Data Availability:** Data gathered and used in this study will be made available by the author upon request.

The remainder of this paper is organized as follows. The next section describes very simply the context of a variance investigation decision. A discussion of outcome and framing effects follows this description. Experimental work is then presented, followed by a discussion of results.

**I. The Variance Investigation Decision**

Most graduate and undergraduate introductory managerial accounting courses teach a simple cost/benefit model for use in making variance investigation decisions. The idea is that managers compare the expected costs under two alternative actions, where the variance is investigated or not investigated, and the action with the lower expected cost is prescribed. The expected costs included in the analysis are (1) an investigation expenditure (IE), which may be the cost of managerial time and/or testing procedures; (2) the cost of correcting any problems found in the system (bringing an "out of control" system back "in control"); and (3) the incremental costs of failing to
correct an "out of control" or problematic system. Managers are assumed to estimate these costs and the associated probabilities (using past data or discussions with engineers, etc.) and then calculate the expected costs in order to arrive at a normative or optimal decision regarding investigation.

If the variance is investigated, IE is spent and the investigation will show whether the system is in or out of control. Thus, the manager gets feedback regarding his/her investigation decision if he/she investigates the variance.1 With an investigation, two outcomes are possible: the system is in control (i.e., has no problems) and IE has been spent, or the system is out of control (i.e., has a problem) and IE has been spent. In either case, the expenditure and the results of the investigation will need to be reported to the next level of management, which will evaluate the performance of the decision maker (DM).

II. The Judgmental Impact of Outcomes

Since the investigation decision is made ex ante (i.e., before results of the investigation are known), the DM should be evaluated based on the ex ante information (Edwards 1984; Fischhoff 1983a) when that information is shared by the DM and the evaluator (Hershey and Baron 1992). Managers who choose actions with lower expected costs are likely to be evaluated more favorably than those who choose actions with higher expected costs. Indeed, Lipshitz (1989) found that those taking normatively appropriate actions were evaluated more favorably than others.

However, research on hindsight or "outcome bias" (Baron and Hershey 1988) shows that ex post information also affects performance evaluation. Studies show that it is difficult to ignore intervening events or outcomes in evaluating the performance of decision makers. For example, Brown and Solomon (1987) found that business students evaluated a capital budgeting committee's decision differently after a project's success or failure was reported than they did before the outcome data were available.2 Similarly, Lipshitz (1989) showed that Israeli Defense Force officers evaluated a regiment commander more favorably when his decision to go to the aid of an attacked force was followed by a significant battle rather than a minor one. Furthermore, a commander who did not go to the attacked force's aid was viewed more favorably if this decision preceded an attack on his own sector (see also Mitchell and Kalb 1981; Baron and Hershey 1988). Each of these studies demonstrates that decision outcomes affect evaluations of the decision maker.3

Thus, in the case of a variance investigation, it is likely that evaluators will consider the outcome of the investigation (the ex post information), as well as the ex ante expected costs. Specifically, when the investigation reveals a problem, the manager's

---

1 Such feedback is certainly less clear if he/she does not investigate.
2 This effect of outcome information, though, was found to interact with the students' prior involvement in the decision task. Perhaps this involvement affected the amount of information actually shared by the DM and the evaluator in the task (Hershey and Baron 1992).
3 Although the term "hindsight bias" is intuitively appealing for referring to this outcome effect, the literature has made a distinction between the impact of outcomes on remembered probabilities and on evaluations. The former is generally called the hindsight bias and the latter the outcome effect. The relation of the two can be imagined quite easily. For example, Fischhoff and Beyth (1975) posited that ex post data cause decision makers to revise estimates of prior probabilities which could lead to changes in the assessed optimality of the original decision. These changes, then, would likely affect the evaluation of the DM.
decision will appear more appropriate and will lead to higher performance ratings than when the investigation indicates no problem with the system. This leads to the following prediction:

H1: The outcome of a variance investigation will be related to the performance rating of the DM, such that his/her performance evaluation will be higher if the system was found ex post to be out of control than if it was found to be in control.

The management and organizational behavior literature indicates that a considerable number of motivational problems arise from the impact of ex post data on performance evaluations (for an introduction to its relation to agency theory, see Kaplan and Atkinson 1989). These theorists argue that it is important that employees believe their evaluations to be fair and to be dependent on their actions (Henderson 1985; Lawler 1981). They state that when the outcomes of managers' decisions are affected by events outside of the managers' control, these outcomes are "contaminated" measures of the managers' performance.

However, there are reasonable explanations for hindsight or outcome effects on performance evaluation (Brown and Solomon 1987). Therefore, these effects should not necessarily be labeled a "bias." In some cases, for example, outcomes may be considered in evaluating a manager in order to spread some risk to that manager. Risk-sharing agreements may be negotiated and contracted quite explicitly, and such use of outcome data should certainly not be considered a bias.

Although the impact of outcome information on performance evaluation is interesting and important, prior research has already demonstrated this effect in other settings. A more interesting issue for the present study is how or why outcome knowledge affects the evaluation. The contribution of the present study is in explicating how outcome knowledge affects the mental representation or decision frame of the evaluator, so as to affect performance evaluation.

III. The Judgmental Impact of Mental Accounts and Decision Framing

Researchers have posited the use of a psychological or mental account to evaluate actions with multiple consequences (Hirst et al. in press; Tversky and Kahneman 1981; see also Thaler 1990). Kahneman and Tversky (1984) indicate that this mental account is a way of framing a decision. The mental account is used to gather and combine the information that will be used in making a decision, while data considered irrelevant to the decision are relegated to a separate account. For example, Tversky and Kahneman (1981) asked subjects to evaluate the following situation:

Imagine that you have decided to see a play where admission is $10 per ticket. As you enter the theater you discover that you have lost a $10 bill. Would you still pay $10 for a ticket for the play?

Eighty-eight percent of the subjects indicated a willingness to purchase the ticket. Another group of subjects was told "you discover that you have lost the ticket. The seat was not marked and the ticket cannot be recovered. Would you pay $10 for another ticket?" In this situation, only 46 percent of the subjects were willing to buy another ticket. Tversky and Kahneman posit that, in the first case, the lost bill was not integrated into the same mental account with the $10 ticket cost. However, in the second
case, the cost of the two tickets (the original and the replacement) were both posted to the same account so that the total cost became excessive.

Mowen and Mowen (1986) provided a more elegant research study of the effects of mental accounting (i.e., the use of mental accounts) or framing. They tested whether students' and business managers' propensity to take advantage of a constant-dollar trade discount would be affected by the size of the bill with which the discount was compared (i.e., the amount in the mental account). Consistent with prospect theory's S-shaped value function (Kahneman and Tversky 1979), both groups were more likely to pursue the discount when the order size was small.

Although the reviewed research has shown that a person's frame or mental account can affect decision making (see also Thaler 1980), there has been little accounting research on framing effects. Perhaps this is because most studies of judgment and decision making in accounting have focused on auditors for whom the decision frame may be fairly well specified (but see Ashton et al. 1988). However, Shields et al. (1987) tested whether auditors' judgments would differ when asked to provide an estimate of audit values for some accounts versus when asked to provide estimates of book value misstatements for those accounts. Although final audit values did not differ under the two "frames," there was no a priori theory predicting such a difference. This study, in contrast, takes an approach similar to Mowen and Mowen (1986), using prior decision research to make predictions about the differences in judgmental frames.

Framing Effects in Variance Investigation

Two decision frames seem to be possible in the variance investigation situation. The investigation expenditure, IE, can be framed as a cost or as a loss. Kahneman and Tversky (1984) suggest that the frame to be invoked will depend on whether there is a benefit or "advantage" to match with the expenditure (i.e., to place in the same mental account). Thus, IE will be considered a cost if some benefit is also posted to the psychological account containing the expenditure. However, if no benefit is perceived to relate to that account, IE will be considered a loss. As an oft-cited example of the impact of a benefit posted to a psychological account, Schoemaker and Kunreuther (1979) showed that presenting a choice between a sure loss and a probabilistic loss as (1) an insurance decision or as (2) a gambling decision led to significantly different choices. Although the monetary effects are the same, people displayed a greater propensity to take the sure loss (i.e., pay the premium) in the insurance context than in the gambling context. Researchers have suggested that this may be true because the term "insurance" implies that a benefit is received for the money (Schoemaker 1980, 83; Slovic et al. 1982) so that the expenditure is considered a cost when described as insurance, but considered a loss when described as a gamble.

Whether a benefit is perceived to result from a variance investigation will likely be affected by the investigation's outcome. Thus, if the system is found to be out of control (and therefore corrected), the firm would have received a benefit in the form of a corrected system. However, if an investigation reveals that the system is already per-

---

4 However, some studies have considered auditors' hypothesis-testing strategies, specifically whether a confirmatory or disconfirmatory "frame" was used (e.g., Trotman and Sng 1989).

5 Further, Thaler and Johnson (1990) and Linville and Fischer (1991) showed that people prefer to integrate (or combine) outcomes in situations where a small loss or expenditure (such as IE) is coupled with a large gain (such as the gain from bringing an out-of-control system back to in-control). Thus, the use of a mental account seems particularly plausible in evaluating a variance investigation decision.
forming correctly, an evaluator may feel that nothing was gained by the search (see section II on the effect of hindsight). This relation is posited in hypothesis H2:

H2: The outcome of an investigation will be related to the perceived benefit of the investigation, such that a greater benefit from the investigation will be perceived when the the system is found to be out of control than when the system is found to be in control.

Although the idea that a benefit implies a cost (while expenditures without benefit are losses) has been presented in many discussions of mental accounting (Tversky and Kahneman 1981; Kahneman and Tversky 1984; Thaler 1985), it has not been tested. Since predicting decision frames has proven to be very difficult in some contexts (Fischhoff 1983b), it would be particularly interesting to test this idea which is stated in hypothesis H3:

H3: Perceived benefits from an investigation will be related to the framing of IE, such that IE would be considered to be a cost to the extent that a benefit is perceived to result from it, and is considered to be a loss to the extent that no such benefit is perceived to result.

Since research has shown that a person’s frame or mental representation of a decision can affect his/her judgment (Kahneman and Tversky 1984; Mowen and Mowen 1986; Thaler 1980), a relation is expected between an evaluator’s frame and the evaluation itself. Thus, if the perception of a benefit leads to framing IE as a cost, a relatively positive evaluation of the manager who made such a payment is expected. In contrast, the manager who is held responsible for a loss (i.e., a payment without perceived benefit) is likely to be evaluated relatively lower. The following hypothesis results:

H4: An evaluator’s decision frame will be related to the performance rating of the DM such that a manager is rated higher when IE is framed as a cost than when IE is framed as a loss.

Thus, experimental work related to mental accounts and decision framing will test whether outcomes of a variance investigation affect perceived benefits, whether these perceived benefits predict the loss and cost frames, and whether the frames are related to performance evaluations of managers making the investigation decision. Figure 1 presents a schematic representation of each hypothesis’ predicted relations and effects. This figure shows that mental accounting and decision framing may provide an explanation for the outcome effect. Hypothesis H1 predicts this link of outcome and performance evaluation. The remaining hypotheses, H2, H3, and H4, then explicate the cognitive mechanisms underlying this link, indicating how and why outcomes affect performance evaluations in this context.

IV. Research Method

Three experiments will be reported: the first two use a common scenario or context and common procedures that can be briefly described as follows. Subjects were students in an undergraduate cost accounting course who had just studied variance analysis and the variance investigation expected cost model as described in section I. The students were presented with a variance investigation decision (i.e., data for IE and the other relevant costs and probabilities) and were asked to make the decision (to be sure they understood the normative model). The students were then told the manager chose
to investigate the variance and were provided with information regarding the outcome of the investigation. Next, they were asked to evaluate the manager's performance using a rating scale, and were asked to provide some judgments regarding the frame they used to evaluate the task.

The first experiment tests the effects of ex post information as posited in hypothesis H1. It also tests the predictions of hypotheses H2 and H3 regarding the decision frames invoked in the variance investigation task and hypothesis H4 regarding the effect of these frames on the performance evaluation.

Experiment One

**Subjects, Design, and Procedure:** One hundred and forty-two students in a cost accounting course participated in the experiment as a part of the class. Students received a small amount of course credit for the task. The task, administered at the end of a lecture on variance investigation, took about five minutes.

This experiment had a 2 x 2 between-subjects design, where the variables were normative decision (NORM) and investigation outcome (OUTCOME). The data presented to the subjects indicated that the normative decision was (1) to investigate or (2) not to investigate (using the simple expected costs model as our normative model). However, as this design variable was unrelated to the hypotheses of this study, it will not be discussed further.\(^6\)

Subjects were presented with the following data related to a variance investigation decision:

Assume that you are VP-Manufacturing for a medium sized firm. Your firm uses a standard costing system with currently attainable standards. You have several production managers working for you. Each one runs his/her own production plant.

\(^6\) NORM was included here to relate the experimental task to the students' course material, where they were learning to use expected costs for making the investigation decision. Certainly, since management theories cited in section II state that DMs should be evaluated on ex ante information, one would expect investigating managers to be evaluated more highly when they are acting normatively than when they are acting contrary to a normative prescription. NORM's significant main effect (shown in table 1) bears out this prediction. NORM is included in all experiments (although operationalized differently in experiment three) to increase comparability.
Pat Smith is one of those production managers. Below is information regarding a variance investigation decision that Pat faced this year.

Pat's plant had a large direct labor efficiency variance this year. Pat had to decide whether to investigate the variance further. The cost of such an investigation (IE) was estimated at $4,000. Other estimates stated that the cost of correcting a problem would be $12,000 and the present value of future costs of an uncorrected out-of-control system would be $22,000. Past data show that the probability that the system is out of control (given this large reported variance) is 50 percent.

The normative decision for the above version is to investigate the variance. Students were asked to show the normative decision to be sure that they understood the expected cost model. Any student making the wrong calculation or judgment was excluded from the analyses. 7

Students were then informed that Pat made the investigation. Some were told that the system was found to be in control, while others were told that it was out of control. Both groups were asked to make the following judgment:

You, as VP-Manufacturing, received a report from Pat detailing all of the costs and estimates involved in the investigation decision, along with Pat's decision.

You need to evaluate Pat's performance for the period. Please give a judgment or opinion on the scale below. Although Pat is not held accountable for direct material or overhead variances, the direct labor variances and investigation decisions are Pat's responsibility.

Pat's performance this period was (make a slash on the scale below):

0

| ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |

Very Poor Below Average Above Very Good

One-half of the experimental subjects were then asked to describe the frame they used to evaluate the investigation expenditure. 8 Specifically, they were presented with four frames (listed in figure 2) and asked to indicate which of the frames best described their own thinking about the task. These frames provide a comment about IE’s benefit to the company (either Benefit or No benefit) and a comment regarding the IE’s loss or cost-like nature (Loss or Cost). For example, frame BL indicates that a benefit has accrued but that IE is a loss. 9 These frame choices are used to test hypotheses H2, H3, and H4.

Results

Performance Evaluations (Hypothesis H1). The performance evaluations were analyzed using a 2 x 2 analysis of variance (ANOVA); the results are presented in panel A of...

* Only 24 out of 282 students (8.5 percent) participating in all experiments made incorrect calculations.

* The remainder of experiment one’s students performed a follow-up task similar to experiment two. Although the results of tests of hypotheses H3 and H4 for these subjects were the same as the results reported for experiment two, results for a test of hypothesis H2 were not. However, it seems quite possible that the insensitivity of the dependent variable used to measure benefit perception may have contributed to the insignificant results here. Students were asked to indicate whether they perceived a benefit by checking off a box stating “yes,” “no,” or “no opinion.” Only six of the 70 participants indicated no benefit. Therefore, although the framing and mental accounting effects were found for these subjects (hypo. H3 and H4), it was decided that a response scale could more sensitively measure benefit perception and could allow a more powerful test of hypothesis H2. This methodology and test are reported in experiment two.

* Six different orders of the four frames were constructed. A chi-squared test indicates that order did not affect the chosen frame (X^2 = 10.09, df = 15, p = 0.81). The subjects had to make a forced choice as there was no option for choosing none of the frames.
Table 1. Outcome had a statistically significant main effect on the performance ratings ($F=9.19, p<0.01$). These performance ratings, presented in panel B of Table 1, indicate that investigating managers were viewed more favorably when a problem was found in the system than when the system was in control. This is the well-known outcome effect posited in hypothesis H1.

Frame Choices and Effects (Hypotheses H2, H3, H4). As the outcome effect is present in this context, hypotheses H2, H3, and H4 posit the mechanisms by which this effect occurs. Tests of these hypotheses are more process-oriented than tests of hypothesis H1, in that subjects were asked how they thought about the task. Specifically here, they were asked to choose among the frames listed in Figure 2.

Hypothesis H2 posits that the outcome of the investigation will affect the perceived benefit from the investigation. A chi-square test of the relevant data indicates a relation of investigation outcome and perceived benefit ($\chi^2 = 15.13, df = 1, p < 0.01$). Specifically, 33 of 36 students who were told that the system was found to be out of control indicated that there was a benefit from investigating the variance (i.e., chose frame BC or BL). In contrast, only 18 of 36 students who were told that the system was found to be in control perceived such a benefit.

Hypothesis H3 implies that students’ frame choices would generally fall into BC and NL, as expenditures with benefits are costs of business (BC) while such payments without benefits are losses (NL). Of the 72 students, 55 chose either BC or NL (76 percent). A chi-square test indicates a relation between perceived benefit and loss/cost judgment ($\chi^2 = 14.82, df = 1, p < 0.01$) as predicted in hypothesis H3. Specifically, 21 students indicated that no benefit accrued from the investigation and 14 of these 21 also indicated the payment was a loss rather than a cost. In comparison, 41 of the 51 students indicating a benefit felt IE to be a cost rather than a loss.

Experiment one also provides a test of hypothesis H4. An F-test indicates that performance evaluations differed for students using the different frames ($F=7.31, df=3, 68, p<0.01$). Mean performance evaluations for each frame were 30.64 (NL), 49.29

<table>
<thead>
<tr>
<th>Frame</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>NL</td>
<td>The firm spent $4,000 on a variance investigation. The firm got nothing in return for that money. There was no benefit from spending that. It seems like that $4,000 was a loss of resources for the company.</td>
</tr>
<tr>
<td>NC</td>
<td>The firm spent $4,000 on a variance investigation. The firm got nothing in return for that money. There was no benefit from spending that. The $4,000 was just another cost of doing business for the company.</td>
</tr>
<tr>
<td>BC</td>
<td>The firm spent $4,000 on a variance investigation. The firm got something in return for that money. There was real benefit from spending that. The $4,000 was just another cost of doing business for the company.</td>
</tr>
<tr>
<td>BL</td>
<td>The firm spent $4,000 on a variance investigation. The firm got something in return for that money. There was real benefit from spending that. It seems like that $4,000 was a loss of resources for the company.</td>
</tr>
</tbody>
</table>
Table 1

Experiment One Performance Ratings

Panel A. Results of the Analysis of Variance:

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>F-stat</th>
<th>Signif.</th>
</tr>
</thead>
<tbody>
<tr>
<td>NORM (N)</td>
<td>9,304.93</td>
<td>1</td>
<td>9,304.93</td>
<td>19.80</td>
<td>0.00</td>
</tr>
<tr>
<td>OUTCOME (O)</td>
<td>4,319.45</td>
<td>1</td>
<td>4,319.45</td>
<td>9.19</td>
<td>0.00</td>
</tr>
<tr>
<td>N x O</td>
<td>67.61</td>
<td>1</td>
<td>67.61</td>
<td>0.14</td>
<td>0.71</td>
</tr>
<tr>
<td>Error</td>
<td>64,848.88</td>
<td>138</td>
<td>469.32</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>78,540.87</td>
<td>141</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Panel B. Mean Performance Ratings:

<table>
<thead>
<tr>
<th>Normative Decision</th>
<th>In-Control System</th>
<th>Out-of-Control System</th>
<th>In-Control System</th>
<th>Out-of-Control System</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>Std. Dev.</td>
<td>n</td>
<td>Mean</td>
</tr>
<tr>
<td>Investigate</td>
<td>56.75</td>
<td>22.36</td>
<td>36</td>
<td>66.41</td>
</tr>
<tr>
<td>Do Not Investigate</td>
<td>39.17</td>
<td>20.51</td>
<td>35</td>
<td>51.59</td>
</tr>
<tr>
<td>Mean</td>
<td>48.08</td>
<td>59.31</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(NC), 51.60 (BL), and 61.61 (BC). Further, the mean evaluations for students indicating a cost and a loss frame, respectively, were 59.81 and 39.37, a statistically significant difference ($t=3.66$, $p<0.01$).

Experiment Two

Results of experiment one supported all of the hypotheses. However, the subjects’ choices were quite constrained in the frame elicitation task used to test hypotheses H2, H3, and H4. Borrowing a methodology from Fischhoff (1983b), experiment one asked the subjects to choose among four frames which were briefly described. Because forced-choice tasks may provide experimental results through experimenter demand and a lack of freedom in responses, a second experiment was developed to provide an additional test of these hypotheses using a different methodology. Here, rating scales were utilized to assess the subjects’ frames, placing fewer constraints on their responses.

Subjects, Design, and Procedure. A new group of 116 cost accounting students participated in the experiment following a lecture on variance investigation. The experimental design was identical to that in experiment one; NORM and OUTCOME were manipulated between subjects. The students were asked to calculate the correct decision for a variance investigation task, were told the manager had chosen to investigate, were given the outcome of the investigation, and were asked to evaluate the manager’s performance.

After completing these tasks, the students rated the investigation expenditure along two dimensions: the degree to which it provided a benefit to the firm, and the degree to which it was a loss or a cost to the firm. The questions and rating scales were presented
in the following form (the order of these questions was varied and this order was found to have a statistically insignificant effect):

In the situation I presented to you, the firm spent $4,000 on a variance investigation. Do you feel that the firm got something in return for that money? Did the $4,000 expenditure result in some benefit to the company? Please make a slash on the scale below to indicate your opinion.

<p>| | | | | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>50</td>
<td>100</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

No, there was no benefit

Yes, there was a benefit

In the situation I presented to you, the firm spent $4,000 on a variance investigation. When you think about that $4,000, do you tend to think of it as a COST or as a LOSS to the company? I just want your opinion. How do you tend to think about it? Please make a slash on the scale below to indicate your opinion.

<p>| | | | | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>50</td>
<td>100</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

LOSS

COST

Opinion

Opinion

Results

Hypothesis H2 predicts that the outcome of the investigation will affect the perceived benefit. The mean benefit rating for the “in control” group was 71.22 while that for the “out of control” group was 84.95. These differed significantly ($t = 3.80, p < 0.01$) again providing support for hypothesis H2.

Further, hypotheses H3 and H4 are also supported. The correlation of perceived benefit and loss/cost rating, as predicted by hypothesis H3, was significant across all subjects ($r = 0.61, df = 114, p < 0.01$). Categorizing responses as “no benefit” or “loss” when below 50 and “benefit” or “cost” when above 50 allows a chi-square test of the relation between perceived benefit and frame. This test also provides support for hypothesis H3 (Yates corrected $\chi^2 = 18.31, df = 1, p < 0.01$).

Finally, the experiment supports hypothesis H4’s predicted relation between performance judgments and decision frames. The correlation of performance judgments and judgments of loss/cost frame was 0.48 in this experiment, a statistically significant relation ($p < 0.01$), and the mean evaluation for students indicating a cost frame (51 to 100 on the scale) and a loss frame (0 to 49 on the scale) were 59.94 and 34.17, respectively, a statistically significant difference ($t = 6.30, p < 0.01$).

Experiment Three

Experiments one and two yielded convergent results, showing that outcome information was related to the perceived benefits of an investigation (hyp. H2), perceived benefits were related to the loss or cost frame (hyp. H3), and this frame was related to the performance evaluation (hyp. H4). However, since student subjects were used in these experiments, generalizability to business managers may be questioned. Therefore, experiment three tests the findings of experiment two using business people as subjects.

---

10 Levene’s test for unequal variance showed a significant difference across the groups ($F = 9.69, p < 0.01$), so the reported t-test used unpoled variance.
Subjects, Design, and Procedure. The local chapter of the IMA (formerly the National Association of Accountants) was asked to assist in an academic study. Leaders of the group agreed to participate if the study could be run by mail and if the materials could be compressed to one sheet (two-sided) of paper. The stimuli material was included in a mailing of the monthly newsletter sent to 350 members. A postage-paid return envelope was also provided and the responses were returned anonymously to the researcher. Fifty-nine responses were received within one month.\textsuperscript{11}

The design was the same as that used in experiment two; the normative decision (NORM) and the outcome of the investigation (OUTCOME) were manipulated in a 2 x 2 between-subjects design. Subjects evaluated the manager's performance, and then rated IE as to whether it provided a benefit to the firm and whether it was a loss or a cost to the firm.

Operationalizing the manipulated variables was more difficult for this experiment.\textsuperscript{12} Relevant to the tested hypotheses, the OUTCOME variable was operationalized as follows:

During the investigation, Pat found that the variance had been caused by a unique and temporary event. As it is very unlikely to happen again, no correction was needed. (In-control condition.)

During the investigation, Pat found that the variance had been caused by a problem in the production system. This problem was not temporary and would continue until corrected. Pat, therefore, corrected the identified problem. (Out-of-control condition.)

Subjects were also asked their job title, the type of firm for which they worked, and how much full-time business experience they had. Finally, because it was expected that people's susceptibility to outcome effects in performance evaluation may be related to the evaluation system under which they work, the subjects were asked the following questions:

When you are evaluated at work, are you held responsible for events and circumstances that are outside of your actual control?

Do you feel it is appropriate, in your firm, to hold people responsible for events and circumstances outside of their actual control?

These will be referred to as the responsibility and appropriateness questions, respectively. Each question was followed by a 0 to 100 scale, as shown below:

\begin{align*}
0 & \mid \cdots \mid \cdots \mid \cdots \mid \cdots \mid \cdots \mid \cdots \mid \cdots \mid \cdots \mid \cdots \mid \cdots \mid \cdots \mid \cdots \mid \cdots \mid \cdots \mid \cdots \mid \cdots \mid \cdots \mid \cdots \mid \cdots \mid \cdots \mid \cdots \mid 100 \\
\text{never} & \mid \text{occasionally} & \text{often} & \text{all the time}
\end{align*}

Results

Debriefing Questions. The IMA members had an average of 15.18 years of full-time business experience, with a range of 1 to 40 years. Job titles were varied but the most

\textsuperscript{11} Certainly, relative to the total chapter membership, this response rate (17 percent) is very low. However, the number of responses is quite close to the number of active chapter members. Average attendance at chapter technical meetings is approximately 50. The author made appeals for responses at two such meetings.

\textsuperscript{12} Because few business people could be expected to recall or use the same normative model, a simplified normative model was used. Subjects were told that the firm's policy stated that "variances greater than 7 percent of standard costs allowed should be investigated by the manager (while those below this should not)." The variance was then described as 8 percent or 6 percent of standard costs allowed, indicating that the investigation should or should not take place.
Table 2
Experiment Three Performance Ratings

Panel A. Results of the Analysis of Variance:

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>F-stat</th>
<th>Signif.</th>
</tr>
</thead>
<tbody>
<tr>
<td>NORM (N)</td>
<td>1,134.64</td>
<td>1</td>
<td>1,134.64</td>
<td>4.88</td>
<td>0.03</td>
</tr>
<tr>
<td>OUTCOME (O)</td>
<td>11,164.55</td>
<td>1</td>
<td>11,164.55</td>
<td>48.01</td>
<td>0.00</td>
</tr>
<tr>
<td>N x O</td>
<td>1,145.31</td>
<td>1</td>
<td>1,145.31</td>
<td>4.93</td>
<td>0.03</td>
</tr>
<tr>
<td>Error</td>
<td>12,790.12</td>
<td>55</td>
<td>232.55</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>26,234.62</td>
<td>58</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Panel B. Mean Performance Ratings:

<table>
<thead>
<tr>
<th>Normative Decision</th>
<th>In-Control System</th>
<th>Out-of-Control System</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>Std. Dev.</td>
</tr>
<tr>
<td>Investigate</td>
<td>54.89</td>
<td>17.96</td>
</tr>
<tr>
<td>Do Not Investigate</td>
<td>37.08</td>
<td>18.25</td>
</tr>
<tr>
<td>Mean</td>
<td>47.42</td>
<td></td>
</tr>
</tbody>
</table>

frequent were controller or assistant controller (13), financial manager or VP-finance (8), and auditor (7). Another 11 subjects were cost accountants, general accountants, or accounting supervisors. Thirty of the 59 subjects were employed in manufacturing firms.

The average responsibility and appropriateness ratings did not differ across the experimental conditions. The average responsibility rating was 44.37, indicating that these professionals were sometimes held responsible for events outside of their control. However, the subjects felt that this was only "occasionally" appropriate in their firm (average appropriateness rating of 25.66). Since these people apparently believe that outcome effects are often inappropriate, it will be interesting to test whether their evaluations are affected by such outcomes.

Analysis of Outcome Data on Performance Ratings. The performance ratings were analyzed via a 2 x 2 ANOVA; the results are reported in panel A of table 2. As predicted by hypothesis H1, OUTCOME had a significant main effect ($F = 48.01, p < 0.01$). However, a significant interaction for OUTCOME and NORM ($F = 4.93, p = 0.03$) indicates that simple main effects must be examined. Panel B of table 2 presents the mean ratings for each cell of the experiment, indicating that the impact of OUTCOME was significant whether the investigation was prescribed by company policy ($t = 2.70, p < 0.01$), or not ($t = 4.29, p < 0.01$). Thus, the outcome effect (hyp. H1) was also present for this subject group.\(^\text{13}\)

\(^\text{13}\) In contrast to student subjects, here the impact of NORM was only felt when the ex post data showed the system to be in control ($t = 2.46, p = 0.01$). Note, however, that here the "normative" model is actually a company policy or rule of thumb rather than an optimal decision model.
Analysis of Frames. Experiment three reinforces the prior findings which support hypotheses H2, H3, and H4. The mean benefit rating for the "in-control" group was 63.29 while that for the "out-of-control" group was 92.21, a significant difference ($t = 4.57, p < 0.01$). This result supports the prediction of hypothesis H2 that the outcome of the investigation will affect the perceived benefit. The correlation of perceived benefit and loss/cost rating across all subjects was significant ($r = 0.47, z = 3.56, p < 0.01$), as predicted by hypothesis H3. Again, categorizing responses relative to their position above or below 50 on the scale and then using a chi-square test demonstrates that the perception of benefit and the loss/cost frames are not independent (Yates corrected $\chi^2 = 8.16, df = 1, p < 0.01$).

In experiment three the correlation of performance judgments and assessments of the loss/cost frame was 0.37, a statistically significant relation ($z = 2.85, p < 0.01$), supporting hypothesis H4. The mean evaluation for subjects indicating a cost frame (51 to 100 on the scale) and for those indicating a loss frame (0 to 49 on the scale) were 63.61 and 41.09, respectively, a statistically significant difference ($t = 3.51, p < 0.01$).

V. Discussion and Extensions

The results of these experiments provide insights into both performance evaluations and the prediction of frames. These two areas will be discussed separately, followed by suggestions for extensions of this research.

Managerial Performance Evaluation

The experimental results indicate that a manager’s performance evaluation is affected by ex post information and the decision frame of the evaluator (see hyp. H1 and H4 in fig. 1). The importance and appropriateness of outcome effects on performance evaluation can only be evaluated within the environment of the particular decision maker and evaluator. The experimental task used in this study abstracted from the contracting and incentive environment of the firm in order to take a first look at the issue.\footnote{Levene’s test for unequal variance again showed a significant difference across the groups ($F = 31.74, p < 0.01$), so the reported $t$-test used unpoolded variance.}

Although hindsight and outcome effects are not necessarily a bias within a given environment, their existence can affect decisions. A rational manager who believes his/her evaluations will be affected by outcomes has reason to supplement the basic variance investigation decision model (or company policy) with these expected outcomes and personal payoffs or even to substitute them for the parameters of the basic model (Magee 1976; Prakash and Rappaport 1977).

It would be interesting to see some modeling of the investigation decision which includes these expected ex post data, outcome effects, and their costs and benefits. Brown and Solomon (1991) suggest that outcome effects may affect managers’ abilities to learn from feedback and may be best dealt with through the design of the evaluation system (rather than by trying to train managers to avoid the effect).

\footnote{Replies to the responsibility question in experiment three showed that firms vary considerably in how much responsibility they assign for uncontrollable factors. The range of responses was from 0 to 94, with a standard deviation of 27.09.}
Implications for Framing Research

Fischhoff (1983b, 103) stated that his framing prediction "results were generally discouraging for the prediction of individuals' choices, generally encouraging for the prediction of group choices." In the current study, frames could be predicted based on perceived benefits of an expenditure. Framing researchers may need to explicate more carefully the mechanisms by which frames are invoked. By knowing more about these mechanisms and links, framing research can become both more predictive and more precise.

The current study also finds a relation between frames and performance evaluation. This result contrasts with Fischhoff's (1983b) difficulty in relating decision frames and decision choices, even though experiment one used a methodology very similar to Fischhoff's. He presented subjects with short explanations of different frames and asked them to indicate the one that seemed most natural or the one that they had to think about the choice. However, subjects' preferences in making the actual choices were unrelated to their indicated frames. In experiment one of the current study, subjects also indicated the frame they used to think about the investigation expenditure, and their frames apparently did affect the performance evaluations they made of the manager.

Although it is difficult to determine why the two studies have different results, Fischhoff's subjects may have had less motivation to appear consistent than did the students and professionals in the current study. Students in the current study may have felt they needed to make consistent frame choices and evaluations since they were returning the materials to their instructor, and the business people may have felt that consistency was important to their perceived professionalism (or they simply may make more consistent judgments as a result of their training). In contrast, Fischhoff's subjects were recruited from a more heterogeneous population, including students from throughout the university and people at an employment agency. Future research on framing's impact on judgments and decisions should consider the effect of motivation and justification (Ashton 1990; Hagafors and Brehmer 1983; Libby and Lipe 1992).

Extensions of the Study

Interesting extensions of this study include testing the proposed cognitive mechanisms for outcome effects in other contexts or tasks where prescriptive models or policies are used. Other such tasks include order or production quantities (EOQ), capital budgeting decisions, product choices, and sampling in an audit setting. Also, comparisons with Fischhoff's work suggest that the relation of decision frames with evaluations or decisions may be affected by accountability or reputation-building. Thus, future research on frame elicitation and decision making could test these effects directly.

Another fascinating question is how the cognitive framework developed here relates to other explanations offered for hindsight or outcome effects. Although no well-developed theory has been advanced for outcome effects, some suggestions have been offered. Brown and Solomon (1991) summarized these in three categories: those using a cognitive reconstruction explanation, those with a self-enhancing motive explanation, and those with an escalation of commitment explanation. While an analysis of these explanations as they relate to the framework developed here is beyond the scope of this article, designing tests which could discriminate among them or explore
their relation to the framework would be a very important extension of this work. Also, Brown and Solomon (1987) showed that prior involvement in a decision task can reduce the outcome effect. An interesting integration of that work with the present study would test the impact of prior involvement on each of the cognitive links posited here to determine which of the links are affected by such involvement. Thus, the present paper has provided a cognitive explanation for the outcome effect and this should provide direction and impetus for future work in this area.

References


