

EVALUATION OF THE VALIDITY OF MANAGEMENT REPRESENTATIONS AS AUDIT EVIDENCE

Diane H. Roberts *

Auditors' fraud risk assessment includes detection of deceptive client representations. Prior psychology deception detection research found low accuracy but high confidence that may lead to inappropriate judgments. In an audit context, subjects were more accurate than chance but had substantial errors. Accuracy was greater for overstatements than for understatements. The warning of low evaluated management integrity aided detection and significantly increased assessed fraud likelihood. Confidence was greater with audio tape cues than with transcript-only; however, accuracy was not significantly greater. The proliferation of Internet fraud makes knowledge about the ability to detect deception with only transcript cues important.

INTRODUCTION

SAS No. 82 (AICPA, 1997) formalized auditors' responsibilities for discovering fraud during an audit to include detection of deception in client representations as a component of fraud risk assessment. Inquiry of client personnel has been found to be a cost effective means of discovering potential errors in the audit planning phase (Hylas and Ashton, 1982; Wright and Ashton, 1989); however, deception detection accuracy rates have been found to approximate chance in both everyday and work-related situations. Low deception detection accuracy is a particular concern in communication channels with limited deception cues such as transcripts-only conditions like the Internet, email, and audit workpapers. Poor calibration due to low accuracy and relatively high levels of confidence in deception detection abilities (DePaulo and Pfeifer, 1986) may lead to inappropriate judgements.

The types of lies generally told by clients tend to be those that make the client company appear to be more financially sound than it actually is. If the auditor perceives the client as more honest than the client actually is, then audit procedures may be reduced or misdirected and audit effectiveness compromised. If the auditor perceives the client to be more deceptive than the client really is, then audit testing may be needlessly extended and audit efficiency reduced. Inaccurate deception detection can also create risks in other business contexts.

* Diane H. Roberts is affiliated with the University of San Francisco

This study uses an experimental approach to examine the accuracy of deception detection and judgment confidence in an audit context where one of the SAS No. 82 risk factors, a significant portion of management's compensation is from bonuses, is present. The impact of high or low management integrity and type of communication channel are examined. The communication channels are audio tape and transcripts-only as cues as the possibility of Internet fraud makes knowledge about the ability to detect deception with only transcript cues important. An overall assessment of the likelihood of the presence of fraud is elicited and related to accuracy of deception detection.

The remainder of the paper is organized as follows. The next section provides a review of the relevant literature from the auditing and psychology domains. The third section details the experimental design and the fourth section presents the results. The final section provides conclusions and recommendations for future research.

LITERATURE REVIEW AND HYPOTHESIS DEVELOPMENT

Use of Unaudited Data in Audit Planning

Evidence the auditor obtains during management inquiry has limitations: it has not been observed directly, and the evidence was derived from the entity itself, thus is not independent. Wild and Biggs (1990) derived a Bayesian decision theoretic model to determine the audit impact of utilizing unaudited book values in analytical procedures. The model indicated that reliance on unaudited book values with inaccurate assessments of no intolerable error unnecessarily exposes auditors to increased risk of incorrect acceptance and audit costs. Auditors' reliance upon unaudited management responses to auditor inquiry without accurate error assessment has a similar risk.

Unaudited financial statements can be considered management's representation of the company's financial status. Professional skepticism is used to evaluate any management representation and the possibility of errors or irregularities considered (AICPA, 1997). Determination of an account's expected value should be based upon sources independent of the audited entity, thus theoretically, unaudited book values should not be used to perform preliminary analytical procedures (AICPA, 1980).

Tversky and Kahneman (1974) found decision makers would use a provided value as an anchor for their judgments and then adjust in the correct direction, but insufficiently compared to a value reached without the anchor's presence. Kinney and Uecker (1982) anticipated that auditors would use this heuristic to determine their analytical review investigation boundaries and the anchor would be the current year's unaudited amount. Auditors who received two prior years' audited data and unaudited current year data in the lower (higher) value range established lower (higher) analytical review investigation boundaries. If the unaudited value contains a material error or irregularity, then the non-investigation range based on that value will incorrectly increase the probability of not investigating the account.

A richer data set of five years of prior audited results did not remove the impact of the unaudited book value anchor, although auditors were slightly less biased in the direction of the anchor (Biggs and Wild, 1985). Unaudited data containing any type of trend influenced auditors (Heintz and White, 1989). A decreasing trend had a greater impact than an increasing trend, perhaps due to conservatism. A trend reversal had a greater impact than did a consistent trend.

Auditors with no or incomplete explanations from management had comparable judgments of the most likely cause of a fluctuation observed in analytical procedures (Bedard and Biggs, 1991). Most auditors either disconfirmed or did not adopt management's incomplete explanation and experience measures by rank did not have a significant effect. This suggests that when management fails to supply the correct explanation, intentionally or not, audit efficiency and effectiveness could be reduced.

Deception Detection

Deception detection is a general domain task practiced throughout life. When outcome feedback is unsystematic and the task's structure is poorly understood, learning from experience is unlikely (Einhorn, 1982). The outcome feedback from deception detection tasks can be accurate or can be confounded by either a disbelieving-the-truth error or a believing-a-lie error. Suspected liars cannot be trusted to provide completely honest outcome feedback, thus reliable information for improving deception detection abilities is unavailable. This may partially explain low levels of accuracy and difficulties in development of appropriate confidence.

Since people have generally been found to be inaccurate at detecting deception, psychology researchers have attempted to determine what, if any, cues lead to success in catching liars. Ekman's (1985) theory of communication channels stated that less controllable channels have greater leakage of cues to deception detection and properly attending to these channels may enable greater detection of deception. Facial expressions are readily controllable but other parts of the body and vocal intonations are more difficult to control. The highly controllable verbal content cues of plausibility, consistency, and concreteness (detailed instead of vague) were found to have Pearson correlations between actual truth and subjects' judgments of .42, .24, and .20, respectively (Kraut, 1978). Planned lies were found to be no more detectable than unplanned lies; however, planned responses, whether deceptive or true, were perceived as being more deceptive (DePaulo et al., 1983).

Visual cues, especially facial visual cues, decrease deception detection accuracy (DePaulo et al., 1983; Riggio and Friedman, 1983; Ekman, 1985). Vocal cues were found to be helpful to detect deceit (Bond et al., 1985), while relying only on transcripts was detrimental to detection accuracy (DePaulo et al., 1983). Body viewing conditions, both with and without sound, had significantly better than chance detection accuracy (Riggio and Friedman, 1983). Training that included valid feedback information was more effective than practice without feedback in improving detection accuracy, but improvement did not generalize to evaluating other people's response (Zuckerman et al., 1984).

Studies whose subjects' profession involved deception detection found that job experience did not significantly improve accuracy (Kraut and Poe, 1980; DePaulo and Pfeifer, 1986; Vrij, 1993). U. S. Secret Service agents were the exception with a 64 percent detection rate (Ekman and O'Sullivan, 1991).

Professional customs inspectors may develop detection strategies on the job (Kraut and Poe, 1980); however, the absence of verifiable outcome feedback did not allow any conclusions about the strategies' normativity. Although some experimental settings approximated the work environment (Kraut and Poe, 1980; Vrij, 1993), the accuracy in these studies was not better than in those studies where the experimental task did not attempt to capture job dynamics (DePaulo and Pfeifer, 1986; Ekman and O'Sullivan, 1991).

Law enforcement officers with an average of seven years of experience were not more accurate than students but were significantly more confident in their judgments (DePaulo and Pfeifer, 1986). Confidence and accuracy have not been found to have a consistent relationship over a broad spectrum of contexts and tasks (Koehler, 1991). Peterson and Pitz (1988) hypothesized that this may be due to the absence of a universal relationship between accuracy and confidence. Other studies (Einhorn, 1980; Hogarth, et al., 1991) focused on the difficulties of learning from feedback as a possible reason for inappropriate confidence.

This prior research and discussion suggest the following hypotheses:

H1: Accounting students will be more accurate than chance (50%) in detecting deception in an audit context.

H2: Accounting students who have access to verbal cues will be more accurate in detecting deception in an audit context than those who access to transcripts-only.

H3: Accounting students' confidence in their judgment will be greater than the accuracy level achieved.

EXPERIMENTAL DESIGN

Subjects

The experiment was completed by 153 advanced accounting students in their final semester of undergraduate study. Deception detection is a general domain task and the students' degree course work provides the background to understand the audit/business context of the task.

Task and Procedures

Background information was provided about a hypothetical, continuing audit client with unqualified opinions in prior audits. The moderate internal control and the materiality threshold were unchanged from the prior year and the base rate (50/50) of honest and deceptive client responses was disclosed. Client responses to auditor inquiries were provided for each of two asset and one liability accounts. For each client response, the students judged its truthfulness or deceptiveness and rated their judgment confidence.

JOURNAL OF BUSINESS AND MANAGEMENT

Then, based upon their belief in the veracity of the information received from all the client responses for an account, the students indicated the percentage likelihood of fraud at the client. Lastly, an overall fraud likelihood judgment was made.

Independent Variables

The experiment is a 2 X 2 factorial design. Management integrity and communication channel are between-subjects factors. The red flags approach was used as a theoretical basis for operationalization of the high and low management integrity levels. Red flags are significant items that should alert the auditor to a higher than normal possibility of management fraud, thus they also serve as indicators of a low level of management integrity. The five most significant indicators from Albrecht and Romney's (1986) study were combined to form the low level, and their opposite formed the high level. These indicators included overlooking controls, and domination of the company by individuals who lived beyond their means and who were characterized as 'wheeler-dealers.' A manipulation check was included in the debriefing questions.

The two communication channels were (1) a prerecorded audio tape condition and (2) a transcript-only condition. The difference was limited to the additional verbal cues on the audio tape. All auditor questions of the client and other information were presented in a written format for all conditions.

Dependent Measures

Students indicated honesty or deception judgments on a seven-point Likert scale with end points of "very honest" and "very deceptive." To obtain the accuracy score the student's response was compared to the known truth or deception status of the client's representation and the percentage of correct judgments was computed. The truth or deception status was determined by comparing the client's response to the facts of the experimental case.

Confidence judgments were indicated on a seven-point Likert scale with the end points of "very confident" and "not at all confident." Each scale point was assigned a numerical value and the responses were summed to yield an overall confidence score for each student. Fraud likelihood judgments were the subjective probability judgments of the students stated as percentages.

RESULTS

To verify the manipulation of the management integrity variable, students were asked to indicate their evaluation of the management's integrity level. The manipulation check levels were assigned values as follows: low, 1; neutral, 2; and high, 3. Students in the low management integrity condition assigned a mean of 1.5775 to the case's management and those in the high management integrity condition assigned a mean of 1.9459. These were significantly different integrity levels per the results of an univariate

ANOVA ($F = 7.2449$, $p = .008$) with the intended higher integrity level perceived as higher by the students. Students primarily identified the low integrity level as low; however, in the high condition 21 students rated it as high, 28 as neutral, and 25 as low. It may have been difficult to assign a high integrity rating when the provided deception base rate was 50 percent. Due to this evaluation, analysis was performed for both the experimental management integrity levels (high and low) and the evaluated management integrity levels (high, low, and neutral). Experimental management integrity results are presented unless the evaluated management integrity level analysis was different.

The first hypothesis suggested accounting students would be more accurate than chance accuracy of 50 percent in detecting deception in an audit context. This hypothesis was supported as overall students were more accurate than chance ($t = 6.51$, $p = 0.000$). Mean accuracy levels by account and experimental group are shown in Table 1. On an individual account basis, students were more accurate than chance for each of the asset accounts; however, the accuracy for the liability of accounts payable was not significantly different than chance ($t = 0.89$, $p = 0.374$). As lower liabilities indicate financial soundness, the deceptive responses for the accounts payable attempted to prevent recording an additional amount and represented understatement errors. The deceptive responses for the asset accounts attempted to increase their values and focused on overstatement. Students had difficulty detecting understatement lies but were more successful with overstatement lies.

The highest accuracy level was attained on the first account in the experiment, inventory, and was statistically greater than chance ($t = 7.45$, $p = 0.000$). As this was the initial account evaluated the increased accuracy level can not be attributed to a learning effect. Further evidence of the lack of a learning effect can be seen from the mean accuracy rates achieved for different client personnel shown in Table 2. The second client responding was always the same individual to simulate the controller or primary financial officer who is the main contact for the audit. The first client responding was specific to that account to simulate the individual accountant responsible for that portion of the financial statements. The highest accuracy levels for the second client, repeated respondent are for the first account, inventory, and decline as more responses are provided. Subject fatigue is possible but unlikely as the overall accuracy for the last account is higher than that of the second account. Although accuracy on the accounts payable responses did not exceed chance, accuracy for the notes receivable was significantly greater than chance ($t = 2.43$, $p = 0.016$).

Table 1
Mean Accuracy of Deception Detection by Account

Inventory	Transcript	Audio Tape	Total Integrity Level
High Management Integrity	.5344*	.6244*	.5772*
Low Management Integrity	.6453*	.6179*	.6314*
Total Communication Channel	.5891*	.6201*	.6045*
Accounts Payable			
High Management Integrity	.4954	.4930	.4942
Low Management Integrity	.5463*	.5023	.5240
Total Communication Channel	.5205	.4978	.5092
Notes Receivable			
High Management Integrity	.4764	.5659*	.5200
Low Management Integrity	.5829*	.4895	.5356*
Total Communication Channel	.5290*	.5267	.5278*
All Accounts			
High Management Integrity	.5044	.5697*	.5362*
Low Management Integrity	.5937*	.5403*	.5666*
Total Communication Channel	.5484*	.5546*	.5515*

* Significantly different than chance accuracy of 50 percent.

Greater accuracy for the highest materiality account, inventory, is a positive indicator of audit effectiveness; however, this enhanced accuracy may not be attributable to increased effort on the part of the subjects in response to the significance of the account. The individuals who provided the responses to be evaluated may have been differentially detectable. All client replies to auditor questions were from male respondents for inventory but the first client respondents for the other accounts were female. There is no clear gender effect although the highest level of account accuracy was for the all male inventory account. However, for the accounts payable and notes receivable accounts, there was greater accuracy for the female first client than for the male second client.

Table 2

Mean Accuracy of Deception Detection for Account Specific Client

First Client	Inventory (Male)	Accounts Payable (Female)	Notes Receivable (Female)
Transcript, High Management Integrity	.5813	.5390	.5385
Transcript, Low Management Integrity	.6661*	.5795	.6400*
Audio Tape, High Management Integrity	.6668*	.5403	.7619*
Audio Tape, Low Management Integrity	.6833*	.5426	.5600
Total Accuracy	.6490	.5416	.6232

Mean Accuracy of Deception Detection for Repeated Client

Second Client	Inventory (Male)	Accounts Payable (Female)	Notes Receivable (Female)
Transcript, High Management Integrity	.5090	.4531	.4141
Transcript, Low Management Integrity	.6237*	.5129	.5353
Audio Tape, High Management Integrity	.5816	.4822	.4900
Audio Tape, Low Management Integrity	.5513	.4609	.4205
Total Accuracy	.5658	.4770	.46542

* Significantly different than chance accuracy of 50 percent.

The second hypothesis proposed that accounting students with access to verbal cues will be more accurate in detecting deception in an audit context than those with access to transcripts only. Hypothesis two was not supported as students with access to the audio tape's additional cues were more accurate, but not significantly so per ANOVA ($F = 0.158$, $p = 0.693$). Experimental management integrity level did significantly effect accuracy ($F = 3.972$, $p = 0.048$) as the low management integrity condition had greater total accuracy in all accounts. The interaction of channel and experimental integrity level was also significant ($F = 15.648$, $p = 0.000$). The low management integrity, transcript group had the greatest accuracy for each account and overall. The high management integrity, transcript group was the least accurate on an overall basis and for inventory and for notes receivable.

An additional ANOVA was performed using the evaluated integrity level. The evaluated integrity level was significant ($F = 3.413$, $p = 0.036$); however, the interaction of channel and evaluated integrity level was not significant. The low evaluated integrity level had the greatest accuracy on an overall basis and for inventory and accounts payable. The high evaluated integrity level had the greatest accuracy for the notes receivable account.

There was a greater range of achieved accuracy rates in the transcript condition than in the audio tape condition. The warning provided by the experimental low management integrity condition information apparently aided detection as students in the low

management integrity, transcript condition were the most accurate for all individual accounts and overall with total accuracy of .5937. One of the factors that contributed to their performance is that this group had the best detection rate for the second client, repeated respondent. The lowest accuracy of any group was .5044 for the high management integrity, transcript condition. This group had the lowest accuracy rate for both the second client, repeated respondent and the inventory account where all other groups were their most accurate. For the audio tape condition, the high management integrity group was more accurate at .5697, while the low management integrity group had accuracy of .5403.

Table 3

Evaluations of Client Responses Compared to Base Rate of 50/50 in Numbers and Percentages of Students

	Inventory		Accounts Payable		Notes Receivable		All Accounts	
	Number	Percent	Number	Percent	Number	Percent	Number	Percent
Deceptive	55	36%	35	23%	81	53%	54	35%
Base Rate	25	16%	17	11%	12	8%	3	2%
Honest	73	48%	101	66%	60	39%	96	63%
Total	153	100%	153	100%	153	100%	153	100%

As accuracy rates are near or slightly above chance, the types of judgments subjects are making of the client replies is of interest and are shown in Table 3. The base rate of deception was provided as 50 percent so students who judged half of the responses to be deceptive had base rate judgments. Those students who judged more than half of the responses as deceptive (honest) considered the client to be more deceptive (honest) than the base rate. For the inventory account more students judged the client's responses to be honest (48%) than to be deceptive (36%). Accuracy was greatest for the inventory account, and it had the highest percentage of base rate judgments (16%). The account with the lowest accuracy was accounts payable which maybe due in part to 66 percent of the students judging the client's replies to be honest. Notes receivable was the only account where more people (53%) judged the responses to be more deceptive than honest (39%). On an overall basis, the client personnel's replies were judged to be more honest (63%) than deceptive (35%) with only two percent achieving the base rate.

Table 4

Mean Confidence by Evaluated Integrity and by Account*

Inventory	Transcript	Audio Tape	Total Integrity Level
High Evaluated Integrity Level	4.2662	4.4508	4.3245
Low Evaluated Integrity Level	3.9174	4.5497	4.2511
Neutral Evaluated Integrity Level	3.1846	4.2659	3.8643
Total Communication Channel	3.9222	4.4416	4.1770
Accounts Payable			
High Evaluated Integrity Level	4.3885	4.4792	4.4171
Low Evaluated Integrity Level	3.8294	4.6471	4.2609
Neutral Evaluated Integrity Level	3.2062	4.1859	3.8220
Total Communication Channel	3.9414	4.4684	4.1959
Notes Receivable			
High Evaluated Integrity Level	4.2704	4.7967	4.4366
Low Evaluated Integrity Level	3.7703	4.6600	4.2399
Neutral Evaluated Integrity Level	3.1800	4.3091	3.8897
Total Communication Channel	3.8579	4.5658	4.2069
All Accounts			
High Evaluated Integrity Level	4.3065	4.5742	4.3911
Low Evaluated Integrity Level	3.8382	4.6171	4.2493
Neutral Evaluated Integrity Level	3.1685	4.2568	3.8526
Total Communication Channel	4.0800	4.4569	4.1907

*Values Assigned to Likert Scale Points: Very Confident, 6; Moderately Confident, 5; Weakly Confident, 4; Neutral, 3; Weakly Not Confident, 1; and Not At All Confident, 0.

The third hypothesis proposed that accounting students' confidence in their judgement will be greater than the accuracy level achieved. Hypothesis three was supported as the ANOVA results reveal that communication channel had a significant effect on confidence levels ($F = 14.177$, $p = 0.000$); however, communication channel had no significant effect on accuracy. Students with access to the audio tape's additional vocal cues were significantly more confident than those with transcripts. The enhanced confidence from the extra cues in the audio tape communication channel was not accompanied by an enhanced accuracy effect so it may be inappropriate or excessive confidence. Correlation of total accuracy and total confidence was .2227 and was significantly different from zero ($p = .006$). This positive correlation indicates that as confidence increases, accuracy increases weakly.

The experimental management integrity level did not significantly impact confidence ($F = 0.794$, $p = 0.374$) but the students' evaluated management integrity level did significantly effect confidence levels ($F = 5.307$, $p = 0.006$). Confidence levels ranged from above neutral (3) to below moderately confident (5) as shown in Table 4. The lowest confidence level was for the neutral evaluated integrity group for all accounts. If the neutral evaluation represents an inability to evaluate the management integrity as either high or low instead of an actual assessment, the comparative lack of confidence is appropriate. For the transcript condition, the greatest confidence was for the high evaluated integrity group. Dealing with a more reputable company gave more confidence

JOURNAL OF BUSINESS AND MANAGEMENT

in assessing responses provided in an impoverished deception cue set. In the audio tape condition, the most confidence was for the low evaluated integrity group on an overall basis and for each account except notes receivable.

A fraud likelihood judgment was elicited after all client representations. ANOVA analysis did not show any significant effect of experimental management integrity level or communication channel. The evaluated integrity level did have a significant effect on the likelihood of fraud per ANOVA results ($F = 3.589, p = 0.032$). Students who judged the client's management integrity level to be lower had higher fraud likelihood judgments as shown in Table 5. There was a significant interaction between evaluated management integrity and communication channel ($F = 3.125, p = 0.047$). Audio tape cues led to more extreme fraud assessments with the highest (lowest) fraud rating on an overall basis in the low (neutral) integrity, audio tape condition. Correlation of the probability of fraud and total accuracy level was a very low .0971 and was not significant ($p = 0.257$).

Table 5

Mean Fraud Likelihood Judgments by Evaluated Management Integrity Level

	Transcript	Audio Tape	Total Integrity Level
High Evaluated Management Integrity	.5364	.5000	.5250
Low Evaluated Management Integrity	.6103	.6789	.6493
Neutral Evaluated Management Integrity	.6417	.4550	.5250
Total Communication Channel	.5881	.5845	.5885

On an overall basis, higher probability of fraud, .6493, was judged in the evaluated low integrity condition compared to the .5250 assessed by those who evaluated the integrity as high or neutral. These fairly high judgments of fraud may be partially attributed to the task. Students were required to focus directly on deceptiveness and were informed of the 50 percent base rate of deceptiveness in the experiment. Also, students and auditors with lesser amounts of experience have been shown to have more conservative judgments in the evaluation of management integrity compared to more experienced auditors (Roberts, 1997).

CONCLUSIONS

Accuracy levels achieved were greater than chance accuracy of 50 percent; however, a substantial error percentage remains (between .4063 and .4956 depending on experimental condition). This is consistent with accuracy levels obtained in other deception detection studies using non-business task contexts. The absence of outcome feedback in the experiment is similar to real life deception detection. Any real world outcome feedback has significant lack of reliability concerns, thus people can only act upon their subjective belief in the veracity of the information they receive. Auditors

should combine deception detection in client representations with other techniques when implementing SAS No. 82's (AICPA, 1997) fraud risk assessment. Considerable caution should be exercised before relying on or acting upon unsupported management representations.

When the client's responses were judged as more deceptive than the base rate, there was excessive disbelief of client representations. In an audit context this may lead to an unnecessary increase in audit work, thus reducing audit efficiency. When client responses were judged more honest than the base rate, the client was able to successfully misdirect the subjects. If successful misdirection from a sensitive client issue is achieved, then audit work maybe reduced. Audit effectiveness is impaired and the risk of an audit failure may be increased.

The proliferation of Internet commerce and the possibility for fraud makes knowledge about the ability to detect deception with only transcript cues important. Although there was no significant communication channel effect on accuracy, both the experimental and the evaluated management integrity levels did have an effect. As the experimental low management integrity warning did aid detection of deception in the transcript condition, it highlights the importance of knowing the background or business reputation of the entity one is dealing with. Under current electronic commerce conditions it is problematic how a user/consumer obtains information about management integrity level. This indicates a significant practice development opportunity for accountants, one which the AICPA has started to tap with its Web Assure product.

Students were more successful in the detection of overstatement lies for the asset accounts than with understatement lies for accounts payable. DeFond and Jiambalvo (1991) found more overstatement of net income errors reported as prior period adjustments. Understatement of liabilities is often accompanied by understatement of expenses, resulting in overstatement of net income. Prevention of liability understatement is thus an important audit issue.

The degree of confidence a person has in their judgment is a significant component of how willing they are to act upon their belief. There was greater confidence than accuracy especially for the audio tape conditions where the additional cues provided as added level of comfort. The transcript condition had lower confidence than the audio tape condition. This was a more realistic assessment and perhaps may indicate a more cautious approach to judgment. The transcript condition's highest confidence was for the high evaluated integrity group which again indicates the importance of knowing about the company one is dealing with.

The greatest (least) accuracy was in the most (least) material account which raises the issue of whether materiality affects accuracy. Future research could directly address this issue by having the same individual(s) provide the responses to be evaluated for all materiality levels. This would remove the effect of the differential delectability of the people making the responses. Deception detection could also be studied in other business or accounting contexts, with income tax being a particularly interesting area.

REFERENCES

- Albrecht, W.S. & Romney, M.B. (1986). "Red-Flagging Management Fraud: A Validation." *Advances in Accounting*, 3, 323-333.
- American Institute of Certified Public Accountants. (1997). *Statement on Auditing Standards No. 82, Consideration of Fraud in a Financial Statement Audit*. New York, NY: AICPA.
- _____, (1980). *Statement on Auditing Standards No. 31, Evidential Matter*. New York, NY: AICPA.
- Bedard, J.C. & Biggs, S. F. (1991). "The Effect of Domain-Specific Experience on Evaluation of Management Representation in Analytical Procedures." *Auditing: A Journal of Practice & Theory*, 10 (Fall), 77-90.
- Biggs, S. F. & Wild, J. J. (1985)., "An Investigation of Auditor Judgment in Analytical Review." *The Accounting Review*, LX (October), 607-633.
- Bond, C. F., Kahler, K. N. & Paolicelli, L. M. (1985). "The Miscommunication of Deception: An Adaptive Perspective." *Journal of Experimental Social Psychology*, 21, 331-345.
- DePaulo, B., Lanier, M. K. & Davis, T. (1983). "Detecting the Deceit of the Motivated Liar." *Journal of Personality and Social Psychology*, 45, 1096-1103.
- _____, & Pfeifer, R. L. (1986). "On-the-Job Experience and Skill at Detecting Deception." *Journal of Applied Social Psychology*, 16, 249-267.
- DeFond, M. L. & Jiambalvo, J. (1991). "Incidence and Circumstances of Accounting Errors." *The Accounting Review*, 66 (July), 643-655.
- Ekman, P. (1985). *Telling Lies*. New York: W. W. Norton & Company.
- _____, & O'Sullivan, M. (1991). "Who Can Catch a Liar?" *American Psychologist*, 46(September), 913-920.
- Einhorn, H. J. (1980). "Overconfidence in Judgment." *New Directions for Methodology of Social and Behavioral Science*, Vol. 4, Pp. 1-16. San Francisco: Jossey-Bass, Inc.
- _____. (1982). "Learning from experience and suboptimal rules in decision making." In *D. Kahneman, P. Slovic, and A. Tversky (Eds.) Judgment under uncertainty: Heuristics and biases*. Pp. 268-283. Cambridge: Cambridge University Press.
- Heintz, J. A. & White, G. B. (1989). "Auditor Judgment in Analytical Review--Some Further Evidence." *Auditing: A Journal of Practice & Theory*, 8 (Spring), 22-39.
- Hogarth, R. E., McKenzie, C. R. M., Gibbs, B. J., & Marquis, M. A. (1991). "Learning From Feedback: Exactingness and Incentives." *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 17 (4), 734-752.
- Hylas, R. E. & Ashton, R. H. (1982). "Audit Detection of Financial Statement Errors." *The Accounting Review*, LVII (October), 751-765.
- Kinney, Jr., W. R. & Uecker, W. C. (1982). "Mitigating the Consequences of Anchoring in Auditor Judgments." *The Accounting Review*, LVII (January), 55-69.
- Koehler, D. J. (1991). "Explanation, Imagination, and Confidence in Judgment." *Psychological Bulletin*, 110 (3), 499-519.
- Kraut, R. E. (1978). "Verbal and Nonverbal Cues in the Perception of Lying." *Journal of Personality and Social Psychology*, 36(4), 380-391.

- _____, & Poe, D. (1980). "Behavioral Roots of Person Perception: The Deception Judgments of Customs Inspectors and Laymen." *Journal of Personality and Social Psychology*, 39 (5), 784-798.
- Peterson, D. K. & Pitz, G. F. (1988). "Confidence, Uncertainty, and the Use of Information." *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 14 (1), 85-92.
- Riggio, R. E. & Friedman, H. S. (1983). "Individual Differences and Cues to Deception." *Journal of Personality and Social Psychology*, 45 (4), 899-915.
- Roberts, D. H. (1997). "Audit Experience: Aid or Detriment to Evaluation of Management Integrity?" *Proceedings of the Institute of Business Administration and Technology 1997 Conference Pp. 20-28*, London, UK (July 7-14).
- Tversky, A. & Kahneman, D. (1974). "Judgment Under Uncertainty: Heuristics and Biases." *Science*, (September 27), 1124-1131.
- Vrij, A. (1993). "Credibility Judgments of Detectives: The Impact of Nonverbal Behavior, Social Skills, and Physical Characteristics on Impression Formation." *The Journal of Social Psychology*, 133 (5), 601-610.
- Wild, J. J. & Biggs, S. F. (1990). "Strategic Considerations for Unaudited Account Values in Analytical Review." *The Accounting Review*, 65 (January), 227-241.
- Wright, A. & Ashton, R. H. (1989). "Identifying Audit Adjustments with Attention-Directing Procedures." *The Accounting Review*, LXIV (October), 710-728.
- Zuckerman, M., Koestner, R. & Alton A. O. (1984). "Learning to Detect Deception." *Journal of Personality and Social Psychology*, 46 (3), 519-528.