

University of Mississippi

eGrove

Proceedings of the University of Kansas
Symposium on Auditing Problems

Deloitte Collection

1-1-1988

Auditor evidential planning judgments

Arnold Wright

Theodore J. Mock

Follow this and additional works at: https://egrove.olemiss.edu/dl_proceedings



Part of the [Accounting Commons](#), and the [Taxation Commons](#)

Recommended Citation

Auditing Symposium IX: Proceedings of the 1988 Touche Ross/University of Kansas Symposium on Auditing Problems, pp. 104-114;

This Article is brought to you for free and open access by the Deloitte Collection at eGrove. It has been accepted for inclusion in Proceedings of the University of Kansas Symposium on Auditing Problems by an authorized administrator of eGrove. For more information, please contact egrove@olemiss.edu.

6

Auditor Evidential Planning Judgments

Arnold Wright

Northeastern University

Theodore J. Mock

University of Southern California

Abstract

The effectiveness and efficiency of an audit rests largely on the nature and extent of evidence gathered, yet there is little research on how auditors make such complex judgments required to plan audits. This study examines the evidential planning decisions of 21 experienced auditors in an experimental setting. The Analytical Hierarchy Process is employed to explicitly investigate the multi-attribute trade-offs made in such judgments.

The results indicate that auditors displayed strong consensus as to the relative importance of key criteria suggested in the professional literature to weigh evidential alternatives. Specifically, competence was considered of greatest concern, followed by sufficiency and costs of gathering the evidence. But in applying criteria to an audit case, the auditors reached quite different conclusions regarding the relative superiority of alternative procedures when evaluated along the various criteria. Differences were also observed concerning the appropriate allocation of audit time, suggesting substantial variations among auditors in the planned portfolio of procedures across engagements.

Introduction

Evidential planning regarding the nature, extent, and timing of procedures entails critical judgments that greatly impact audit effectiveness and efficiency. The audit planning process should result in a cost-effective portfolio of procedures which are likely to identify material errors at an acceptably low level of audit risk.¹ In deciding upon an appropriate plan, a number of broad categories of procedures, such as detailed tests, analytical review, and observation are normally available. These procedures vary qualitatively and quantitatively along a number of criteria such as competency, sufficiency and cost [SAS 31, AICPA, 1987]. The auditor's task is to select a combination of these procedures to conduct in order to gather sufficient, competent evidence to support an overall opinion on the financial statements. Therefore, evidential planning judgments represent complex, multiple-criteria decisions.

¹ Audit risk is the risk that the financial statements are materially misstated without the auditor's knowledge (SAS 47).

Despite the importance of these decisions, there has been little empirical evidence of how such judgments are made in practice. As a result, there are many unanswered questions, including: Do auditors explicitly consider and weigh evidential criteria? If so, how? What is the relative importance attached to various criteria? How are common procedures evaluated along criteria? What is the level of consensus among auditors in planning judgments? Research into these types of questions may suggest useful decision tools to assist the audit judgment process.

The purpose of this paper is to acquire a better understanding of the multi-attribute trade-offs made in evidential planning decisions. Specifically, the study addresses three implicit/explicit auditor judgments embodied in the planning process: (1) the weighting placed on widely-cited evidential criteria; (2) the evaluation of alternative audit procedures along these criteria; and (3) the resulting allocation of audit hours to the procedures. The level of consensus of each of these judgments is also explored.

To illustrate, assume an auditor considers three criteria to be important in selecting audit procedures: competency, sufficiency, and cost. The relative importance (weightings) of each of these criteria in a given situation will significantly impact the final evidential choices made. If competency is considered of primary importance, with sufficiency and cost of little concern, the auditor is likely to search for the evidence of highest quality (competence) with little regard for availability or costs. After evaluating competing procedures along each pertinent criterion, the final decision is to allocate available resources (e.g., audit hours) among the various procedures.

The next section of the paper contains an overview of evidential planning and the prior research in this area. The methodology and results of this study are then described, with the final section devoted to a discussion of the major results and their implications for future research and practice.

An Overview of Evidential Planning

Figure 1 provides a model of evidential planning factors and considerations. The auditor's overriding goal is to gather evidence through tests to address various audit objectives and thus be able to express opinions for both the accounts (micro-level) and for the overall financial statements (macro-level). As depicted in Figure 1, there are frequently several evidence sources to achieve a given audit objective. For example, the "existence" assertion of an accounts receivable may be tested by sending confirmations, examining subsequent cash receipts, or looking at shipping/sales documents. Auditors must decide which of these procedures to conduct, i.e., the "nature" of the tests. Of course, any or all of the procedures may be planned, since often some degree of corroborating evidence is sought. Once the choice of procedures is established, the auditor must determine the extent and timing of tests.

The choice of procedures is implicitly a multi-attribute judgment. Based on a review of SAS 31, a number of salient criteria for selecting audit procedures are given in Figure 1. There are a few required procedures from the professional standards, such as inventory observation and receivable confirmations. Such requirements are, however, minimal and most audit procedures performed on an engagement are the result of choosing among alternative

summary, in planning the nature, and subsequently the extent of procedures, the auditor is implicitly or explicitly weighing a number of multiple criteria.³

As discussed earlier, there has been limited empirical research on evidential planning. Lewis et al. [1983] asked auditors to allocate budgeted hours to various procedures. An experimental group was provided with a decision aid (Analytical Hierarchy Process—AHP) as a means of structuring the judgment process. The results indicated that the decision aid significantly affected planning decisions. The experimental (decision aid) group allocated more hours to analytical review and less to detailed tests than the unaided control group. This greater focus towards analytical review is consistent with the current trend in the auditing profession towards such tests as a cost-effective means of detecting material errors [Hylas and Ashton, 1982; SAS 56 *Analytical Procedures*, AICPA, 1988; Tabor and Willis, 1985].

Arrington et al. [1984] examined the choice of analytical review procedures, given explicit consideration of five criteria: effectiveness attributes—statistical performance, model robustness, and understandability; and efficiency attributes—cost and ease of application. Using AHP, three academicians who had published research in the area, and three experienced auditors evaluated five alternative analytical review approaches along these criteria. The attributes considered of greatest importance were statistical performance and model robustness. Subjects differed a great deal in their evaluations of each analytical review approach on several of the criteria, resulting in a lack of consensus as to preferences of approaches. In general, the practitioner preferences displayed a narrower range than those of academicians with a random-walk model favored overall. In contrast, academicians preferred a regression approach with random-walk being the second choice. Despite the small sample size employed, the lack of consensus in evidential preferences was perhaps the key finding of this study, and one of concern.

Other studies have examined evidential planning judgments in response to changes in risks. However, such studies have focused on the extent of testing judgments and have not explored the underlying multi-attribute considerations of such judgments. For example, Mock and Turner [1981] examined the sample size (extent) decisions of auditors in a realistic case for four procedures in the revenue cycle. They found a low level of consensus among subjects and the existence of a significant anchoring effect on the initial, planned sample size. Auditors were responsive to changes in the internal controls, planning larger samples when controls deteriorated. Joyce and Biddle [1981] also studied sample size decisions when controls varied. Consistent with Mock and Turner, the results indicated that auditors adapt samples to the controls. However, a significant control by order effect was present, suggesting that auditors recognize trends in controls and apply other heuristics from experience. For

³ In many situations, the choice of tests may seem automatic or obvious given the audit objective. For example, to determine the existence of petty cash, a count of the fund is normally done by the auditor. Even though this procedure appears to be evident, various criteria are implicitly considered. That is, why even conduct the test unless it meets minimal standards as to relevance, reliability, etc.? Also, there usually are alternative tests that could be conducted; e.g., the custodian of the fund could be asked to sign a representation letter attesting to the fact that the petty cash fund in question does exist or merely be asked (inquiry) whether it exists.

example, when controls become stronger, there is a reluctance to reduce samples. However, deteriorating controls led to substantial increases in samples. Joyce and Biddle also found significant variation among participants in the planned levels of testing.

Although the research to date suggests that the evidential planning decisions of auditors are responsive to client and other risk changes, a concern is the low level of observed consensus in such judgments. Importantly, few of the prior studies examined the multi-attribute nature of these decisions, which is the focus here.

Methodology

Task

Twenty-one practicing auditors were presented with a comprehensive, realistic case ("Modern Appliances Manufacturing Co.") and were asked to evaluate alternative evidential sources along various key criteria for the inventory account.⁴ Subjects then decided which evidence to focus on by allocating 150 available audit hours to three broad evidential areas: analytical review, physical observation, and detailed tests.

As a common frame of reference, the case provided illustrative audit programs for each evidential area. Since there are numerous procedures that may fall under each of these areas, it was believed that a benchmark program was necessary to reduce confusion and avoid serious confounding of the results. For example, what is "analytical review"? Different auditors may have various images about what constitutes necessary analytical review procedures for this case, e.g., ratios, regression, and/or industry comparisons. The programs provided were developed with the consultation of practicing auditors and were later pilot tested. The procedures appear to be representative of widely-used tests for a manufacturing client with strong controls, as in the case here. To maintain task simplicity and minimize required subject time, the study examined audit planning judgments for these major evidence areas rather than the selection of detailed individual audit procedures.

The Modern Appliances case contained extensive background information necessary to plan substantive tests. First, information on the client, including product lines and comparative financial statements, was provided. Second, the inventory/purchases internal control system was described in detail, reflecting an environment of strong controls. Compliance tests further revealed that controls were functioning properly.⁵

⁴ Due to the exploratory nature of this study, auditors evaluated evidence for each criterion as related to the overall audit objective that the inventory account was "fairly presented." In practice, evidence may be evaluated along these criteria for each detailed audit objective such as existence and valuation. As will be described later, incorporating specific audit objectives would have geometrically expanded the subject time needed and resulted in having to significantly narrow the scope of the research, perhaps to addressing only one detailed objective. Given the early state of our knowledge here, it was decided that a broader focus was appropriate. However, future research is needed to address specific audit objectives and verify the generalizability of the major results.

⁵ A copy of the complete case may be obtained from the authors upon request.

The initial evaluation of evidence alternatives was based upon criteria cited in Statement on Auditing Standards (SAS) 31. Specifically, subjects were randomly assigned to two experimental groups, varying as to the number of criteria considered. The first group focused on three criteria (cost, sufficiency, and competence), while the second group examined five criteria (cost, sufficiency, relevance, reliability, and bias). The last three criteria are a finer partitioning of “competence.” Placing the subjects into two groups provided the opportunity to examine whether evidential choices are affected by the number of criteria considered. Definitions of each criterion were provided to subjects from SAS 31, as indicated in Table 1.

In considering the criteria, auditors made two sets of judgments for the case: (1) establishing the relative importance of each criterion, and (2) evaluating the three evidential choices along the various criteria. For example, a subject would first assess the relative importance of the criterion , “cost of

Table 1
Evidential Criteria

SET A	
1. Cost:	The additional cost of obtaining the audit evidence which is being evaluated
2. Competency:	<p>The overall quality of audit evidence, which is based on two general factors:</p> <ul style="list-style-type: none"> a) accurate measurement (valuation) resulting from lack of bias (preparer influence) and reliability (accurate accounting system), and b) relevance: the pertinence of the evidence to the audit objective examined.
3. Sufficiency:	The quantity or “weight” of evidence relative to what is needed to satisfy audit objectives. Audit evidence is usually considered sufficient if it is persuasive rather than convincing.
SET B	
1. Cost:	The additional cost of obtaining the audit evidence which is being evaluated.
2. Bias:	The amount of error or misstatement in audit evidence which may result from preparer influence (e.g., management)
3. Reliability:	The amount of error in audit evidence which is a result of inaccuracies in measuring and compiling data.
4. Relevance:	The pertinence of the evidence to the audit objective examined.
5. Sufficiency:	The quantity or “weight” of evidence relative to what is needed to satisfy audit objectives. Audit evidence is usually considered sufficient if it is persuasive rather than convincing.

gathering the evidence” as compared to other criteria. Then, he would evaluate the merits of the audit procedures for each criterion. “In terms of *cost*, which evidence (analytical review, observation, or detailed tests) is preferable?” As will be described in a later section, subjects used the Analytical Hierarchy Process [Saaty, 1980] to arrive at both of these judgments.

Evidential planning decisions are, thus, viewed as a function of the relative weighting placed on key criteria and the judged superiority of alternative sources of evidence on each of these criteria. Finally, subjects planned the allocation of efforts (audit hours) for the three alternative procedure areas (evidence sources). Auditors were allowed to take whatever time was needed to complete the task and anonymity was guaranteed.

Subjects

Participants were from three of the Big Eight firms. Subjects were provided on the basis of time availability and, thus, do not represent a random sample. Table 2 reports demographic data on the participants. As indicated, auditors had, on average, over five years of public accounting experience and were primarily at the supervisory and managerial levels. Therefore, subjects had the extensive experience and background necessary for the tasks examined—the planning of substantive procedures and allocation of audit time. A one-way ANOVA revealed no significant differences ($p < 0.10$) between experimental groups as to years of audit experience. A Chi-square test also did not reflect significant differences in staff level.

Analytical Hierarchy Process (AHP)

Subjects made evidential judgments on a computer terminal in an interactive mode, utilizing the AHP developed by Saaty [1984, 1980, 1978]. AHP is a systematic, multiple criteria method for making unstructured decisions. A judgment is decomposed into a hierarchical framework—from the most general level to specific choices. The decision maker then evaluates criteria/alternatives at each level through a series of pairwise comparisons. For example, at the most general level the subject would be asked, “Which criterion is more

Table 2

Demographic Data on Subjects

Experimental Group	n	Mean Experience	Staff Level	Frequency
Three Criteria	11	5.4 Years (3-10 Years)	Seniors Supervisors Managers	9% 36% 55%
Five Criteria	10	5.5 Years (3-10 Years)	Seniors Supervisors Managers	10% 20% 70%

Note: Numbers in parentheses indicate range of experience.

important—cost or competence?’’ The individual then indicates his or her degree of preference (weights) on a scale of one to nine (equal importance—absolute importance). The number of pairwise comparisons represents every combination of criteria. Thus, the three criteria group made six comparisons (three criteria taken two at a time) and the five criteria group made ten comparisons. Once these sets of comparisons were completed, subjects were asked to evaluate pairwise comparisons of the three evidence sources along each criterion; e.g., ‘‘In evaluating cost, which procedure is cheaper—analytical review or detailed tests?’’ A measure of the strength of preference was then obtained on a scale of one to nine, where one indicates very little preference and nine represents absolute preference.

Using matrix algebra, a maximum eigenvalue is calculated and a normalized eigenvector is derived from the weights. This eigenvector sums to 1.00 and measures the auditor’s relative trade-offs at each level of the hierarchy on an interval scale. The approach entails a linear, additive, compensatory model.

AHP has been used in many decision settings and has several advantages: ease of understanding, high test/retest reliability, and ability to deal with complex decisions [Saaty, 1980]. A number of recent auditing studies have employed AHP [Lewis et al., 1983; Arrington et al., 1984; Lin et al., 1984; Boritz and Jensen, 1985]. The principal disadvantages of the approach are that AHP: (i) does not consider heuristics; (ii) it is a linear, additive model, while judgment may not be so; and (iii) although it provides adjustments, AHP does not present a normative way to deal with inconsistent responses.⁶

This study focuses on multi-attribute decision making and thus the pairwise comparisons made during the decision process are of greatest concern. The Analytical Hierarchy Process is, therefore, useful here as a vehicle to structure the decision process.

Results

Relative Importance of Evidential Criteria

Table 3 reports the frequency of preferences in pairwise criteria comparisons, suggesting the relative importance of each criterion. A frequency near 50 percent indicates wide disagreement in choice among auditors, while 100 percent reflects unanimity. The results suggest reasonably clear choices as to desired evidential criteria. For subjects in the three criteria group, the order of importance was competency, sufficiency, and (a distant third) cost. These preferences are in agreement with the professional literature (SAS 31); i.e., competence and sufficiency are paramount, with cost a secondary consideration.⁷

⁶ See Jensen [1983, 1984] for a review of the literature on AHP, and a critical analysis. The AHP program used in this study checks for consistency of responses and adjusts the values of the normalized eigenvector for inconsistent weights employing a method developed by Lusk [1976]. The data were further examined for the level of transitivity logic errors. The level of such errors was found to be quite low (9% for the three criteria group and 3% for the five criteria group), suggesting that consistency was not a problem for the auditors in the experiment.

⁷ The low weighting of the cost criterion found also may be because auditors are aware that SAS 31 indicates that cost should be of lower importance, and they are responding in a normative manner, whereas on actual audits cost plays a more dominant role. Future empirical research would be needed to address the validity of this plausible alternative explanation.

Table 3**Pairwise Rankings of Evidential Criteria**

Comparison	Criterion Preferred	% of auditors with Indicated Preference	Strength of Preference* (Mean)
Three Criteria Group:			
Cost vs. Sufficiency	Sufficiency	82	4.9
Cost vs. Competence	Competence	100	6.3
Suff. vs. Competence	Competence	82	5.2
Five Criteria Group:			
Cost vs. Sufficiency	Sufficiency	90	5.7
Cost vs. Bias	Bias	80	5.1
Cost vs. Reliability	Reliability	100	4.5
Cost vs. Relevance	Relevance	90	6.8
Sufficiency vs. Bias	Sufficiency	80	3.4
Suff. vs. Reliability	Sufficiency	60	3.0
Suff. vs. Relevance	Relevance	70	6.4
Bias vs. Reliability	Reliability	60	3.8
Bias vs. Relevance	Relevance	80	4.8
Reliability vs. Relevance	Relevance	70	6.1

* Scale of one (equal importance) to nine (absolute importance).

The order of significance for the five criteria group was: relevance, sufficiency, reliability, bias, and cost. Cost was again seen as least important. Sufficiency, reliability and bias were close choices, all perceived as of about equal, intermediate importance when compared to relevance. Therefore, the two groups displayed consistent responses reflecting relevance/competency as the most important evidential quality, while cost of gathering evidence was considered a secondary factor. Sufficiency fell in the middle.

AHP weightings in Table 4 also reflect this ordering. Competence was the primary criterion (mean weighting .63), with sufficiency (.26) and cost (.11) as secondary factors (three criteria group). For the five criteria group, relevance (.37) and sufficiency (.24) were considered the most important criteria; reliability (.19) and bias (.15) followed in importance. Cost (.05) was judged as a distant minor factor.

Evaluation of Evidential Alternatives

After judging the relative importance of evidential criteria, auditors were asked to evaluate the three procedure areas along each of the criteria studied. The results of these choices are shown in Table 5. For example, when auditors in the three criteria group compared analytical review to observation in terms of cost, 100 percent felt that analytical review was less costly.

However, in general, Table 5 reveals a lack of consensus among participants in applying the criteria to judge the quality of alternative evidence sources. This occurred despite the reasonably strong consensus described

Table 4

**Relative Importance of Evidential Criteria
Measured by AHP Weightings**

Three Criteria Group			Five Criteria Group		
Criterion	Mean*	Standard Deviation	Criterion	Mean*	Standard Deviation
Cost	.11	.09	Cost	.05	.03
Sufficiency	.26	.14	Sufficiency	.24	.16
Competence	.63	.17	Bias	.15	.12
			Reliability	.19	.15
			Relevance	.37	.18

* Represents AHP normalized weightings of relative importance; Scale zero to one.

earlier regarding the relative importance of the various evidential criteria. This lack of consensus is reflected in many of the pairwise comparisons. For example, for the three criteria group, four out of the nine comparisons indicated a lower than 65 percent level of agreement, while the five criteria group had nine out of 15 comparisons below 65 percent. In contrast, in evaluating the criteria, only one of 13 comparisons fell below 65 percent.

The overall ranking of evidential alternatives for the various criteria (Tables 5 and 6) also reflects the difficulties in achieving consensus. While both groups felt analytical review was the least costly to obtain, contradictory results appear in evaluating sufficiency. The three criteria group chose analytical review as superior regarding sufficiency, while subjects in the five criteria group ranked analytical review as of lowest quality along this criterion. The five criteria group could not reach any meaningful consensus in two cases (evaluating cost and relevance for observation and detailed tests) and demonstrated a lack of clear consensus on analyzing all evidence sources as to reliability and relevance. Recall that relevance was considered of greatest importance and yet an evaluation of procedures on this dimension produced great disagreement.

Table 6 indicates the AHP normalized weights for each of the three evidential sources as judged for the various evidential criteria. Both groups considered analytical review to be the least costly procedure to conduct. However, beyond this evaluation, a clear consensus is not present on all other evidential judgments. For example, for the five criteria group, the three forms of evidence are viewed as essentially equal in terms of relevance as measured by the mean weightings. Observation and detailed tests are ranked closely together on all five criteria. Further, the standard deviations of weightings evaluating analytical review are close to, or exceed, the mean on all criteria except cost for both groups, suggesting wide disagreement on the relative merits of this source of evidence.

In summary, despite strong agreement on the relative importance of various evidential criteria, auditors displayed low consensus in applying these

Table 5

Evaluation of Audit Procedures

Criterion	Pairwise Comparison	Choice	% of Subjects	Mean*	S.D	Choice	% of Subjects	Mean*	S.D
Three Criteria Group									
Cost	AR vs. OB	AR**	100%	5.9	3.0	OB	0%	—	—
	AR vs. DT	AR**	91	6.2	2.5	DT	9	5.0	0
	OB vs. DT	OB**	82	3.7	2.0	DT	18	3.0	2.8
Sufficiency	AR vs. OB	AR**	55	2.7	1.5	OB	45	6.2	2.7
	AR vs. DT	AR**	82	3.1	2.1	DT	18	5.0	0
	OB vs. DT	OB**	73	3.9	2.4	DT	27	3.3	0.6
Compe- tence	AR vs. OB	AR	36	5.8	2.5	OB**	64	5.0	2.0
	AR vs. DT	AR	36	3.0	1.8	DT**	64	3.6	2.2
	OB vs. DT	OB**	55	4.0	2.8	DT	45	5.2	2.0
Five Criteria Group									
Cost	AR vs. OB	AR**	90%	5.9	2.0	OB	10%	5.0	0
	AR vs. DT	AR**	100	5.0	2.6	DT	0	—	—
	OB vs. DT	OB	50	4.8	1.5	DT	50	2.8	1.1
Sufficiency	AR vs. OB	AR	20	6.5	2.1	OB**	80	5.5	2.0
	AR vs. DT	AR	20	3.5	0.7	DT**	80	5.4	2.1
	OB vs. DT	OB**	60	4.8	2.2	DT	40	4.3	2.5
Bias	AR vs. OB	AR	10	3.0	0	OB**	90	5.1	1.8
	AR vs. DT	AR	10	3.0	0	DT**	90	4.8	1.1
	OB vs. DT	OB	40	4.3	1.3	DT	60	4.7	2.0
Reliability	AR vs. OB	AR	40	3.8	1.0	OB**	60	5.3	1.6
	AR vs. DT	AR	40	3.8	2.5	DT**	60	5.2	1.8
	OB vs. DT	OB	40	3.3	1.7	DT**	60	3.3	2.1
Relevance	AR vs. OB	AR	40	4.3	2.5	OB**	60	4.0	2.8
	AR vs. DT	AR	40	2.5	1.3	DT**	60	4.3	2.9
	OB vs. DT	OB	50	2.6	1.7	DT	50	3.0	0.7

* Scale one (Equal Importance) to nine (Absolute Importance)

** Majority Preference

Note: Evidential Choices—AR - Analytical Review

OB - Observation

DT - Detailed Tests

criteria to evaluate the illustrative audit programs. Such disagreement is of concern since this result may suggest that, given the same facts, two auditors may plan a widely varying portfolio of audit procedures. This concern is

Table 6**Relative Quality of Evidential Sources**

Criterion	Analytical Review		Observation		Detailed Tests	
	Mean*	S.D.	Mean*	S.D.	Mean*	S.D.
Three Criteria Group						
Cost	.65	.18	.21	.11	.14	.10
Sufficiency	.36	.25	.40	.22	.24	.20
Competence	.28	.25	.37	.26	.35	.24
Five Criteria Group						
Cost	.64	.19	.20	.18	.16	.12
Sufficiency	.20	.25	.43	.23	.37	.25
Bias	.15	.16	.39	.26	.46	.23
Reliability	.27	.23	.36	.24	.37	.18
Relevance	.29	.21	.34	.21	.37	.17

* Represents AHP normalized weightings of relative importance; scale zero to one.

addressed in the next section where the final planned allocations of audit hours are examined.

Allocation of Audit Hours—Final Evidential Judgment

After making criteria evaluations, auditors decided on the allocation of audit hours. Table 7 provides summary data on these judgments. A one-way ANOVA indicated no significant differences ($p < 0.10$) in the allocation of audit hours between the three and five criteria groups. The results here, however, reflect a relatively low level of consensus, consistent with prior studies of evidential planning (e.g., Mock and Turner, [1981]). This low consensus may be the result of the earlier findings, indicating that auditors displayed a high level of disagreement in applying evidential criteria to actually evaluate alternative evidential sources. This finding will be explored further in the final section of the paper.

Discussion

A significant finding in this study was that, although auditors were in close agreement as to the relative importance of various evidential criteria, there was not a strong consensus in applying these criteria to evaluate the merits of alternative procedures or in planning the allocation of audit hours. For example, subjects had widely disparate judgments regarding the effectiveness and sufficiency of analytical review.

The divergent allocation of audit hours among procedures found here is disturbing, since such widely varying audit plans suggest that, in practice, engagements may differ substantially in efficiency and/or effectiveness. Future

Table 7**Allocation of Audit Hours**

Experimental Group	Mean Response (Audit Hours—150 total hours)		
	Analytical Review	Observation	Detailed Tests
Three Criteria	46 hours (19)	38 hours (14)	64 hours (29)
Five Criteria	39 hours (20)	39 hours (19)	72 hours (31)

Note: Numbers in parentheses indicate standard deviation.

research might consider the efficacy of various decision aids to assist auditors in program planning. For example, perhaps continued use of a multi-attribute method such as AHP may result in explicit consideration of evidence trade-offs and lead to greater consensus.

An additional extension to the current study is to focus on various audit objectives. Given the early state of the research in this area, this study looked at audit planning for the inventory account overall. The number of paired comparisons required, and hence subject time, grows very rapidly as additional audit objectives are considered. However, planning in practice is more complex and should explicitly consider needed evidence to address all the relevant audit objectives pertaining to an account. Are auditors evaluating alternative procedures to obtain cost-effective evidence for each objective? Is there a redundancy of tests on some objective(s) while little or no evidence on others?

The auditors in this study assessed evidential criteria and weighted alternative procedures along these criteria with respect to a specific audit case. Thus, corroborating findings, perhaps examining different account areas, evidence alternatives, and/or risk situations are needed to enhance the validity of the results.

The weak consensus in evaluating alternative procedures as to relevance, the criterion considered of greatest importance, is of concern. Perhaps auditors in practice have great difficulty, as suggested here, in appropriately considering this criterion. The evaluation of relevance is further complicated by the fact that audit procedures may address multiple objectives. For example, receivable confirmations provide evidence as to existence, valuation, and cut-off. Thus, future research may address how auditors can and do operationalize this important, yet difficult criterion in practice. For example, some accounting firms have developed program planning materials where various common procedures are ranked (e.g., strong, moderate, weak) with respect to relevance for each key audit objective for the account examined.

These questions illustrate the fact that, as noted earlier, we have little understanding of how evidential planning occurs in practice. Certainly, much more work is needed to begin to evaluate current practices and to identify tools to aid auditors in arriving at such complex, vital judgments. The results

reported here suggest that such tools appear needed for evidential planning and potentially offer significant returns in improving audit efficiency and effectiveness.

References

- American Institute of Certified Public Accountants, *Statement on Auditing Standards No. 56: Analytical Procedures*, New York (1988).
- American Institute of Certified Public Accountants (AICPA), *Codification of Statements on Auditing Standards*, New York (1987).
- Arrington, C. E., W. Hillison, and R.E. Jensen, "An Application of Analytical Hierarchy Process to Model Expert Judgments on Analytical Review Procedures," *Journal of Accounting Research* (Spring 1984), pp. 298-312.
- Boritz, J. E., and R. Jensen, "An Hierarchical Assertion-Oriented Approach to Planning Audit Evidence-Gathering Process," *University of Southern California Audit Judgment Symposium* (1985).
- Cushing, B., and Loebbecke, J., *Comparison of Audit Methodologies of Large Accounting Firms* (American Accounting Association, 1986).
- Hylas, R., and R. Ashton, "Audit Detection of Financial Statement Errors," *The Accounting Review* (October 1982), pp. 751-765.
- Jensen, R. E., "Aggregation (Composition) Schema for Eigenvector Scaling of Criteria Priorities in Hierarchical Structures," *Multivariate Behavioral Research* (1983), pp. 63-84.
- , "An Alternative Scaling Method for Priorities in Hierarchical Structures," *Journal of Mathematical Psychology* (September 1984), pp. 317-332.
- Joyce, E., and G. Biddle, "Anchoring and Adjustment in Probabilistic Inference in Auditing," *Journal of Accounting Research* (Spring 1981), pp. 120-145.
- Lewis, M., T. Lin, T. Mock, and A. Wright, "Evaluation of Audit Evidence in the Planning Process," *Proceedings: Symposium on Auditing Research V* (University of Illinois, 1983).
- Lin, W., "Multiple Criteria Decision Making in Auditing: The State of the Art," *The Auditor's Report* (Winter 1987).
- Lin, W., T. Mock, and A. Wright, "The Use of the Analytic Hierarchy Process as an Aid in Planning the Nature and Extent of Audit Procedures," *Auditing: A Journal of Practice and Theory* (Fall 1984), pp.89-99.
- Lusk, E., "Evaluation of Prospective Audit Clients: An Eigenvalue Priority Assignment Model," *Proceedings of Audit Symposium II* (University of Illinois, 1976).
- Mock, T., and J. Turner, *Internal Accounting Control Evaluation and Auditor Judgment*, Auditing Research Monograph 3 (AICPA, 1981).
- Saaty, T., "A Scaling Method for Priorities in Hierarchical Structures," *Journal of Mathematical Psychology* (September 1984), pp. 317-332.
- , *The Analytic Hierarchy Process* (New York: McGraw-Hill, 1980).
- , "Modeling Unstructured Decision Problems—The Theory of Analytical Hierarchies," *Mathematics and Computers in Simulation* (September 1978), pp. 147-158.
- Tabor, R. H., and J. T. Willis, "Empirical Evidence on the Changing Role of Analytical Review Procedures," *Auditing: A Journal of Practice and Theory* (Spring 1985), pp. 93-108.
- Wright, A., "The Impact of Prior Working Papers on Auditor Evidential Planning Judgments," *Accounting Organizations and Society* Vol. 13, No. 6 (1988) pp. 595-606.