Journal of Business & Leadership: Research, Practice, and Teaching (2005-2012)

Volume 2
Number 2 *Journal of Business & Leadership*

Article 3

1-1-2006

Effect of Outcome Knowledge On Repeated Performance Evaluations

Russell Calk
New Mexico State University

Maryanne Mowen
Oklahoma State University

Journal of Business and Leadership: Research, Practice, and Teaching
2006, Vol. 2, No. 2, 231-241

EFFECT OF OUTCOME KNOWLEDGE ON REPEATED PERFORMANCE EVALUATIONS

Russell Calk, New Mexico State University
Maryanne Mowen, Oklahoma State University

Outcome information has been shown to have a significant impact on performance evaluations in a wide variety of decision settings. Studies of outcome effects on performance evaluation typically examine a one-time decision. Often, however, managers must repeatedly evaluate performance. This study examines how outcome information affects the revision of multiple performance evaluations over time. The results have important implications for research into the sensitivity of evaluators to the sequence of the receipt of positive and negative information. Implications of these results for performance evaluation in multiple areas of business are discussed.

INTRODUCTION

Outcome effects have been shown to have a significant impact on performance evaluations (Ghosh and Lusch, 2000; Luckett and Eggleton, 1991). The result is robust across a wide variety of decision settings including financial distress (Tuttle and Stocks, 1998; Fisher and Selling, 1993), capital budgeting (Brown and Solomon, 1987, 1993; Cheng et al., 2003), investment choices (Frederickson et al., 1999), personnel decisions (Highhouse and Gallo, 1999), corporate branding strategy (Agrawal and Maheswaran, 2005), and accounting variance investigations (Lipe, 1993). The typical study of the influence of outcome effects on performance evaluations examines an isolated individual in a one-time, well defined decision setting under conditions that ignore organizational characteristics (Ashton, 1990).

Performance evaluation, however, is a dynamic process and is important to both the evaluator and the evaluatee. Frequently, managers are required to repeatedly evaluate the performance of the same person or team over a period of time. In this setting, evaluators continuously revise and update prior-period evaluations as new information becomes available, a process Hogarth and Einhorn (1992) call belief revision. Several studies of audit decision making have shown that prior period evaluations are revised in this manner (Wilks, 2002; Ashton and Ashton, 1988; Asare, 1992; and Tubbs et al., 1990, 1993). In outcome-based performance evaluation schemes, evaluators focus their attention on actual outcomes (Brown and Solomon, 1993; Frederickson, 1992; Frederickson et al., 1999).

The purpose of this study is to link the studies of outcome-based performance evaluation and of belief revision by examining how outcome information impacts belief revision over time in the process of repeated performance evaluations. We base our study on the Hogarth and Einhorn (1992) model which posits that recency effects outweigh primacy effects. That is, later outcomes have a stronger impact on evaluations than earlier outcomes. In addition, this model suggests that the impact of the more recent outcomes diminishes over time. Our study is designed to test that assertion. Our student subjects were given five evaluations to make over a period of five days. The results show that there is a significant contrast effect such that recent information that contrasts with previous information is dramatically overweighted. Furthermore, the effect does not diminish over time.

This study is the first study of outcome evaluations over multiple time periods. As such, it contributes to the important area of management accounting research on performance evaluation. It also holds implications for other streams of literature in which evaluations must be made over a series of time periods, such as the determination of a bank's loan portfolio strength, variance investigation, or advertising agency efforts. The study may also hold implications for nonbusiness fields such as the evaluation of rankings of sports teams.

The remainder of this paper is organized as follows. The next section discusses a model of systematic belief revision. Sections 3 and 4 describe the methodology and present an analysis of the results. The final section offers some concluding remarks and suggestions for future research.

Model of Systematic Belief Revision

The periodic evaluation of employees by managers involves the systematic revision of prior evaluations based upon the introduction of new evidence. With their belief adjustment model, Hogarth and Einhorn (1992), propose a structure for systematic belief revision. While other models of belief revision exist, the HE Model has been shown to "capture both the direction and magnitude of auditors' belief revision." (Krishnamoorthy, 1999: 105). We believe that the manager's performance evaluation task is sufficiently similar to the evaluation task of auditors that the HE Model provides a good theoretical basis for considering belief revision.

In the HE Model, beliefs are revised based upon the effect that the introduction of new evidence has on an initial anchor. In an outcome-based performance evaluation scheme, the prior period's evaluation serves as the anchor and subsequent actual outcomes are treated as new evidence and included into belief adjustment model.

According to the HE Model, not all evaluators will place the same weight on a new bit of information. Furthermore, the impact of new information on the initial anchor is dependent on the direction of the evidence. That is, a person may be so invested in a belief or decision (the anchor) that evidence confirming the anchor will be weighted much more heavily than disconfirming evidence. Over time, as information accumulates and evaluators become more committed to their beliefs, the impact of new evidence will diminish.

For example, suppose that the actual outcome reported by a subordinate manager is significantly lower (higher) than the expected outcome. As a result, the evaluator gives the subordinate a low (high) performance evaluation. For the next period, the actual outcome is again lower (higher) than expected. The evaluator may or may not place a significant subjective weight on this new information so that the evaluation may or may not change. Thus, after several periods, the impact of the reported results on the performance evaluation should diminish so that regardless of how actual results compare to expectations, the evaluation will remain essentially unchanged from the previous period.

The model of this phenomenon developed by Hogarth and Einhorn (1992) appears as follows:

$$S_k = S_{k-1} + \alpha S_{k-1}[s(x_k) - S_{k-1}] \quad (1)$$

and

$$S_k = S_{k-1} + \beta S_{k-1}[s(x_k) - S_{k-1}] \quad (2)$$

where:

S_k is the evaluation for period k.
α is the sensitivity towards reported results that disconfirm the previous evaluation, and $0 \leq \alpha$.
β is the sensitivity towards reported results that confirm the previous evaluation, and $\beta \leq 1$.
x_k is the reported result in period k.
$s(x_k)$ is the subjective weight that the evaluator places on the reported results in period k.

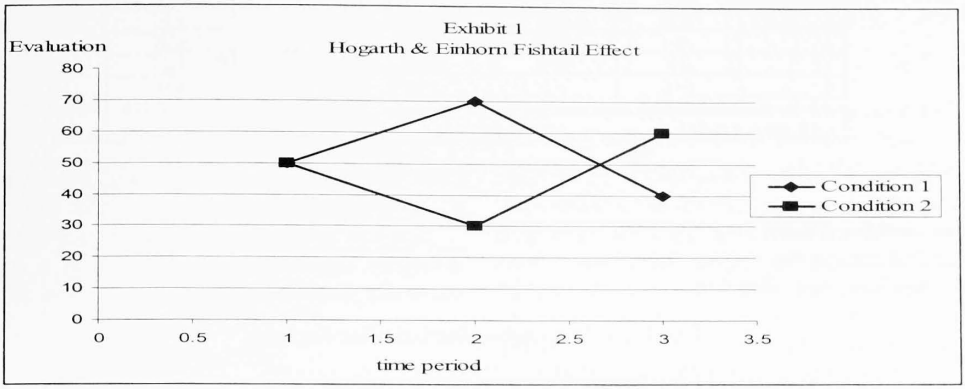
The model simply states that the evaluation in period k is equal to the prior period's evaluation plus some adjustment. The adjustment is a function of both the subjective weight placed on the current period's reported results and the evaluator's sensitivity towards this additional information. Equation (1) applies to reported results that disconfirm the previous evaluation. Equation (2) applies to reported results that confirm the previous evaluation.

The values of α and β are functions of both individual and situational variables. Some evaluators may weight confirming results higher than disconfirming outcomes, or vice versa. The same evaluator may also place different weights on reported results in different situations.

In Hogarth and Einhorn's (1992) experimental work, subjects were split into two groups - condition 1 and condition 2. In both conditions, subjects were given an initial piece of information and asked to make an evaluation. Then the subjects in condition 1 were given a second piece of information that was better than the first; the subjects in condition 2 were given a second piece of information that was worse than the first. As expected, the subjects given the better (worse) information raised (lowered) their evaluations. Finally, the subjects in condition 1 were given a third piece of information that was worse than the second piece, while the subjects in condition 2 received a third piece of information that was better than the second piece. Importantly, the average of the three pieces of information was identical in both conditions. The results revealed that the third evaluation was much lower for the first condition and much higher for the second condition. This is known as the "fishtail effect." Shown graphically in Exhibit 1, subjects in condition 1 received a disconfirming third piece of information that was negative, and that resulted in a considerably lower evaluation than that given the first piece. Subjects in condition 2, on the other hand,

https://scholars.fhsu.edu/jbl/vol2/iss2/3
DOI: 10.58809/CXRA3757

received a disconfirming third piece of information that was positive, and those subjects awarded an evaluation that was considerably higher than that given the first piece.



Hogarth and Einhorn (1992) use the results of this experiment to argue that recency (the most recent piece of information) has a stronger effect on people's evaluations than does primacy (the initial piece of information). Even so, and regardless of individual and situational variables, over a long series of evaluations, they believe that evaluators should become more committed to their beliefs and, therefore, the values of α and β should approach zero. As this happens, the performance evaluation in period k should become equivalent to the evaluation in period k-1. In other words, the HE Model posits that over time, as one receives more and more information, s/he should become more committed to his/her views, and new information should have little to no effect. This leads to the following hypothesis.

Hypothesis: The impact of outcome information on performance evaluation will diminish over time.

To study this hypothesis, we created an experiment in which subjects used reported accounting results to repeatedly evaluate the performance of the same team. In a results-based control system, accounting information is often used to assess the actual results (McNair 1994). The reported accounting results serve as outcome information in the performance evaluation scheme.

METHODOLOGY

The experiment is a mixed design with four between subject treatments and repeated measures of evaluations across five time periods. The variable manipulated is the Ss' rating of the groups' performance.

There were four between subjects treatments. Each subject was presented with a folder containing the scenario in appendix A, the manipulation check in appendix B, and the first of five evaluation forms. Each evaluation form was similar to the one presented in appendix C. The scenario asked the subjects to role-play that they were the regional manager whose job included the monthly evaluation of the performance of an interdisciplinary team. The team's objective was to reduce total assembly costs to the value-added level. The makeup of the team, the problem they faced, the decision that was reached, and the expected results were the same across all treatments. Only the actual results varied across the four treatments.

In treatments 1 and 2, the actual results exceeded expected results in the first four periods. In period 5, however, actual results exceeded expected results for treatment 1, but actual results were less than expected results for treatment 2. In other words, subjects in treatment 1 received a confirming outcome for the team in period 5, while subjects in treatment 2 received a disconfirming outcome. In treatments 3 and 4, the actual results were less than expected results in the first four periods. In period 5, however, actual results continued to be less than expected results for treatment 3, but actual results were greater than expected results for treatment 4. In other words, subjects in treatment 3 received a confirming outcome for the team in period 5, while subjects in treatment 4 received a disconfirming outcome. A mapping of the actual outcomes for the treatments is shown in exhibit 2.

Exhibit 2: Treatment Mapping

	Period				
Treatment	1	2	3	4	5
1	A>e	A>e	A>e	A>e	A>e
2	A>e	A>e	A>e	A>e	a<E
3	a<E	a<E	a<E	a<E	a<E
4	a<E	a<E	a<E	a<E	A>e
A > e means that the actual results were greater than expected a < E means that the actual results were less than expected.					

Repeated measures of the evaluations were collected on five consecutive class days. The folders were returned to the subjects with one additional "monthly" evaluation form included. Although the subjects were identified on the folders, they were not identified with the completed evaluation forms. Ten subjects who missed at least one of the five class meetings were excluded from the study.

Subjects were 38 undergraduate students enrolled in principles of managerial accounting at a southwestern university. Exhibit 3 provides details about the subjects.

Exhibit 3: Descriptive Statistics for Sample

Major		**Course Grade**	
Accounting	18	A	7
Economics	1	B	10
Finance	4	C	12
Management	13	D	6
Marketing	2	F	1
Gender		**Ethnicity**	
Female	18	African American	2
Male	20	Hispanic	8
Age		White	26
Range	18 - 47	Other	2
Median	22		
Mean	25		

There were no significant differences in academic ability (as measured by course grade), age, classification (sophomore, junior, etc.), ethnicity, or gender between groups.

As a control to ensure that evaluations across the four treatments would vary only because of the outcome effect, the subjects repeated the decision process faced by the hypothetical group by completing the manipulation check. Subjects who did not reach the same decision as the group were excluded from the study. In total, four subjects were eliminated. Fisher and Selling (1993) show that the effect of outcome information diminishes when the evaluator observes the decision rule of the person being evaluated. Thus, having the subjects repeat the decision process of the group being evaluated could bias the results. However, the purpose of this study is not to test for outcome effects. Instead, given that an outcome effect exists, we are trying to determine whether the impact of the outcome information on the periodic performance evaluation diminishes over time.

Four subjects, two from treatment 2 and two from treatment 4, were eliminated. One subject's final evaluation represented a large deviation in direction from those of the other subjects in the group. That subject's group received information that showed actual results to be much lower than expected. This subject then significantly raised his/her evaluation. The other three Ss were eliminated in order to balance the groups in terms of number of subjects (so five subjects remained in each group). Thus, the final sample size was twenty. Inclusion of the three subjects has no impact on the results.

Results

As a manipulation check to ensure that subjects relied upon the actual results as part of the evaluation process, the data were first tested for the presence of an outcome effect. Table 1 presents an ANOVA comparing the mean evaluations for each of the four treatments in period 1. The overall model is significant ($p = 0.01$). Since treatments 1 and 2 and treatments 3 and 4 are identical in the first period, the following contrast is also tested.

$$L_1 = \bar{t}_{11} + \bar{t}_{21} - \bar{t}_{31} - \bar{t}_{41} = 0$$

where t_{ij} = the mean evaluation for treatment i in period j. The contrast is significant ($p = 0.0033$), thus showing the presence of an outcome effect in the first period.

Table 1: ANOVA for Period 1 of Treatments 1 - 4

Source	df	Sum of Squares	Mean Square	F - value	p-value
Model	3	1406.15	468.71	4.756	0.0148
L_1	1	1170.45	1170.45	11.877	0.0033
Error	16	1576.80	98.55		
Corrected total	19	2982.95			

Is the outcome effect diminishes over time. To examine this issue, the mean evaluations within each treatment are compared over the five periods. Tables 2 - 5 present the ANOVA results for each treatment. The following sixteen orthogonal contrasts are tested to determine the differences in mean evaluations over time within each of the four treatments.

$$L_1 = \bar{t}_{i1} + \bar{t}_{i2} + \bar{t}_{i3} + \bar{t}_{i4} - 4\bar{t}_{i5}$$
$$L_2 = \bar{t}_{i1} + \bar{t}_{i2} + \bar{t}_{i3} - 3\bar{t}_{i4}$$
$$L_3 = \bar{t}_{i1} + \bar{t}_{i2} - 2\bar{t}_{i3}$$
$$L_4 = \bar{t}_{i1} - \bar{t}_{i2}$$

If evaluators do become so committed to their beliefs that they are reluctant to revise evaluation scores, we would not expect any of the contrasts to be significant.

Interestingly, L_1 is significant for treatments 2 and 4 ($p = 0.01$ and $p = 0.01$, respectively), but none of the other fourteen contrasts is significant. Significant revisions only occur when outcomes that contradict prior periods occur. Over time, the β value in equation (2) approaches zero, so that current evaluations are equal to previous evaluations, but the α value in equation (1) remains large.

Table 2: ANOVA for Periods 1 - 5 of Treatment 1

Source	df	Sum of Squares	Mean Square	F – value	p-value
Model	4	44.24	11.06	0.173	0.9500
L_1	1	0.64	0.64	0.010	0.9213
L_2	1	38.40	38.40	0.600	0.4476
L_3	1	4.80	4.80	0.075	0.7870
L_4	1	0.40	0.40	0.006	0.9390
Error	20	1281.60	64.08		
Corrected total	24	1325.84			

Table 3: ANOVA for Periods 1 - 5 of Treatment 2

Source	df	Sum of Squares	Mean Square	F – value	p-value
Model	4	3751.44	937.86	9.146	0.0002
L_1	1	3745.44	3754.44	36.527	0.0001
L_2	1	4.27	4.27	0.042	0.8397
L_3	1	0.83	0.83	0.008	0.9296
L_4	1	0.90	0.90	0.009	0.9254
Error	20	2050.80	102.54		
Corrected total	24	5802.24			

Table 4: ANOVA for Periods 1 - 5 of Treatment 3

Source	df	Sum of Squares	Mean Square	F – value	p-value
Model	4	152.56	38.14	0.133	0.9684
L_1	1	34.81	34.81	0.121	0.7316
L_2	1	30.82	30.82	0.107	0.7470
L_3	1	8.53	8.53	0.030	0.8642
L_4	1	78.4	78.4	0.273	0.6070
Error	20	5741.60	287.08		
Corrected total	24	5894.16			

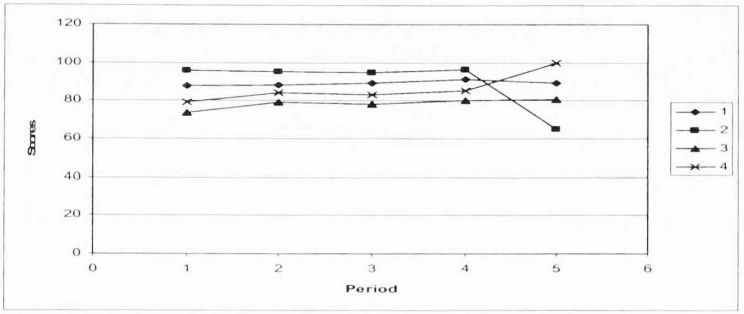
Table 5: ANOVA for Periods 1 - 5 of Treatment 4

Source	df	Sum of Squares	Mean Square	F – value	p-value
Model	4	1294.00	323.50	3.460	0.0265
L_1	1	1190.25	1190.25	12.730	0.0019
L_2	1	33.75	33.75	0.361	0.5547
L_3	1	7.50	7.50	0.082	0.7776
L_4	1	62.50	62.50	0.668	0.4234
Error	30	1870.00	93.50		
Corrected total	34	3164.00			

In other words, when the group reports actual results that are consistent with the prior period's evaluation, no revision in the evaluation takes place. However, when disconfirming results are reported, a significant revision in the current evaluation takes place. Thus, the direction of the evidence seems to be a significant factor.

The impact of direction of evidence is perhaps better illustrated by the graph of the data shown in exhibit 4. The mean evaluations for treatments 1 and 2 and treatments 3 and 4 are relatively equal for the first four periods, as one would expect. In the fifth period, as outcomes that are consistent with previous periods are reported for treatments 1 and 3, the mean evaluations remain unchanged. However, as outcomes that contradict previous periods are reported for treatments 2 and 4, significant revisions of the mean evaluations occur.

Exhibit 4: Mean Evaluations for Treatment *j* in Period *t*



Notice that in the fifth period the mean evaluation for treatment 4, which reported lower than expected outcomes in every period except the last, is even higher than the mean evaluation for treatment 1, which reported higher than expected outcomes in every period. The expected and reported outcomes are equivalent in period 5 for both treatments. A t-test of this difference reveals that it is statistically significant ($t = 2.54$, $p = 0.0174$). Likewise, the expected and actual results for treatments 2 and 3 are the same in the fifth period, but the Ss' mean evaluation for treatment 2 is lower, although the result misses statistical significance ($t = 1.33$, $p = 0.1101$).

DISCUSSION

This study links the research on outcome-based performance evaluation and belief revision by investigating the implications of the Hogarth and Einhorn (1992) model for performance evaluation over a series of time periods. In their model of belief revision, Hogarth and Einhorn (1992) assert that over time, as decision-makers become committed to a particular belief, their sensitivity towards all new information, confirming or disconfirming, will diminish. In the context of the evaluation of team performance in choosing between alternative projects, the results suggest that the sensitivity towards confirming evidence does diminish, but the sensitivity towards disconfirming evidence remains quite strong. Thus, we find that a series of similar performance measures, whether positive or negative, results in flat evaluations. However, if the last measure in the series is either higher or lower than the previous measures, there is a strong effect on the evaluation. In fact, a positive ending measure after a series of lower evaluations results in a higher evaluation than is given by subjects who made a series of higher evaluations and then saw a negative ending measure. This is the fishtail effect of the HE model.

There are at least two important implications from

https://scholars.fhsu.edu/jbl/vol2/iss2/3
DOI: 10.58809/CXRA3757

these results. First, once an evaluatee is given a high (low) performance evaluation based on reported accounting results that are higher (lower) than expected, that person's subsequent performance evaluations are not likely to ever become much higher (lower). On the other hand, an evaluatee with a sequence of high (low) performance evaluations who reports disconfirming accounting results such those results are lower (higher) than expected will likely experience a significant decrease (increase) in performance evaluation score. In other words, a person with a high evaluation has nowhere to go but down, and a person with a low evaluation has nowhere to go but up. This was true in our experiment even though the potential always existed to raise or lower evaluations.

Second, in our experiment, the reported accounting information resulted in high or low performance evaluations regardless of the decision process used or the decision reached by the group being evaluated. This suggests that evaluatees would not be "punished" for poor decisions as long as reported accounting results are initially greater than expected and disconfirming accounting results do not occur. Of course, it would be hard to argue that a decision was poor if actual results are consistently higher than expected. More important, perhaps, is that evaluatees may be "rewarded" for poor decisions if disconfirming information is reported such that actual results are greater than expected and previous evaluations were low. Also, they may be "punished" for good decisions when disconfirming accounting information is reported such that actual results are less than expected and their previous evaluations were high. If evaluatees become aware of this dynamic, those with consistently high evaluations could reasonably be expected to become risk averse and those with low evaluations should be expected to become risk seeking in a manner consistent with the premise of prospect theory (Kahneman and Tverksy, 1979).

These findings have implications for multiple areas of accounting and the other business disciplines. The present research evaluated the performance of teams. Additional research should investigate whether the findings apply to performance evaluation of individuals. For example, audit manager evaluations of junior members could be observed to see if the same pattern of results holds. For another example, one can extrapolate to financial analysts' assessments of quarterly earnings reports of individual companies and investigate whether the same pattern of results would hold.

The authors propose that the findings may generalize from the performance of accounting teams to teams in other disciplines, such as marketing, finance, management, or information sciences. Other general business contexts in which the research paradigm may apply include evaluations in the fields of banking, marketing, and management. For instance, in banking the performance of loan officers is evaluated on the basis of the performance of their loan portfolios. In a marketing context, sales managers evaluate sales personnel. Advertising executives evaluate the work of their staff on advertising. In management, team performance is often the basis of employee performance evaluation.

Despite the belief of Hogarth and Einhorn (1992), that sensitivity towards additional evidence will diminish, our results show that sensitivity towards disconfirming evidence is quite strong. This implies that there may be a decision bias, a single-data point bias, which is remarkably similar to the fundamental attribution error. Nisbett and Ross (1980) proposed that people have a bias to attribute the cause of outcomes to the person rather than the situation. The problem with this is that there can be many situational reasons for the change in outcome, unrelated to actual performance. These situational factors may include unexpected changes in the availability of materials and labor, changes in overall economic conditions, or simple variations in luck and chance. Any of these factors can influence outcomes over short periods of time. The implication for evaluators is that changes in ratings or assessments should be based on more than one data point. Alternatives include using moving averages, for example, as a means of avoiding a single data point bias.

In business, the balanced scorecard is an attempt to weight a variety of outcomes in evaluating overall performance. (Kaplan and Norton, 1996; Ittner et al., 2003). This continuing sensitivity toward disconfirming evidence is clear in business education. Students frequently ask professors to discount an earlier poor grade - pointing out that a more recent higher grade shows "what they are really capable of." (It should be noted that students, in these authors' experience, have never argued the opposite.) Students also tend to rely on situational factors as an excuse for lower grades. It may be that the averaging model commonly used by professors to assign final grades is a way of reducing weight on a lone data point. In a similar fashion, statistical control charts can help to mitigate analysis of a lone data point by showing control limits and

spotlighting whether or not a data point is outside those limits.

Future Research

This study modeled the dynamic process of performance evaluation by examining sensitivity toward new information over a number of time periods. Hogarth and Einhorn (1992) offer no suggestion as to how much time is required for decision-makers to become committed to their beliefs. While this exploratory study used a time line of five periods, future research should examine the effects of outcome information in shorter and longer time horizons to determine if the results shown persist.

Hogarth and Einhorn (1992) state that evaluators' sensitivity to either confirming or disconfirming evidence will vary among individuals as well as different decision contexts. The revisions to prior period evaluations given in equations (1) and (2) are a function of both the evaluator's *sensitivity* towards the current period's reported results and the *subjective weight* placed on that information. If the weight placed on the disconfirming evidence is large, the revision could be significant even if the evaluator's sensitivity towards new information has diminished. This study did not distinguish between the weighting of the evidence and the sensitivity towards the evidence. Future research should investigate the determinants of the sensitivity and subjective weighting variables as well as their interaction in making an evaluation. In addition, the magnitude of the difference from one period to the next may interact with the sensitivity and subjective weight. This magnitude may have a stronger impact than the absolute value of the change alone. It is possible that larger initial differences may strengthen the primacy effect. This kind of research offers more insight into potential biases of managers and the way in which those biases might impact evaluations. For example, Allison et al. (1990) found that outcome bias is dampened when need for accuracy is high. Mackie et al (2001) found that outcome bias occurred when people were favorable toward the process that generated the outcome. These are fruitful areas for further research.

There are several limitations to the present research. Future research should employ nonstudent samples and should include a larger sample size. Secondly, a limitation of the present research is the use of scenarios. There are two alternative approaches to scenarios. One is to create true experiments in which subjects act as evaluators and observe the behavior and outcomes of confederates and evaluate those individuals. The second ideal situation is the use of actual business decision making over a series of time periods. Moving beyond performance evaluations to financial evaluations, it is possible to assess the stock market's reactions to companies' quarterly earnings reports and use market data to test the results. In this domain one might expect to see a recency effect as market expectations either are or are not met.

Another fruitful area for further research is an examination of the impact of changes in direction of the information over time. This could be thought of as mixed (confirming and disconfirming) information. That is, would a "good news" - "bad news" - "good news" series of disclosures eventually lead to a steady-state evaluation, or would recency dominate?

In sum, this study provides insight into the dynamic process of repeated performance evaluations. Surely, this is one of the most complex and stressful tasks facing managers. The results support the Hogarth and Einhorn (1992) contention that positive or negative information received after a baseline evaluation results in a strong contrast effect. However, our results show that the impact of that evidence on a series of previous evaluations does not appear to diminish over time. Instead, it remains strong. An understanding of the decision process used for performance evaluations should help managers in a wide variety of business fields to better determine the impact of outcome information on performance evaluation.

REFERENCES

Agrawal, N., & Maheswaran, D. 2005. Motivated reasoning in outcome-bias effects. **Journal of Consumer Research,** 31: 798-805.

Allison, S., Worth, L., & King, M. 1990. Group decisions as social inference heuristics. **Journal of Personality and Psychology,** 58: 801-811.

Asare, S. 1992. The auditor's going concern decision: Interaction of the task variables and the sequential processing of information. **Accounting Review,** 67: 379-393.

Ashton, R. 1990. Pressure and performance in accounting decision settings: Paradoxical effects of incentives, feedback, and justification. **Journal of Accounting Research,** 28: 148-186.

https://scholars.fhsu.edu/jbl/vol2/iss2/3
DOI: 10.58809/CXRA3757

Ashton, A., & Ashton, R. 1988. Sequential belief revision in auditing. **Accounting Review,** 63: 623-641.

Brown, C., & Solomon, I. 1987. Effects of outcome information on evaluations of managers' decisions. **Accounting Review,** 62: 564-577.

Ashton, A., & Ashton, R. 1993. An experimental investigation of explanations for outcome effects on appraisals of capital budgeting decisions. **Contemporary Accounting Research,** 10: 113-118.

Cheng, M., Schulz, A., Luckett, P., & Booth, P. 2003. The effects of hurdle rates on the level of escalation of commitment in capital budgeting. **Behavioral Research in Accounting,** 15: 63-85.

Emby, C., A. M. G. Gelardi, and D. J. Lowe. 2002. A research note on the influence of outcome knowledge on audit partners' judgments. **Behavioral Research in Accounting**, 14: 87-104.

Fisher, J., & Selling, T. 1993. The outcome effect in performance evaluations: decision process observability and consensus. **Behavioral Research in Accounting,** 5: 58-77.

Frederickson, J. 1992. Relative Performance Information: The effects of common uncertainty and contract type on agent effort. **Accounting Review,** 67: 647-699.

Frederickson, J., Peffer, S., & Pratt, J. 1999. Performance evaluation judgments: Effects of prior experience under different performance evaluation schemes and feedback frequencies. **Journal of Accounting Research,** 37: 151-165.

Ghosh, D., & Lusch, R. 2000. Outcome effect, controllability and performance evaluation of managers: Some field evidence from multi-outlet businesses. **Accounting, Organizations and Society**, 25: 411-425.

Hankinson, H., & Lloyd, S. 1994. Change management. **Handbook of cost management, E2-1-E2-23**. New York: Warren Gorham Lamont.

Highhouse, S., & Gallo, A. 1997. Order effects in personnel decision making. **Human Performance,** 10: 31-46.

Hogarth, R., & Einhorn, H. 1992. Order effects in belief updating: The belief-adjustment model. **Cognitive Psychology**, 24: 1-55.

Ittner, C., Larcker, D., & Meyer, M. 2003. Subjectivity and the weighting of performance measures: Evidence from a balanced scorecard. **Accounting Review,** 78: 725-758.

Kahneman, D., & Tverksy, A. 1979. Prospect theory: An analysis of decision under risk. **Econometrica,** 47: 263-291.

Kaplan, R., & Norton, D. 1996. **The balanced scorecard: Translating strategy into action**. Boston, MA: Harvard Business School Press.

Krishnamoorthy, G., Mock, T., & Washington, M. 1999. A comparative evaluation of belief revision models in auditing. **Auditing: A Journal of Practice and Theory**, 18: 105-127.

Lipe, M. 1993. Analyzing the variance investigation decision: the effect of outcomes, mental accounting, and framing. **Accounting Review**, 68: 748-764.

Luckett, P., & Eggleton, I. 1991. Feedback and management accounting: A review of research into behavioral consequences. **Accounting, Organizations, and Society,** 16: 371-394.

Mackie, D., Ahn, M., Asuncion, A., & Allison, S. 2001. The impact of perceiver attitudes on outcome-biased dispositional inferences. **Social Cognition**, 19: 73-91.

McNair, C. 1994. Responsibility accounting and controllability networks. **Handbook of cost management, E4.1-E4.33**. New York: Warren Gorham Lamont.

Nisbett, R., & Ross, L. 1980. **Human inference: Strategies and shortcomings of social judgment.** Englewood Cliffs, New Jersey: Prentice-Hall.

Reimers, J., & Butler, S. 1992. The effect of outcome knowledge on auditors' judgmental evaluations. **Accounting, Organizations and Society,** 17: 185-213.

Tan, H., & Lipe, M. 1997. Outcome effects: The impact of decision process and outcome controllability. **Journal of Behavioral Decision Making,** 10: 315-325.

Tubbs, R., Gaeth, G., Levin, I., & Van Osdol, L. 1993.

Order effects in belief updating with consistent and inconsistent evidence. **Journal of Behavioral Decision Making,** 6: 257-269.

Tubbs, R., Messier, W., & Knechel, W. 1990. Recency effects in the auditor's belief revision process. **Accounting Review,** 65: 452-460.

Tuttle, B., & Stocks, M. 1998. The use of outcome feedback and task property information by subjects with accounting-domain knowledge to predict financial distress. **Behavioral Research in Accounting**, 10: 76-108.

Wilks, T. 2002. Predecisional distortion of evidence as a consequence of real-time audit review. **Accounting Review**, 77: 51-71.

Russell Calk is an associate professor of accounting at New Mexico State University. He received his Ph.D. in accounting from Oklahoma State University. His current research interests include behavioral accounting, financial planning, and business ethics. He has published in Accounting Education: An International Journal, Journal of Business Inquiry, Southwestern Business Administration Journal, and Journal of State Taxation.

Maryanne Mowen is an associate professor of accounting at Oklahoma State University. She received her Ph.D. in economics from Arizona State University. Her current research interests include the impact of internal disclosures on market valuation post-Sarbanes-Oxley and the impact of time on business decision making. She has published in the Journal of Management Accounting Research, Journal of Economic Psychology, and Decision Sciences. She has also coauthored textbooks in management and cost accounting.

Appendix A

As part of its continuous improvement program, the Southwest Region of Alpha Co. has revised its control framework from the traditional functional structure and has implemented several cross-functional teams. Rather than conducting individual performance evaluations on each team member, each team is evaluated collectively. Performance reviews, bonuses and other incentives of the team members are determined from this team-based performance evaluation. As Regional Manager, you are personally accountable for the performance of the entire region. It is your responsibility to conduct monthly performance evaluations for each team. You are allowed to determine the criteria used in the evaluation.

The team for this study consists of a cost accountant, an industrial engineer and an operations manager. The purpose of the team is to identify ways of improving the existing process of assembling component parts so that the cost is reduced to the value-added level. The engineer's responsibility is to identify alternatives for improving the process. The accountant determines the relevant costs of implementation and cost savings of each alternative. The operations manager determines the likelihood of achieving the estimated cost savings for each alternative. Although the expected net financial benefit should be considered as part of the decision process, the team members are allowed and encouraged to use professional judgment in deciding which alternative to implement.

Appendix B

On June 1, the team met to discuss and choose between two improvement alternatives. Available resources are sufficient to choose either alternative A or alternative B, but not both. Projected costs and benefits for each alternative have been submitted to you as part of the performance evaluation file and are given below.

Alternative A

Month	Projected cost of implementation	Projected cost savings	Probability of cost savings
1	$5,000	$4,000	0.5
2	$4,000	$6,000	0.5
3	$3,000	$10,000	0.5
4	$2,000	$10,000	0.5
5	$1,000	$12,000	0.5
total	$15,000	$40,000	

Alternative B

Month	Projected cost of implementation	Projected cost savings	Probability of cost savings
1	$8,000	$6,000	0.5
2	$6,000	$12,000	0.5
3	$4,000	$12,000	0.5
4	$1,000	$14,000	0.5
5	$1,000	$16,000	0.5
total	$20,000	$60,000	

Calculate the total net expected benefit (loss) from Alternative A.
Calculate the total net expected benefit (loss) from Alternative B.
Which alternative provides the greatest net expected benefit? (Circle one): A B
After considering both alternatives, the team chose to implement alternative B.

Appendix C

For the month of June, the following results were reported to you by the team.

Targeted cost of implementation	Actual cost of implementation	Target cost savings	Actual cost savings
$8,000	$8,000	$6,000	$8,000

Evaluate the team for the month of June on a scale from (0 - 100) were zero is very poor and 100 is excellent.

Team evaluation _____