Seeing What You Want to See: Perceptual Biases of Auditors

Larry M. Parker
Case Western Reserve University

Timothy J. Fogarty
Case Western Reserve University

How auditors perceive evidential material lies at the heart of the value of the audit. The belief that accounting and business information has an objective meaning that can be reliably and consistently interpreted is foundational for the value of the auditor's work-product to users. This study tests this bedrock proposition with an experiment administered to a large number of auditors from several accounting firms. The results show that internal control predispositions are highly consequential for accounts payable judgments. Statistical analysis of the data suggests that auditors seek to avoid cognitive dissonance. They also fall prey to the self-fulfilling prophecy.

INTRODUCTION

Auditing will always be a process of acquiring and evaluating evidence. Relevant evidence that is accurately perceived should serve as the foundation of quality audit opinions, and in turn, contribute to financial statements that correspond to economic reality.

One of the largest threats to the realization of the above objective is the self-fulfilling prophecy. Auditors that have strong expectations about what they will find, tend to perceive confirmation of these expectations, even when the evidence is mixed. To the extent that the self-fulfilling prophecy limits auditor ability to objectivity assess evidence, the quality of professional work may be compromised.

Similar to the self-fulfilling prophecy, cognitive dissonance limits the integrity of evidence evaluation. Decision making proceeds best when multiple data points align in their importance and consequence. When data is internally inconsistent, cognitive functioning can be thought of as in disequilibrium. Auditing does not have strong and multiple cause-effect chains, but rather depends upon the totality of the evidence. When this data refuses to be properly aligned, a loss of confidence about decisions might result.

Although these topics have been part of a larger inquiry into audit judgment over the last quarter century, they have not been specifically considered together. The present research studies the self-fulfilling prophecy and cognitive dissonance as part of an accounts payable confirmation task. Using a broad sample of auditors from many firms, this paper reports on the extent auditors can neutrally evaluate evidence.

The balance of the paper is organized unto four sections. The first provides a brief literature review that grounds the statement of the hypotheses. The means by which the hypotheses were tested is described

in a second section that provides ample detail about measures and design. The results of these tests and a discussion of them are contained in the third and fourth section.

LITERATURE REVIEW AND HYPOTHESES DEVELOPMENT

The collective belief that auditing processes are capable of informing the extent to which financial statements are fairly presented depends upon the extent to which auditors can be systematic about evidence collection, aggregation and interpretation (Joyce, 1976). Central in this regard are the perceptions that underlie concept validity and inter-judge reliability (Einhorn, 1974).

The inability of subjects, both novice and expert, to behave in accordance with statistical models has been well documented for some time (Newell and Simon, 1972; Slovic and Lichtenstein, 1971). Subjects were inadequately appreciative of sample size (Kahnman and Tversky, 1972) and variance (Beach and Scopp, 1968). Early study progressed to the systematic study of auditor decision making. This resulted in the enumeration of several judgment biases including recency, availability, insufficient adjustment from an established anchor, and representativeness (see reviews by Biddle and Joyce, 1978; Johnson et al. 1989; Nelson and Tan, 2005). Auditors act as if they were strong believers in heuristics (Kinney and Uecker 1982) and poor estimators of prior probabilities (Corless, 1972; Felix, 1976). Unquestionably, auditor knowledge and predisposition matter to the quality of the highly specialized judgments that are asked of the practicing auditor (Tan et al. 2002; Hogarth, 1991).

Many studies have manipulated the order of audit cues. Routinely, recency effects have been documented (e.g., Tubbs et al. 1990; Cushing and Ahlawat, 1995). These figure prominently among other order effects (e.g., Ahlawat, 1999; Ahlawat and Fogarty, 2001). Recent work suggests however that these findings may be less endemic, and more specific to the auditor sample involved (e.g., Early 2002).

Past judgment research suggests that auditors are more likely to be influenced by evidence that confirms their original opinions, and somewhat resistant to that which questions those positions (Bamber et al. 1997). Originally, this decisional frame was interwoven with recency effects (e.g., McMillian and White, 1993). Other work has suggested that confirmation bias can exist apart from a particular ordering of audit evidence (Pei et al. 1992). Early et al (2008) formulate the issue more broadly by specifying knowledge effects, wherein auditors are unable to discount previously processed information.

Internal control has served as an important parameter for auditor decision research for some time. Whereas the cue that internal control is strong produces auditor consensus, the idea that it is weak triggers uncertainty and high inter-subject variation (Soloman et al. 1982; Felix and Niles, 1988). Conclusions reached about internal control tend to have persistent and consistent consequences (Ashton, 1974), perhaps even altering the general level of skepticism that auditors bring to their work (Quadackers et al. 2009). Internal control evaluations stand out at all judgment levels (Whittington, 1978), such that halo effects are possible (Mock and Turner, 1981). Although internal control states have traditionally been used to contextualize audit decisions, more recent work has repositioned it as a dependent variable (e,g., Bronson et al. 2006).

In the post-Sarbanes Oxley Act era, internal control quality has increased in its importance. However, most of this work has centered on the separate internal control report required under that act and how that might complicate the interpretation of the audit opinion by external constituents (e.g., Schneider, 2009; Schneider and Church, 2008). Nonetheless, this emphasis should not gainsay the pivotal nature of internal control evaluation for the plan and execution of the audit (Messier, 2009).

The Self-Fulfilling Prophecy

Merton (1957) defined the self-fulfilling prophecy as situations wherein subjects, taken by a false definition of conditions, act in accordance with their illusion and bring into existence the very phenomenon that they had wrongfully first assumed. This chain of events has many applications ranging from the individualistic/personal to the societal.

Disaggregating the self-fulfilling prophecy to its elements allows us a heightened appreciation of its dynamic. Typically, situations are defined based on an authority relationship. Power or influence is seen

to be exerted upon the subject in some sort of coercive, mimetic or normative manner. The initial situation is subsequently observed by others who independently must interpret the original action. This sets in motion a change in substantive conditions that might otherwise not have occurred.

The self-fulfilling prophecy is distinctly a product of selected perception. A full and neutral examination of the facts in a situation would usually reveal the triggering action to be unfounded. One could also argue that self-fulfilling prophecies are the product of idiosyncratic ways in which individuals construct cognitive schema. If these include a fallacious cause-effect linkage, evidence to the contrary will be ignored. Thus, people do not actually see their environment. Instead their perceptions are the result of a systematic filtering process.

Much of the self-fulfilling prophecy occurs without the conscious awareness of the human subject. Small cues perhaps emitted without purpose, initiate subtle changes in reinforcement that have consequences capable of shading results. These patterns are often sufficiently latent and minuscule that they go undetected by the observer who is taken in by alternative explanations.

On a more sociological theme, the self-fulfilling prophecy is encapsulated in W.I. Thomas' quote "If men define situations as real, they are real in their consequences," (Thomas and Thomas, 1928). This reminds us that much about society is an agreed upon state that is not capable of proof apart from the consensus that sustains it. Most prominently, transactions are only possible by virtue of the trust that we are usually willing to inject into situational definition.

The self-fulfilling prophecy has not been ignored in the auditing literature. However, it has tended to be confined to a single issue. Auditors may be reluctant to issue going concern opinions due to the collective belief that doing so is apt to bring about the corporate collapse that the going concern signal puts into issue (Louwers et al. 2003). This connection does not appear to be warranted by the archival evidence (Citron and Taffler, 2001), although collateral actions like auditor switches persist in its believed power (Tucker et al. 2003). Since this work is predicated on macro-level beliefs in "how things work," (see also Guiral-Contreras et al. 2008), this research is different in nature to the micro-level subconscious sense of the self-fulfilling prophecy that this paper employs. The self-fulfilling prophecy has been invoked when auditors confuse their interpretations with the facts (Moeckel and Plumlee, 1989) and as they adopt decision-making strategies that facilitates reaching pre-determined conclusions (Boiney et al. 1997). These studies are more in keeping with the decisional context of the current study.

Auditing provides promising application of these ideas. Corporate clients receiving external audits provide a very complex field from which some cues can be stressed and others readily suppressed. Auditors working under considerable time pressure and other sorts of collateral stress cannot afford the luxury of in depth epistemological exercises.

Cognitive Dissonance

Festinger (1957) conducted lab experiments that formed the basis of our understanding of cognitive dissonance. Subjects strove to escape from situations where they were presented with contradictory information. Because the individual strives for cognitive consistency, dissonance creates psychological discomfort. Avoidance strategies are many, ranging from the evasion of entire situations to irrational perception.

The prospects of dissonance may underlie why individuals do not faithfully evaluate new information and use it to improve their cognitive schema. Seemingly helpful new input is not used in this way because it would conflict with pre-existing beliefs which the new information would otherwise cause to be challenged.

Cognitive dissonance can be understood as the progressive application of constraints. Decision makers seek confirmation for the elimination of certain interpretations of the data. When these constraints rule out previously favored rationales, dissonance occurs. Festinger (1957) suggested that in the face of such a situation, people would cling to one extreme or the other even if that meant being indifferent to inconsistent cues. Thus, the cognitive effort involved in testing initial assumptions can be minimized. Cognitive dissonance is therefore conservative in that it predicts that rationalizations will be at the ready to resist change.

Cognitive dissonance and the self-fulfilling prophecy take two different paths to the same result. Both are built upon strong psychological rigidity that precludes the fair evaluation of new information. Whereas self-fulfilling prophecy stresses the consequences of the action that results from the flawed perception, cognitive dissonance provides more emphasis on the mechanism whereby new information is not taken on its face value.

Auditing provides a fertile ground for cognitive dissonance. Environmental judgments tend to be non-mechanical and difficult to reject by virtue of one exception or another. The audit is a team effort that requires extreme disaggregation such that re-assembly is somewhat esoteric. Whether the parts add up to the whole is not closely observed. In short, the accountability that might check the tendency to make bad decisions as a result of cognitive dissonant conditions tends to be poor.

Hypotheses

The essence of both the self-fulfilling prophecy and cognitive dissonance is the inability to see things as they actually are. In other words, perception is somewhat selective. With regards to the self-fulfilling prophecy, action in the furtherance of a belief has to be responsive to a condition not deemed to be existent by most people. Cognitive dissonance requires the existence of tension between the subject's perception and the subject's idea of reality.

As a threshold approach to the area, a test of the existence of frames of reference is appropriate. The extent to which context matters to professional judgment is essential to the study of perceptual bias. Therefore:

> *H_1: Perceptions of internal control strength are inversely related to errors detected by auditors.*

The essence of the self-fulfilling prophecy is the creation of the imagined state of affairs and its persistence through time. This state, once initiated, is unlikely to be critically re-examined even in the face of contrary evidence.

Strong internal control should make processing errors less frequent and more remediatable in the normal processing of transactions by the client. The existence of errors might cue the reconsideration of the quality of internal control. However, the power of the self-fulfilling prophecy is that one's initial judgment work as a frame of reference for subsequent interpretations. Thus, adjustment in the direction of new information becomes more difficult.

Therefore, attention on individuals whose evaluations of internal control were strong is in order. Those who initially evaluated internal control as weak should be less surprised by evidence of errors discovered by the audit. Those that were not asked, or were not given information about internal control, have not taken the first step toward a self-fulfilling prophecy.

> *H_2: Auditors that perceive a strong internal control environment will persist in that conclusion after reviewing evidence of control errors.*

Another way of appreciating the sequence of information that auditors encounter is through the theory of cognitive dissonance. If audits are aimed at producing opinions about the financial statements, a deep and lengthy chain of evidence would be helpful if it were also consistent. Information that points in different direction will reduce the harmony that can be expected to exist around proper interpretation.

Cognitive dissonance does not predict a particular behavioral reaction to the discomfort that it entails. When cues are seen as incompatible, people might revert to a belief in the first cue or shift their allegiance to this second cue (assuming that escape from the exercise is not possible). Therefore, evidence consistent with cognitive dissonance could be found in the degree that subject belief loses its unanimity.

> *H_3: Auditors that perceive a strong internal control environment will experience more dissonance after reviewing evidence of control errors.*

In sum, the three hypotheses offer us insight into how auditors might react to their own judgments and subsequently encountered evidence. The first issue is how judgments shape the very definition of proper audit work. The second and third provide different yet consistent possibilities for belief revision.

METHOD

Design

Three subject groups were organized for purposes of a between-subjects design. These varied according to the information they received about the state of internal control. Group 1 acted as a control group since its members received no internal control state information. The other two groups were provided with a set of documents suggesting that internal control was weak (Group 2) or was strong (Group 3) from an interval control questionnaire. In order to make this a strong and memorable group partition, subjects assigned to these groups were asked to formulate an internal control conclusion by completing a questionnaire about internal control (see Appendix A).

The experimental materials pertained to the confirmation of accounts payable. Subjects were then given work papers prepared by a hypothetical auditor in which the disposition for eleven separate discovered exceptions were presented. From this, subjects were asked to agree or disagree with the disposition recommended by this individual. In essence, subjects were asked to play the role of an audit supervisor reviewing the work of a junior auditor.

The confirmation of accounts payable was chosen as a task area because it was believed to be within the experiential realm of most practicing auditors. At the same time, the confirmation process was purposefully taken away from its most conventional application (accounts receivable) so as to distance subjects from content areas wherein overly rich preconceptions and firm policies might exist.

After subjects had provided an *ex ante* evaluation of internal control (or been given no such information in the case of the control group) and had judged the propriety of accounts payable confirmation audit work, subjects were asked to again judge the strength of internal control. Since all subjects were exposed to the same information about errors in accounts payable, a revised opinion about internal control strength could be anticipated. The case materials also solicited standard demographic information. Although none of the hypothesis proposed that such variables might quantify the results, an assessment of possible influence was made using this information.

Subjects

In order to mitigate firm or office effects, a broad-based sample of auditors was utilized. In total, auditors from 18 offices of 11 firms in three different states were included. All subjects were auditors, and most had more than two years of experience. This experience level was a desirable feature since it increased the likelihood that subjects were recently familiar with the work that was described in the experimental materials. Subjects were randomly assigned to one of the three treatment groups described above. In total, 215 subjects were asked to complete the task. This process yielded 211 usable auditor responses for the final sample. The fact that all the respondents were auditors is important in light of recent findings that attitudes are patterned by experience and training areas (Pinsker et al. 2009).

The experiment was administered to groups in a classroom setting that was part of the subjects' continuing education obligation. Usually, this was done by one of the authors. Tests were conducted to determine if the results were in any way affected by the differences in who administered the experiment. No significant differences were found. After the completion of the task and the collection of the materials, the substance of the case was discussed with the participants.

The case materials were pretested with students in an auditing course. This resulted in changes to the wording of the instruction and minor modifications to improve the flow of the experimental process. The suggestion to give subjects more time to complete the work (from 30 to 40 minutes) also came from pretesting. The realism of the auditing content had been previously reviewed by both practicing accountants and accounting professors.

Measures

Both initial and *ex post* internal control evaluations were measured with a seven-point Likert scale that varied from 1.0 (very weak) to 7.0 (very strong). The materials presented to auditors about internal control are reproduced in Appendix A. The variation in these descriptions were not made so extreme so as to force subjects to crowd the end positions of this scale. The use of the same scale for both the first and second internal control evaluations enabled their direct comparison.

Subjects were asked to categorize the hypothetical auditor's disposition of each of the accounts payable exceptions as either proper or improper. This forced choice enabled the mathematical aggregation of satisfactory and unsatisfactory resolutions. The latter provides a measure of perceived error (range 0-11) and is particularly important for purposes of the hypothesis tests. All eleven items were adapted from an international firm's training program, and are reproduced in Appendix B.

Dissonance, for purposes of H_3, was measured with the variance of the internal control measure. This anticipates that individual subjects are likely to resolve inconsistent efforts in different ways. Collectively, this should be revealed by less internal consensus in the measure.

Confidence in audit judgments has been regarded as nearly equally important as the substance of those judgments (Pincus, 1991). Auditing as a regrettably incomplete evidential environment, depends heavily upon confidence (Moeckel and Plumlee, 1989). In this study, confidence was used as a manipulation check metric. A single Likert scaled item was used to used to measure auditor confidence.

RESULTS

As expected, group differences were observed in the initial evaluation of internal control. Subjects in Group 3 rated internal control to be much stronger than subjects in Group 2. This design intention proved successful. Subjects in all groups were then asked to evaluate the resolution of accounting payable findings.

Test of Hypothesis 1

Subjects assigned to Group 2 (weak internal control) reported on average of 7.9 errors. This was higher than Group 1 (control group – no internal control information) who averaged 7.5 errors. Group 3, who were led to believe than internal control was strong, averaged 6.9 errors.

Table 1 reports the results of an ANOVA tests of these differences. The significant p-value suggest that group status mattered. Subjects who were led to believe certain attributes about internal control tended to perceive a statistically significant different frequency of reported error ($p<.01$).

The overall significance found in Table 1 necessitates the examination of pairwise difference. The average means were compared via t-tests. Group 1 and Group 3 were significantly different ($t=1.87$, $p<.05$) indicating that those auditors that were told of a strong internal control environment tended to report fewer errors than those told nothing about internal control. *A fortitori*, those told that internal control was strong (Group 3) behaved differently than those that were told that interval control was weak (Group 2). This was an even stronger effect ($t=3.03$, $p<.01$). However, all pairwise differences did not prove statistically different since Group 2 (weak internal control) did not differ from the control group (Group 1: no information) with $t=1.20$, $p>.10$ results.

The results indicate that auditors can vary in their opinions about proper dispositions of accounts payable exceptions discovered during an audit. However, consensus opinions about auditing exist such that "correct" behavior can be stipulated. The experiment provided information about auditor behavior that the firm from whose materials these items were taken would conclude that eight erroneous actions were taken. Therefore, deviations from that point can be taken as movements in an incorrect direction. This suggests that the tendency for Group 3 to report fewer errors in the context of what they believed to be strong internal control is more curious than is the tendency for Group 2 to report more error in what they believed to be weak internal control. The latter is buttressed by the closer approximation of the consensus normative position. On balance, Hypothesis 1 is supported.

TABLE 1

ONE-WAY ANOVA AND DESCRIPTIVE STATISTICS
Initial Evaluation of Internal Control by Subject Groups 1, 2, and 3

Source	df	Sum of Squares	Mean Squares	F Ratio	F Probability
Between groups	2	270.8581	135.429	204.259	0.0000*
Within groups	208	137.9097	0.663		
Total	210	408.7678			

Group	N	Mean	Standard Deviation	Standard Error	95% Confidence Intervals for Mean
1. No internal control information	72	0	0	-	-
2. Weak internal control information	69	2.6422	1.0686	0.1286	2.3955 to 2.9089
3. Strong internal control information	70	5.4429	0.9268	0.1108	5.2219 to 5.6639

Test of Hypothesis 2

The second hypothesis enacted the logic of the self-fulfilling prophecy to the extent that it anticipated that auditors could not be swayed from their initial judgments of a strong internal control.

Upon observation of the errors in the accounts payable area, both treatment groups with prior internal control information lowered their estimates of the strength of internal control. This change was significant ($p<.01$) for both Group 3 (strong internal control) and for Group 2 (weak internal control). However, the parallel movement in this direction was not so powerful as to blur the distinctiveness of the initial group assessments. Group 2 and Group 3 still were significantly different in their final assessment of internal control, with the two groups taking predictably opposite positions straddling the control group.

The adjustment made by Group 3 (strong internal control) was smaller than that of Group 2. This difference is significant at $p<.01$. Group 3 auditors were less able to move away from the position that their initial judgment had established. This supports the existence of a self-fulfilling prophecy and is consistent with Hypothesis 2.

Test of Hypothesis 3

The final hypothesis pertained to the cognitive confusion that might have been created for subjects in the experiment that were given cues leading to disparate conclusions. Auditors in Group 1 were presented information about audit errors in a vacuum of expectations. Shown many accounts payable problems, Group 2 was presented error information consistent with their prior beliefs about weak internal control. However, Group 3 had to face the inconsistency between their *a priori* assessment of internal control and *ex post* errors.

Group 3's initial evaluation of internal control produced a standard deviation of 0.9268, a statistic not materially different than the initial standard deviation of Group 2 of 1.0686. With the mental processing by the subjects of errors in payables, Group 3's standard deviation increased by more than 50% to 1.4374 indicating much high variance in the internal control evaluation. Looking at the same information, Group 2's standard deviation dropped to 0.9996, indicating a stronger consensus that internal control was poor.

The same trajectory can be seen in the standard errors, with Group 3's increasing from 0.1108 to 0.1718, and Group 2's decreasing from 0.1286 to 0.1203. Changes in variability were significantly different for the two groups at $p<.01$. This supports Hypothesis 3.

Manipulation Check

Because subjects were not told that internal control was either strong or weak, but instead had to formulate that conclusion, how assured they were in believing their reported conclusion about internal control is an empirical question. As part of the concluding section of the experiment, subjects were asked about their confidence in their rating of this state. Their answer indicates the firmness of their convictions.

The results suggest that this manipulation was very successful. First no subject in either Group 2 or Group 3 reversed the intended direction. In other words, no one in Group 2 rated internal control as "strong" and no one in Group 3 rated internal control as "weak." However, subjects did vary in the confidence of their conclusion. A strong majority of each group expressed strong confidence in their internal control opinion. Specifically 88% of Group 2 and 83% of Group 3 signified they were somewhere between "confident" and "highly confident." Confidence data was collected on a 5-point Likert scale, with "confident" assigned to the third (middle) position. For these purposes, responses 3-5 were collapsed ("confident," "highly confident," "highest confidence").

Demographic Variables

A large set of personal and organizational differences were collected from subjects as part of the experiment. If the groups differed from each other in these ways, it might cloud the interpretation of the main results. Chi-square tests were conducted for the variables of age, gender, highest degree earned, CPA exam completion status, years experience, years in auditing, size of office, and size of town that the office was located within. None of these variables proved significant, even at the $p<.10$ level.

DISCUSSION

Standard procedure in auditing requires that auditors form an opinion about internal control, so that the magnitude of substantive testing can be properly selected (Hooks, 2011). Following the passage of the Sarbanes Oxley Act in 2002, a heightened focus upon internal control as an irreplaceable element of defense against fraud and misreporting was mandated for publicly traded companies. How internal control states enter into the decision making of auditors should remain the fulcrum of auditing research.

The results of this paper indicate that auditors are prone to a self-fulfilling prophecy in their assessment of evidence. The strength of internal control creates a powerful frame of reference that leads auditors to shade their interpretation of the conduct of the audit. Strong internal control creates a more beneficial tolerance for that which is done with the unusual cases that the audit turns up. Less strong internal control leads auditors to be more circumspect in their evaluations of their own actions.

The results also show how inconsistency in the evidence erodes the consensus that auditors usually bring to their judgments. Shown a long litany of exceptions after they concluded that internal control was strong caused some to being to reverse their positions, but had others more likely to cling to their original position. We need to know more about the group positions that auditors must come to in order that closure on the work be achieved.

In the mostly highly aggregated form, this study suggests that auditors exhibit tendencies to see what they want to see. Auditing can hardly be conceived of as a neutral collection of evidence. Instead, auditing is a rather precarious perceptual process that is subject to bias and artifactual influence.

Another strength of this research is that it allows subjects to form their own conclusions about internal control. Rather than merely telling the auditors that control was strong or weak, the experiment gave them raw materials from which such judgments could be formed. This design feature therefore attempted to mitigate format and aggregation effects (see Agoglia et al. 2009). Auditors may always be subject to what Fischhoff (1977) calls knowledge bias since they enter situations with prior knowledge of previously processed information that they may be unable to ignore in making their own judgments. This

information may come from the client, from last year's engagement team or from the working paper file. What the auditor is given creates a mostly unintentional cognitive effect that might require a restructuring intervention to overcome (Earley et al. 2008).

Limitations

The use of a hypothetical auditor working on the accounts payable exceptions was done to add to the realism of the experiment. However, this may have created other unintended effects. For example, subjects may have sympathized with this actor, especially if they saw him as a inexperienced auditor dealing with matters above his current abilities. This reaction could have produced a leniency effect that might have suppressed subject willingness to identify more errors. Alternatively, subjects may have formed an adverse impression of the hypothetical auditor. This negative "halo effect" might have led to an overly harsh evaluation of his actions, and to the tendency to find more erroneous behavior on the accounts payable. Since these are offsetting biases, the leniency and the halo effect may have cancelled each other, leaving no residual effect on the data. However, since one cannot measure the magnitude of these possibilities in this study, a limitation is appropriate.

The research design did not search for possible experience effects. The literature on this issue is quite mixed, making this issue worthy of focused study rather than an ancillary dimension. Relevant here is differing sensitivity to the ethical environment (Martimov-Bennie and Pflugrath, 2009) and to client-resisted adjustments (Trotman et al. 2009) according to audit experience. Unfortunately, experience does not seem to tell a consistent story with regards to how evidence is viewed (see also Christ, 1993; Asare et al. 2009).

The study also presents subjects with audit situations of a difficulty level which cannot be experimentally manipulated. Although subjects did not perceive the experiment to be difficult to understand, some of the cues given to subjects may have varied in difficulty. This opens the prospect that subjects may have behaved differently had they been presented with a different mix of difficulty in the accounts payable task. Like most experiments using professional subjects, the need to present realistic scenarios, and the constraint that the task require only average degrees of expertise, is a balancing act. Auditor expertise matters in many contexts (Kent et al. 2006).

The design of this research could also be thought of as outcome effects. By injecting a certain set of exceptions after the judgment of internal control, there is a tendency to overly blame the individuals involved. Future research should examine the extent to which this tendency is a function of initial position about the hypothetical level of human agency (see Peecher and Peircey, 2008). Following recent movement in the direction of more actors such as audit supervisory review (Ismail and Trotman, 1995) and insistent audit clients (Wolfe et al. 2009), more elaborate contexts for auditor judgments might be necessary for future extensions.

REFERENCES

Agoglia, C., Beaudoin, C. and Tsakumis, G. 2009. "The Effect of Documentation Structure and Task-Specific Experience on Auditors' Ability to Identify Control Weaknesses." *Behavioral Research in Accounting* 21: 1-17.

Ahlawat, S. 1999."Order Effects and Memory for Evidence in Individual versus Group Decision Making in Auditing." *Journal of Behavioral Decision Making* 12:71-88.

Ahlawat, S. and Fogarty, T. 2003. "An Analysis of Group Influences on Going Concern Auditor Judgments." *Advances in Accounting Behavioral Research* 6: 27-52.

Asare, S., Cianci, A. and Tsakumis, G. 2009. "The Impact of Competing Goals, Experience and Litigation Consciousness on Auditors' Judgments." *International Journal of Auditing* 13: 223-236.

Ashton , Robert H. 1974. “Cue Utilization and Expert Judgments: A Comparison of Independent Auditors With Other Judges.” *Journal of Applied Psychology*, 59: 437-444.

Bamber, E.M., Ramsay, R. and Tubbs, R. 1997. “An Examination of the Descriptive Validity of the Belief-Adjustment Model and Alternative Attitudes to Evidence in Auditing.” *Accounting, Organizations and Society* 22: 249-268.

Beach, L. R., and Scopp, R. S. 1968. “Intuitive Statistical Inferences About Variances.”*Organizational Behavior and Human Performance*, 3: 109-123.

Biddle, Gary C. , and Joyce, Edward J. 1978. “Anchoring and Adjustment in Probabalistic Inference in Auditing.” Unpublished working paper prepared for the Peat, Marwick, Mitchell & Co. Foundation.

Boiney, L. Kennedy, J. and Nye, P. 1997. “Instrumental Bias in Motivated Reasoning: More When More is Needed.” *Organizational Behavior and Human Decision Processing* 72:1-24.

Brophy, J. 1983. “Research on the Self-Fulfilling Prophecy and Teacher Expectations.” *Journal of Educational Psychology* 75(5): 631-661.

Christ, M. 1993. “Evidence on the Nature of Audit Planning Problem Representation: An Examination of Auditor Free Recalls.” *The Accounting Review* 68: 304-322.

Citron, D. and Taffler, R. 2001. “Ethical Behavior in the UK Audit Profession: The Case of the Self-Fulfilling Prophecy Under Going Concern Uncertainties.” *Journal of Business Ethics* 29:353-363.

Corless, John C. 1972.“Assessing Prior Probability Distributions for Applying Bayesian Statistics in Auditing.” *The Accounting Review* 47: 556-566.

Cushing, B. and Ahlawat, S. 1996. “Mitigation of Recency Bias in Audit Judgment: The Effect of Documentation.” *Auditing: A Journal of Practice and Theory* 15: 110-129.

Earley, C. 2002. “The Differential Use of Information by Experienced Novice Auditors in the Performance of Ill-Structured Tasks.” *Contemporary Accounting Research* 19: 595-614.

Earley, C., Hoffman, V. and Joe, J. 2008. “Reducing Management’s Influence on Auditors’ Judgments: An Experimental Investigation of SOX 404 Assessments.” *The Accounting Review* 83: 1461-1485.

Eder, D. 1981. “Ability Grouping as a Self-Fulfilling Prophecy: a Micro-Analysis of Teacher-Student Interaction.” *Sociology of Education* 54: 151-162.

Einhorn, Hillel J. 1974.“Expert Judgment: Some Necessary Conditions and an Example.” *Journal of Applied Psychology*, 59: 562-571.

Felix, W. L. 1976.“Evidence on Alternative Means for Assessing Prior Probability Distribution for Audit Decision Making.” *The Accounting Review*. 800-807.

Felix, W. and Niles, M. 1988. “Research in Internal Control Evaluation.” *Auditing: A Journal of Practice and Theory** 7:43-60.

Festinger, Leon. 1957. *A Theory of Cognitive Dissonance*. Stanford, California: Stanford University Press.

Guiral-Contreras, A. Ruiz-Barbadillo, E. and Rodgers, W. 2008. "To What Extent is the Going Concern Judgment Influenced by the Self-Fulfilling Prophecy?" Available at SSRN http://SSRN.com/ABSTRACT=1267105.

Hogarth, R. 1991. "A Perspective on Cognitive Research in Accounting." *The Accounting Review* 66: 277-290.

Hooks, K. 2011. *Auditing and Assurance Services*. (New York: J. Wiley & Sons).

Ismail, Z. and Trotman, K. 1995. "The Impact of the Review Process in Hypothesis Generation Tasks." *Accounting, Organizations and Society* 20: 345-357.

Jeremias, J. 2001. Cognitive Dissonance and Resistance to Change: The Influence of Commitment Confirmation and Feedback on Judgment Usefulness of Accounting Systems." *Accounting, Organizations and Society* 26:141-160.

Johnson, P., Jamal, K. and Berryman, R.G. 1989. "Audit Judgment Research." *Accounting, Organizations and Society* 14: 83-99.

Joyce, Edward J. 1976. "Expert Judgment in Audit Program Planning." *Journal of Accounting Research*, 14: 29-60.

Jussim, L. 1986. "Self-Fulfilling Prophecies: A Theoretical and Integrative Review." *Psychological Review* 93(4): 429-445.

Kahneman, D., and Tversky, A. 1972. "Subjective Probability: A Judgment of Representativeness." *Cognitive Psychology*, 3: 430-454.

Kent, P., Munro, L. and Gambling, T. 2006. "Psychological Characteristics Contributing to Expertise in Audit Judgment*." International Journal of Auditing* 10: 125-141.

Kinney, W., Vecker, W. 1982. "Mitigating the Consequences of Anchoring in Auditor Judgments." *The Accounting Review*, 57: 55-69.

Louwers, T., Messina, F. and Richard, M. 2007. "The Auditor's Going Concern Disclosure as a Self-Fulfilling Prophecy: A Discrete Time-Survival Analysis." *Decision Sciences* 30:805-824.

Martinov-Bennie, N. and Pflugrath, G. 2009. "The Strength of an Accounting Firm's Ethical Environment and the Quality of Auditor's Judgments." *Journal of Business Ethics* 87: 237-253.

McMillian, J. and R. White. 1993. Auditors' Belief Revisions and Evidence Search: The Effect of Hypothesis Frame, Confirmation Bias, and Professional Skeptism." *The Accounting Review* 68(3):443-465.

Merton, Robert K. 1968. *Social Theory and Social Structure*. New York: The Free Press.

Mock, Theodore J., and Turner, Jerry L. 1979. "The Effect of Changes in Internal Controls on Audit Programs." *Behavioral Experiments in Accounting II*, Thomas J. Burns, ed. Columbus, Ohio: The College of Administrative Science, The Ohio State University, pp. 277-302.

Moeckel, C. and Plumlee, R. 1989. "Auditors' Confidence in Recognition of Audit Evidence." *The Accounting Review* 64:653-666.

Newell, A., and Simon, H. 1972. *Human Problem Solving*. Englewood Cliffs, New Jersey: Prentice-Hall, Inc.

Peecher, M. and Piercey, M. 2008. "Judging Audit Quality in Light of Adverse Outcomes: Evidence of Outcome Bias and Reserve Outcome Bias." *Contemporary Accounting Research* 25: 243-274.

Pei, B. Reed, S. and Koch, B. 1992. "Auditor Belief Revision in a Performance Auditing Setting." *Accounting, Organizations and Society* 17:169-183.

Pincus, K. 1991. "Audit Judgment Confidence." *Behavioral Research in Accounting* 3:39-65.

Pinsker, R., Pennington, R. and Schafer, J. 2009. "The Influence of Roles, Advocacy and Adaption to the Accounting Decision Environment." *Behavioral Research in Accounting* 21: 91-111.

Quadackers, L., Groot, T. and Wright, A. 2009. "Auditors' Skeptical Characteristics and their Relationship to Skeptical Judgments and Decisions." Available at SSRN, http://SSRN.com/abstract=1478105.

Schneider, A. 2009. "Auditors' Internal Control Opinions: Do They Influence Judgments about Investments?" *Managerial Auditing Journal* 24: 709-724.

Schneider, A. and Church, B. 2008. "The Effect of Auditors' Internal Control Opinions and Loan Decisions." *Journal of Accounting and Public Policy* 27: 1-18

Schneider, A., Gramling, A., Hermanson, D. and Ye, Z. 2009. "A Review of the Academic Literature on Internal Control Reporting under SOX." *Journal of Accounting Literature* 27:1-46.

Slovic, P., and Lichtenstein, S. 1971. "Comparison of Bayesian and Regression Approaches to the Study of Information Processing in Judgment." *Organizational Behavior and Human Performance*, 6: 649-744.

Tan, H., Ng, T. and Mak, B. 2002. "The Effects of Task Complexity on Auditors' Performance: The Impact of Accountability and Knowledge*." Auditing: A Journal of Practice and Theory* 21: 81-96.

Thomas, W.I. and Thomas, D. 1928. *The Child in America: Behavior Problems and Progress* (New York: Knopf).

Trotman, K., Wright, A. and Wright, S. 2009. "An Examination of the Effects of Auditor Rank on Pre-Negotiation Judgments." *Auditing: A Journal of Practice and Theory* 28: 191-203.

Tubbs, R., Messier, W. and Knechel, W. 1990. "Recency Effects in the Auditors' Belief Revision Process." *The Accounting Review* 65: 452-460.

Tucker, R. Matseemura, E. and Subramanyam, K. 2003. "Going Concern Judgment: An Environmental Test of the Self-Fulfilling Prophecy and Forecast Accuracy." *Journal of Accounting and Public Policy* 22:401-423.

Uecker, Wilfred, and Kinney, William. 1977. "Judgmental Evaluation of Sample Results: A Study of the Type and Severity of Errors Made by Practicing CPAs." *Accounting, Organizations and Society* 2:269-275.

Whittington, Oliver R. 1978. "An Examination of the Effects of Analytical Evidence and Time Budgets on Audit Judgments." Unpublished Ph.D. dissertation, University of Houston.

Wolfe, C. Mauldin, E. and Diaz, M. 2009. "Concede or Deny: Do Management Persuasion Tactics Affect Auditor Evaluation of Internal Control Deviations?" *The Accounting Review* 84: 2013-2027.

APPENDIX A
INTERNAL CONTROL QUESTIONNAIRE

1. Is the accounts payable department clearly designated as the group responsible for verification and approval of invoices for payment?

2. Is the accounts payable department independent of:
   (a) The purchasing department?
   (b) Other persons requesting specific expenditures?
   (c) Cashier or persons signing checks?
   (d) The receiving department?

3. Do the procedures for verification of invoices in the accounts payable department require a signature on every voucher to show that the following steps have been taken:
   (a) Comparison of quantities billed on the invoice with quantities called for on the purchasing order and shown as received by the receiving report?
   (b) Comparison of the prices, discounts, credit period, and terms of shipment per the invoice with those specified on the purchase order?

4. Is an accounts payable trial balance prepared monthly and balanced to the general ledger control account?

5. Are monthly statements from vendors regularly reconciled with accounts payable ledgers or open vouchers?

6. Are returned purchases controlled in a manner which assures that the correct vendor will be charged for the appropriate amount?

7. Are advance payments to vendors recorded as receivables are controlled in a manner which assures that they will be recovered by offset against vendors' invoices?

8. If payments are made to vendors for shipments direct to customers, are the procedures in use adequate to assure that the sales are billed to customers?

9. Is there an adequate record of open purchase orders and commitments?

10. Are debit memos issued to vendors for discrepancies in invoice prices, quantities, or computation? Are these memos approved by an executive outside the accounts payable department?

11. Are debit balances in vendors' accounts brought to the attention of the credit and purchasing departments?

12. Are there established procedures to call attention to invoices not paid within the discount period?

13. Are all disbursements, excepting petty cash, made by check?

14. Are voided checks mutilated and retained for inspection?

15. Do checks require two signatures? Is there dual control over machine signature plates?

16. Are checks signers persons who have no access to accounting records, cash receipts, or bank reconciliations?

17. Are vouchers and other supporting documents reviewed by the check signer? Is the presignature review effective?

18. Are checks mailed directly by the signer and not returned to the accounts payable department for mailing?

19. Are the voucher and supporting documents stamped or impressed "paid" to prevent duplicate payment before being returned to accounts payable for filing?

APPENDIX B
DETAIL ON AUDITOR ACTIONS AND DISPOSITIONS: ACCOUNTS PAYABLE EXCEPTION WORK

Number	Type/Detail	Consensus Judgment
1.	Unrecorded invoice not traced to reason.	Unsatisfactory
2.	Unrecorded asset and liability for goods in transit for which client's explanation was accepted without documentation.	Unsatisfactory
3.	Materiality judgment made in the field about an error in client's records.	Unsatisfactory
4.	Payment to vendor in transit – no additional audit work done.	Satisfactory
5.	Returned goods verified as journal entry but not checked against shipping documents or traced to credit ledger.	Unsatisfactory
6.	Client list of names and addresses of suppliers obtained but not independently collaborated.	Unsatisfactory
7.	Discounting of invoices accepted per client explanation without evidence of consistently followed policies.	Unsatisfactory
8.	Endorsements of cancelled checks unexamined.	Unsatisfactory
9.	Inability to locate evidence of a return shipment not followed up and resolved.	Unsatisfactory
10.	Travel voucher's amount was transposed – source documents examined.	Satisfactory
11.	Duplication of charge to be paid by client to vendor brought to client's attention.	Satisfactory