1-1-1980

Investigation of a measurement based approach to the evaluation of audit evidence

Theodore J. Mock

Arnold Wright

4

An Investigation of a Measurement Based Approach to the Evaluation of Audit Evidence

Theodore J. Mock*
University of Southern California

Arnold Wright*
University of Nevada, Reno

Several definitions of the auditing function have been proposed, e.g., Mautz and Sharaf (1961, p. 15); American Accounting Association (1972). At the root of these definitions is the notion that the accumulation of *evidence* is the cornerstone of the auditing, attestation process. There is also widespread recognition that the "quality" of the many types of audit evidence varies considerably (AAA, 1971; Toba, 1975; Robertson, 1976; AICPA, 1973). In evaluating the propriety of a given assertion, an auditor must weigh the quality as well as the quantity of evidence gathered. Some forms of evidence are compelling such as observation of marketable securities while others are merely suggestive. Thus, evidence evaluation is a complex, vital decision. How should an auditor, then, consider the many factors involved and arrive at an appropriate judgment? What guidelines or tools are available to aid in this difficult task? Although there have been numerous attempts to provide theoretical frameworks which examine the nature of evidence,[1] the concepts proposed are generally vague and not operational. Most importantly, an overall approach for evaluating evidence has not been presented.

The purpose of this paper is to investigate the usefulness of a measurement based approach as an integrative, operational process to evaluate audit evidence. Towards this end the paper addresses four main topics. The first section considers the nature of audit evidence and its role in the audit process. This discussion is followed by a review of the literature. Several evidence evaluation frameworks are identified and analyzed. The third section introduces the measurement based approach and illustrates its use in audit evidence evaluation. The final section discusses the implications of this approach for practice and explores avenues for future research.

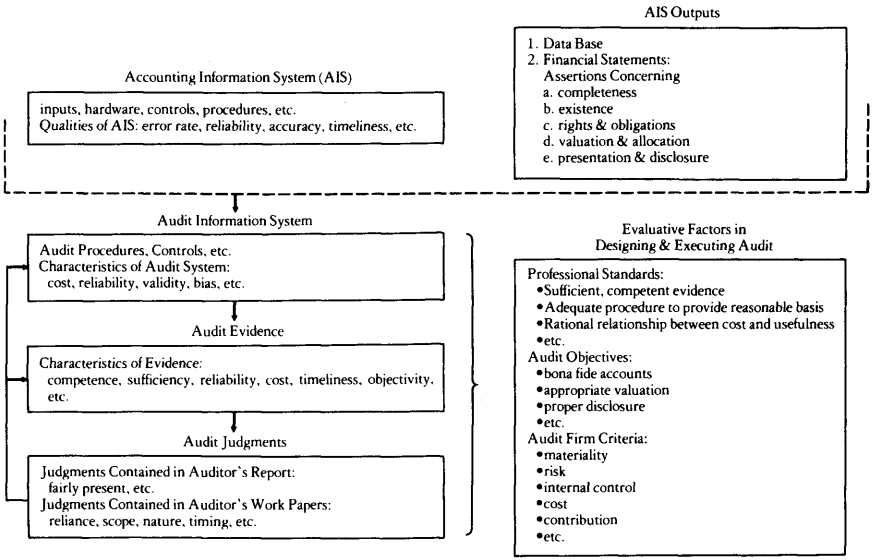
The Nature of Audit Evidence and its Role in the Auditing Process

Figure 1 presents a model of the role of evidence in the audit process. The model contains three major elements:

* The authors wish to gratefully acknowledge the comments and ideas expressed by the participants at the University of Southern California Accounting Research Forum.

(1) the Accounting Information System (AIS);
(2) the Auditing Information System (ADIS); and
(3) the evaluative factors in planning audit tests and analyzing the evidence gathered.

Figure 1
ROLE OF EVIDENCE IN THE AUDIT PROCESS



The primary output of the AIS is the financial statements. Underlying these statements are several broad assertions represented by management:

completeness;
existence;
rights and obligations;
valuation or allocation;
presentation and disclosure (AICPA Auditing Standards Board, 1979).

The vital role of the audit process is to independently test whether these assertions appear warranted based on the evidence accumulated and, thus, express an overall opinion as to the fairness of the financial statements.

The ADIS attempts to gather evidence to provide the basis for various audit judgments. Toba (1975, p. 9) emphasizes this significant function of evidence as "the basis on which one ought to fashion one's belief or draw some conclusion with respect to the proposition established." The value of an audit lies in the "warranted assertions" (AAA, 1972) made by the auditor. Warranted assertions are those believed to be appropriate based on the evidence examined and the circumstances. The audit opinion, thus, adds credibility to the financial statements because of the declared belief by the professional auditor that management assertions are appropriate.

Figure 1 identifies three key evaluative factors that appear to be significant considerations (criteria) in weighing various types of audit evidence to support a given assertion:

(1) professional standards, e.g., SAS#1, Section 330 (1973);
(2) audit objectives (AICPA Auditing Standards Board, 1979); and
(3) audit firm criteria.

The first two factors are self explanatory. Audit firm criteria are matters relating specifically to the given assertion under investigation in its client setting. Some of the important firm criteria noted are (SAS#1, Section 330, 1973 and Anderson et al., 1970):

- risk of assertion;
- evidence cost;
- materiality; and
- internal control.

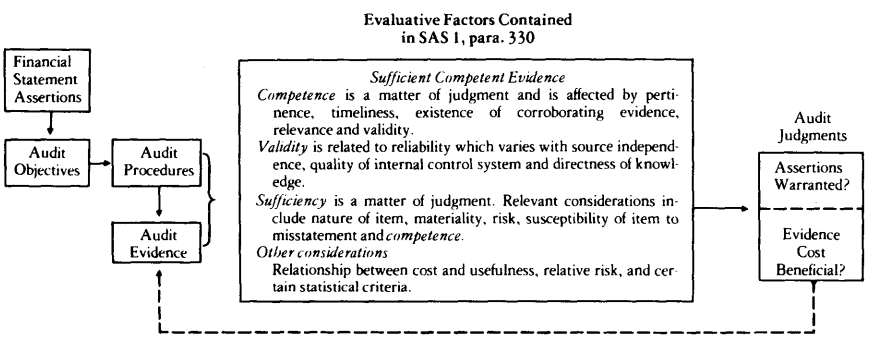
The model of the role of evidence in the audit process, as discussed in the preceding paragraphs, highlights the function of evidence as the means of affording the auditor the basis to state beliefs regarding financial statement assertions. Therefore, the auditor must be able to evaluate the "adequacy" of various evidential matter necessary to support a warranted assertion. The evaluation of evidence is a very complex decision involving numerous variables (Peat, Marwick, Mitchell and Co., 1976, p. 20-24; AAA, 1972, pp. 34-50). Thus, guidance and aids in this area appear vitally needed. Additionally, an overall systematic approach that encompasses the major factors to be considered would be highly desirable and promising.

The next section examines some of the existing evidence evaluation approaches found in the professional and research literature. An overall summary and critique of these approaches is then presented. Finally, the need for an integrated process is discussed.

Existing Approaches to Evaluating Audit Evidence

Professional Standards. SAS No. 1, Section 330 is the primary normative framework available in the professional literature on evidence evaluation. Figure 2 outlines the criteria advanced by this framework.

Figure 2
AUDIT EVIDENCE AS INDICATED IN SAS 1, 330



The third standard of field work states:

Sufficient competent evidential matter is to be obtained through inspec-

tion, observation, inquiries, and confirmations to afford a reasonable basis for an opinion regarding the financial statements under examination (para. .01). [emphasis added]

The decision as to what constitutes "sufficient" evidence is considered to be a "matter of professional judgment," although four factors are mentioned as significant to the decision (para .09):

(1) the nature of the item examined;
(2) materiality;
(3) the risk involved, which is dependent upon the adequacy of internal control and the susceptibility of an item to misstatement; and
(4) the competence of the evidence available.

"Competent" evidence is defined as that which is both valid and pertinent. A final criterion is presented in paragraph .10: "In the great majority of cases, the auditor finds it necessary to rely on evidence that is persuasive rather than convincing." Thus, SAS No. 1, Section 330, entitled "Evidential Matter," provides three general criteria to guide practitioners in assessing the quality of audit evidence:

(1) competency;
(2) validity; and
(3) sufficiency.

Additionally, a number of other considerations such as cost effectiveness and risk are noted in SAS No. 1.

The criteria provided by SAS No. 1, although identifying significant concepts and issues, appear to have several shortcomings:

(1) the concepts noted are vague and not operational;
(2) measurement of the criteria is not addressed; e.g., how does one measure "validity?"
(3) a scientific, systematic approach is not presented. Thus, the reliability and validity of evidence gathered employing this standard is open to question; and
(4) on occasion, terminology and concepts appear to be used in an inconsistent, imprecise manner which may result in confusion. For example, validity is said to be directly related to reliability (para .08). Such concepts have distinct, separate scientific meaning, as will be noted later.

A later exposure draft (AICPA Auditing Standards Board, 1979) outlines the nature of major audit assertions and related objectives and substantive tests but does not address the issue of evidence evaluation criteria.

Research findings. Mautz and Sharaf (1961) propose essentially the same evaluation criteria as Section 330:

> In the degree of influence it exerts on the mind of the auditor, audit evidence varies from compelling through persuasive to inconclusive. . . . Audit evidence must be reviewed critically with respect to its *validity* and *pertinence* before it is permitted to influence the mind of the auditor with respect to an assertion at issue (p. 110).

Robertson (1976) similarly defines competent audit evidence as: relevant, objective, and free from bias.

Toba (1975) provides valuable insights as to the nature and evaluation of evidence. The concept of the *weight of evidence* is discussed. Evidence is divided into two categories: confirming evidence and supporting evidence. Confirming evidence establishes the validity of a proposition. Supporting evidence merely makes a proposition more tenable.

> Statement (evidence) q may be said to have *confirming* power for statement p if statement q is well established and renders p more probable than not − p (expressed as $\bar{p}$). In symbols, for q to have confirming power:
>
> $P(p \mid q) > P(\bar{p} \mid q)$
>
> Statement (evidence) q may be said to have *supporting* power for statement p if the probability $P(p \mid q)$ is greater than the prior probability of statement p. In symbols, for q to have supporting power:
>
> $P(p \mid q) > P(p)$ (Toba, 1975, p. 9)
>
> Toba also classifies propositions as:
>
> (1) elementary, stating facts or events;
> (2) general, describing a value judgment or generality;
> (3) immediate, self evident; and
> (4) demonstrable, subject to proof.

A general proposition cannot be directly proven, as can an elementary statement, but must be rephrased into elementary propositions while maintaining equivalence between the general proposition and the surrogate elementary statements. Demonstrable propositions can be proven to some degree of confidence (probability), while immediate statements do not require proof. Thus, in Toba's framework the evaluation of internal control is not evidential matter, but an elementary proposition to be proven, providing supporting evidence as to the fairness of the financial statements. Finally, Toba notes that auditing is essentially a heuristic, demonstrative process of persuasion rather than an investigative, learning approach.

While Toba does much to develop a general theory of evidence, the concepts provided are broad and not operational. For example, evidence is said to have confirming power if it is "well established and renders p more probable than not-p." However, neither a definition nor criteria are proposed to determine what constitutes "well established" evidence in a given situation. Further, a basis to analyze the strength of various forms of evidence is not offered.

Kissinger (1977) addresses a number of "deficiencies and oversights" in Toba's paper and extends the framework. He especially disagrees with Toba's conclusions as to the conditions necessary for fair presentation. Kissinger proposes twelve general propositions considered in an audit and symbolically presents the necessary and sufficient conditions for the various types of opinions.

Sneed (1978) examines similarities in evidence accumulation problems and objectives for historians and auditors.[2] Several evidence evaluation criteria are advanced:

(1) authenticity;
(2) credibility;
(3) reliability; and
(4) relevance.

Schandl (1978) proposes five "Principles of Evidence":

(1) availablity (sufficient evidence needed);
(2) independence;
(3) directness (reliability, distance source is removed from the assertion tested);
(4) confirmation (need corroborating evidence); and
(5) bias.

These principles should be jointly considered in examining the strength of various evidence sources.

A Statement of Basic Auditing Concepts (ASOBAC) (AAA, 1972) deals extensively with the nature of audit assertions and the investigative process. However, the issue of evidence evaluation is only incidentally addressed. ASOBAC merely notes that evidence must be competent and sufficient to provide adequate belief that an assertion is warranted. A prerequisite to competency is "intersubjectivity" (objectivity).

There has been extensive research on the nature of the audit decision process and the role of evidence in this process (Kinney, 1975; Scott, 1973; Tracy, 1969; Elliot and Rogers, 1972). These works discuss the usefulness of various decision models to auditing such as the Bayesian method. However, an integrated approach to assess the strength of a particular type of evidence is not proposed. Instead, these papers provide a general framework to organize and direct audit efforts.

Summary and Critique of Existing Evidence Evaluation Approaches

Generally accepted auditing standards (SAS No. 1, Section 330) emphasize the importance of obtaining "sufficient competent" evidential matter to provide a reasonable basis for the auditor to express an opinion. However, only vague criteria are provided to assess the adequacy of evidence gathered, i.e., validity; competency and sufficiency. Mautz and Sharaf (1961), Schandl (1978), Robertson (1976), and others attempt to provide criteria to evaluate evidence but do little to clarify or provide more concrete guidelines.

Toba (1975) and Kissinger (1977) present a general theory of evidence and outline the heuristic decision process employed during an audit. These works provide a theoretical framework to address the purposes and evaluation of evidence. However, the model proposed is not operational and does not outline an approach to evaluate the strengths of various forms of evidence.

The existing approaches, thus, suffer from two major deficiencies:

(1) they provide only heuristic intuitive concepts, i.e., they are not operational nor subject to empirical testing; and
(2) they do not provide an integrated, scientific approach to the evaluation of evidence.

The first deficiency has resulted in several heuristic beliefs in auditing to access the strength of evidence. While many of these heuristic rules are undoubtedly useful and valuable, some of these "rules of thumb" may result in serious errors in audit judgments. Few of these beliefs have been empirically verified.

An example of such a heuristic rule is contained in SAS No. 1:

> When evidential matter can be obtained from independent sources outside an enterprise, it provides greater assurance of reliability than that secured solely within the enterprise (para .08).

This belief has, for example, resulted in heavy reliance by auditors on accounts receivable confirmations. Recent empirical results suggest that, in fact, such evidence has a high error rate (Sorkin and Meuwissen, 1978).

Need for an Integrated, Scientific Approach

Many individuals would argue that the evaluation of audit evidence is by its very nature a matter of professional judgment and, thus, cannot be subject to any scientific approach. The same argument existed for many years about the entire auditing process. Undoubtedly, auditing should not be viewed as a purely precise scientific discipline. However, Mautz and Sharaf (1961) have demonstrated the applicability and advantages of a scientific approach to auditing. Since that time several scientific notions have been applied successfully such as statistical sampling (noted in SAS No. 1, Section 330) and the explicit recognition of audit assertions or hypotheses (AICPA Auditing Standards Board, 1979).

The scientific approach tends to add rigor, precision, and greater reliability to an endeavor, thus, improving quality control. As Anderson et al. (1970) state:

> The auditor must be able to say that he has *enough* evidence to sustain or refute one of the evidential propositions. Thus, *quantification* of the evidence needed, however crude the measurement, is an essential aspect of the discipline (p. 527).

Two additional factors that support the need for a scientific approach to evidence evaluation are:

(1) difficulties generally encountered by decision makers in arriving at complex judgments; and
(2) the threat of government intervention into the auditing profession and the extensive legal exposure facing CPAs.

Research findings (Libby and Lewis, 1977; Slovic and Lichtenstein, 1977) on human information processing indicate that individuals often arrive at judgments that seriously deviate from normative models in complex decision settings, e.g., demonstrate poor accuracy, low consensus, poor consistency. For instance, several studies of auditor's judgment have found widely varying recommendations even when auditors were presented with identical problem situations (Ashton, 1974; Joyce, 1978; Weber, 1978; Wright, 1979). This type of finding also has been observed with respect to auditor's evaluation of the nature of audit evidence. In a series of five related field experiments, Mock and Turner (1978, 1980) found that experienced auditors frequently differed as to whether three audit procedures were compliance, substantive or dual purpose tests. Such findings strongly suggest that, if quality control is to be maintained at high levels, professional judgment alone cannot be relied upon. Guidance, training, a rigorous approach, and/or other tools are needed.

The Metcalf (1976) and Moss (1977) reports allege that a number of bankruptcies and frauds have led to loss of confidence in the auditors' opinion. Federal government intervention was recommended. The prospect of interven-

tion, coupled with the legal exposure accompanying certified audits, attests to the vital need for the profession to maintain and improve quality control and to provide proper documented support for audit judgments. A systematic, scientific approach appears promising in addressing these concerns for the assessment of audit evidence.

In response to these needs, the measurement based approach is now introduced. An illustration of the use of this approach in the area of inventory is then presented.

A Measurement Based Approach to the Evaluation of Audit Evidence

The preceding review of the audit evidence evaluation problem has indicated that this is a complex, multiple-factor problem. Professional standards suggest that the decision as to what constitutes "adequate" audit evidence should be based upon factors such as competency and sufficiency and is primarily a matter of "professional judgment."

From the perspective of information economics and measurement theory, the issue of developing sufficient competent evidence might be viewed as the issue of designing and implementing an efficient audit information system (ADIS) which provides audit evidence of acceptable quality. Our discussion begins with a consideration of those factors which affect the factual quality of audit evidence (a viewpoint based upon measurement theory concepts). Then the question of the cost-effectiveness of the ADIS is considered.

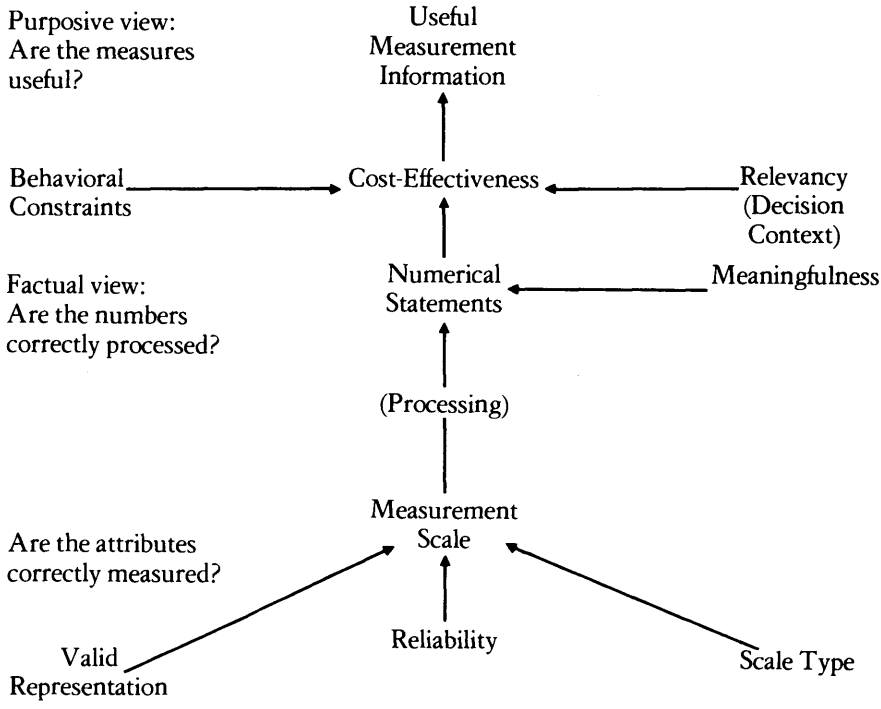
The factual quality of evidence. Research directed at the evaluation of the data, evidence, or measures provided by an information, audit, or measurement system[3] may focus on two interrelated questions. The overriding question is whether the evidence is useful. This question is labeled the *purposive view* in Figure 3 which depicts the basic factors underlying measurement system evaluation. The purposive view will be discussed in detail in the following section.

A second view identified in Figure 3 asks two related questions concerning the factual quality of measures. Are the attributes correctly measured and are the assigned numbers correctly processed? An attribute may be thought of as a characteristic of an object such as accounts receivable or an event such as a sale. Relevant attributes in auditing might include the *reliability* of a system of accounting controls over payroll, the *bona fides* of a receivable or the *net realizable value* of obsolete inventory.

As indicated in Figure 3, the factual quality of measures of such attributes depends on three criteria: (1) reliability, (2) valid representation or validity, and (3) scale type. In discussing each of these criteria, our concern will be with the advantage, if any, of these criteria as compared to existing audit evidence criteria as discussed in professional standards and the literature in general. Some of the potential advantages of the measurement based approach include improved guidance, a more systematic approach, and more operational and rigorous definitions.

Reliability. The notion of the reliability of a measurement procedure emphasizes the errors inherent in that process. Measurement error may be the result of a number of factors including calibration errors, observer errors, and sampling errors. Ackoff (1962, p. 208) reports upon experiments which showed important differences in observer error rates among auditors who were testing credit-

Figure 3
A BASIC FRAMEWORK FOR MEASUREMENT SYSTEM EVALUATION



Adapted from: Mock and Grove (1979).

compensation forms. Sampling error (sampling risk) has been an important audit consideration for many years and results when the audit procedures are not applied to the entire population of interest. Sorkin and Meuwissen (1978) have researched alternative audit confirmation procedures and have reported significant differences in terms of their reliability.

An important feature concerning reliability as a potential criterion for audit evidence evaluation is the previous research which has been done on defining and operationalizing the concept. Definitions of reliability focus on the stability and accuracy of a measurement procedure. Stability implies that a measurement procedure applied to the same object or event should arrive at identical (or at least very similar) numerical assignments (assuming the attributes have not changed in the period between measurements). Clearly the reliability of audit tests could be measured (estimated) in this manner.

Reliability has also been conceptualized in terms of accuracy defined as a function of the difference between a measure (Y) of an attribute and its true value (X). Thus the error (e) in a test is mathematically defined as $e = X - Y$ and reliability is operationalized as the average error ($\bar{e} = \Sigma_i(X_i - Y_i)/N$), the mean squared error

$\sigma^2 = \Sigma_i(X_i - \bar{e})^2/N)$ or some other transformation of e. Further details of this approach can be found in Mock (1976) and Mock and Grove (1979). Clearly the reliability aspect of measurement procedures offers several alternative approaches that would assist the auditor in obtaining operational measures of the reliability of audit procedures and evidence.

Validity. The issue of validity or whether the numerical assignments are a valid representation of the attributes being measured is a more difficult concept to define and operationalize. Rigorous definitions have been developed in formal measurement theory which emphasize that the assigned numerals should be related in the same manner as the measured objects or events (i.e., the measures should be a homomorphism of the measured phenomenon). The auditor has the same objective in mind when deciding that the reliability attribute of one system of controls is .95 and another is .85. Hopefully, the first system is *in fact* more reliable than the second. A number of operational approaches to empirically testing validity in this sense are reported in Mock (1976).

In the behavioral sciences a valid measure is one that measures the attributes it is designed to reflect, e.g., an IQ measure actually reflects intelligence. Many audit attributes may be as difficult to validly measure as intelligence: nonsampling risk, audit risk, audit materiality, and reliability of an accounting control. Several approaches to operationalizing the validity of a measurement system have been developed in the behavioral sciences including construct and criterion validation (see Kerlinger, 1973). Such approaches would seem to be applicable to the evaluation of the validity of audit evidence.

Scale type and meaningfulness. The third evaluative factor identified in Figure 3 is the scale type of the measurement system being evaluated. Possible scale types include nominal, ordinal, interval and ratio scales. Scale type is important in that it relates directly to the *meaningfulness* of numerical statements which are based upon processed (aggregated) measures. The meaningfulness of a numerical statement, inference, or assertion may be determined analytically (see Ackoff, 1962 or Mock 1976). The analogous issue in auditing would be whether audit assertions were or were not meaningful. This would depend partly upon the underlying scale type of the audit procedures. Potential areas where the meaningfulness of audit assertions and aggregations may be questioned include (1) reliability scores and compliance rates which are aggregated into reliance factors, (2) ordinal reliance judgments and internal control questionnaire enumerations which are factored into beta risk and ultimately sample size decisions and (3) the aggregation of multiple, related tests such as negative and positive confirmations into an audit judgment.

As is apparent in the preceding, consideration of the factual qualities of measurement systems may have some application in evaluating audit evidence and procedures. Purposive aspects are also important.

The purposive view. The question in Figure 3 associated with the "purposive view" asks whether the obtained measures, related measurement procedures, and numerical statements are useful. Although factual level criteria are depicted as impacting usefulness, they are not sufficient criteria. The designer of the measurement, information, or audit system will also need to consider relevancy, cost, and certain behavioral constraints.

In a management information system, relevancy is dictated by a decision problem or decision context. In auditing, relevancy seems to depend on the particular

audit assertions which are being evaluated. A recent exposure draft concerned with evidential matter (AICPA, 1979), translates audit assertions into audit objectives and the auditor is expected to implement the necessary audit procedures to achieve these objectives. The exposure draft presents an illustration but does little in the way of providing guidance in evaluating alternative audit procedures except to suggest that the evidence be "adequate to achieve the audit objectives" and "provide a reasonable basis concerning the validity of individual assertions" (paraphrased from AICPA, 1979).

Cost and behavioral constraints may also affect the usefulness of alternative measurement and auditing procedures. Cost is clearly an important factor in comparing packages of audit procedures in light of the reliance that may be placed on a system of internal accounting controls (Turner and Mock, 1980). Behavioral constraints may include tendencies of auditors to anchor on last year's audit program, halo effects and idiosyncratic information search heuristics (Mock and Turner, 1978).

Although both the identified considerations in the factual and purposive views could be discussed in much greater depth, we now turn to the question of application and then present an illustration. References are included in the bibliography for those interested in further details.

Application of Evaluative Factors Contained in the Measurement Based Approach to Auditing

The evaluation of audit evidence and audit procedures involves two issues that are identified in Figure 2: Are management's assertions as contained in the financial statements warranted? Does the benefit of the evidence collected justify its cost? To consider the applicability of the evaluative factors identified in the preceding section, Figure 4 replaces the criteria contained in Figure 2 with the measurement based criteria. The questions contained in the figure are identical except that "Are the numbers correctly processed?" is replaced by "Are the assertions correctly drawn?" As discussed earlier, this question may be evaluated in terms of the meaningfulness of the assertions given the results of an audit test of a certain reliability, validity and scale type. As before, the issue of the correctness of the attribute evidence is directly a function of reliability and validity. Note also that the figure replaces measurement scale with audit test (procedure).

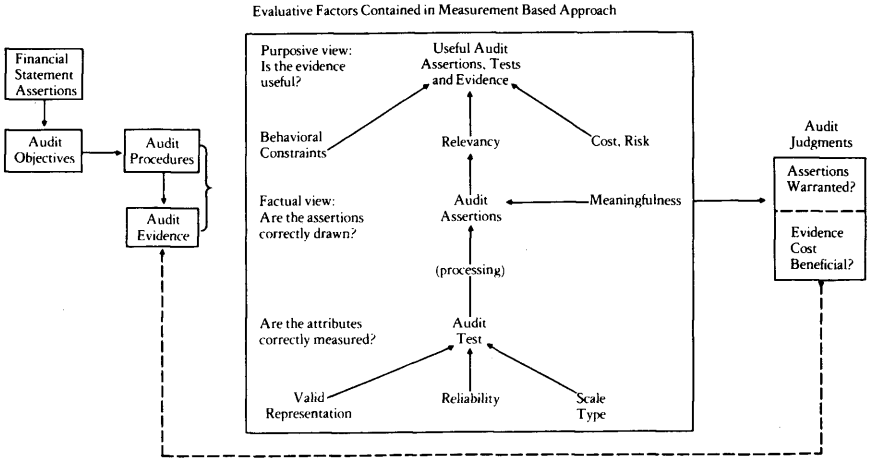
Given that the factual qualities of an audit test or a set of audit procedures is known or knowable (recall that each of these underlying factors has been operationalized), the analysis shifts to the usefulness of the evidence. As before, important evaluative factors include cost, relevance, and behavioral constraints. In addition, audit risk is added to the figure. As will be discussed in the illustration, risk is a necessary addition to the evaluation which is somewhat unique to the audit information system.

Some previous research has been completed on the procedures one might apply in using a measurement based approach in evaluating information systems in general (Mock and Grove, 1979). In the case of audit evidence evaluation, the following reformulated steps seem appropriate.

Step 1: Identify audit assertions to be evaluated.
Step 2: Identify the financial statement attributes which need to be investigated.

Figure 4

AUDIT EVIDENCE EVALUATION AS INDICATED IN MEASUREMENT BASED APPROACH



Step 3: Identify alternative audit tests (programs) which may provide evidence on the appropriate attributes. Analyze each alternative in terms of the validity and reliability of the evidence and the meaningfulness of possible audit assertions.

Step 4: Analyze the usefulness of each alternative in terms of relevancy, cost, audit risk, and behavioral constraints.

Illustration of the Use of the Measurement Based Approach: Audit of Inventory

To illustrate the application of the measurement based approach, the effectiveness of the procedure of inventory observation is now examined. The analysis parallels the four step process outlined earlier.

Step 1: Identify audit assertions to be evaluated.

The general assertion is that inventory is fairly presented on the financial statements. To test this assertion, the auditor must partition the general proposition into several elementary assertions such as (AICPA Auditing Standards Board, 1979):

completeness
existence
rights and obligations
valuation
allocation
presentation and disclosure

Step 2: Identify the financial statement attributes which need to be investigated.

The attribute of inventory *existence* will be the focus of this illustration. Of course, during the course of the audit all of the above attributes would be addressed through the various audit tests.

Step 3: Identify alternate audit tests (programs) which may provide evidence on the appropriate attributes.
Analyze each alternative.

For brevity, only inventory observation is examined as a source of evidence to test the assertion of existence. In fact, audit tests are often interrelated and may provide corroborating evidence. For example, analytical review may also be relied upon as evidence of inventory existence. Audit assertions may also be related. For instance, the attributes of *existence* and *valuation* of a jeweler's inventory are affected by the grade and quality of diamonds.

For simplification, such interrelationships are not dealt with in this illustration. However, such complexities do not appear to impair the value of the measurement based approach. This approach can alert the auditor to apparent problems in a particular source of evidence. Greater reliance can then be shifted to other evidence sources or strategies can be taken to reduce the problems (errors) suggested by the measurement based approach. Three alternate forms of evidence that may be evaluated along with observation are:

(1) outside expert observation;
(2) greater reliance in and testing of the purchasing system; and/or
(3) analytical review with limited observation.

Unless the choice is obvious, each of the alternative audit tests should be analyzed employing the measurement based approach to select the most advantageous procedure.

At the factual level the following measurement characteristics of inventory observation could be examined:

(1) *validity:* can an auditor appropriately identify an inventory item when viewed?
method of testing: a series of field experiments to see if auditors can spot deliberate misrepresentations of inventory.
(2) *reliability*[4]: to what extent are there errors in auditor test counts? How does the count plan affect reliability?
method of testing:
(a) field experiments where various auditors take controlled test counts;
(b) field experiments where the error rate is determined under alternate count plans.
method of measurement: reliability could be operationalized as the mean error or mean squared error. Analysis of variance could also be utilized in a multi-variate approach.
(3) *scale type:* no apparent problems

Step 4: Analyze the usefulness of each alternative in terms of relevancy, cost, audit risk, and behavioral constraints.

Inventory observation appears to be a highly relevant source of evidence since it directly tests the assertion of existence. However, observation is a costly procedure and is of unknown effectiveness. There are several behavioral constraints to consider such as limited auditor experience and environmental biases (Wright, 1979). Additionally, inventory is usually a high risk area in that it is often a material amount and, if misstated, could result in the financial statements being misleading. The ultimate risk of testing the assertion of existence is significantly dependent on the internal control system. Ultimate risk may be stated as (Robertson, 1976, p. 368):

$$UR = \beta \times (CUPL) \times (1\text{-}SP)$$

where: UR = ultimate risk
β = probability substantive tests will fail to detect a material error.
CUPL = probability the internal control system has allowed and failed to detect a material error.
SP = probability supplemental procedures will uncover a material error.

At the purposive level the overall cost-effectiveness of a source of evidence must be determined by weighing all of these variables (cost, relevance, risk, behavioral constraints). If two procedures appear equally effective, the less costly one should be employed. The implications of the measurement based approach, as illustrated by the evaluation of inventory observation, are addressed in the final section. Avenues for future research and conclusions are then presented.

Implications and Avenues for Future Research

From the illustration several implications of the measurement based approach are brought forth. First of all, the systematic nature of the process is displayed. The approach requires specific statements as to assertions, measurement criteria, methods of testing effectiveness, and overall considerations in arriving at a conclusion. Secondly, the effectiveness of a procedure is subject to empirical testing.

The analysis indicates several hypothesized tradeoffs in relying on inventory observations:

(1) unknown validity;
(2) unknown reliability;
(3) significant costs;
(4) high relevance; and
(5) behavioral considerations.

Observation is traditionally considered among the most competent forms of evidence (Windel, 1961, SAS #1, Section 330, 1973). The measurement based approach analysis suggests that there may be important validity problems in relying on such evidence. Specifically, do the auditors assigned to observe the count and take test counts have the expertise to identify various inventory items? For example, can an auditor correctly differentiate between types of electronic printed circuit boards or a transistor and a capacitor? Often the auditors performing such tasks have limited technical knowledge in the client's industry. Obviously, the ability of the auditor to identify the inventory and, thus, the validity of this evidence source varies among industries. The illustration indicates that the validity of the procedure should be empirically tested when in question. An important general hypothesis, thus, emerges from this analysis for future auditing research:

- Inventory observation may produce evidence of limited validity.

The heuristic evidence evaluation approaches discussed earlier tend to lead auditors to accept inventory observation as a highly compelling form of evidence. A similar possible misperception regarding the apparent strength of accounts receivable confirmations was alluded to earlier. It is of interest to note that both inventory observation and accounts receivable confirmations are required under

generally accepted auditing standards. Such requirements may have resulted in unwarranted perceptions of the quality of these forms of evidence.

Another implication of the approach is that alternate evidence sources should also be evaluated when a decision is not evident. Alternate sources to inventory observation were noted earlier.

Several additional avenues for future research in this area appear promising. An obvious extension would entail attempting to apply the approach to evaluate alternate evidence sources in practice. Much work needs to be done to operationally define and obtain agreement on the measurement of the criteria presented in this paper. The problem of the interrelationship of various audit evidence is not addressed in this study. Another significant area of research would be to use the approach to identify overlapping, duplicate forms of evidence that may be unnecessary and, thus, lead to inefficiencies in the audit.

The purpose of this paper is to introduce the measurement based approach as a systematic means of evaluating audit evidence. Grounded in measurement theory and the scientific method this approach appears to offer greater rigor and precision than traditional heuristic evidence evaluation procedures. Further research in applying and refining the approach is, thus, greatly encouraged and needed.

Footnotes

1. For example: see Mautz and Sharaf (1961); Mautz, (1964); Arens, (1970); AAA ASOBAC (1971); Kissinger, (1974); Toba, (1975).

2. Evidence evaluation is also a significant concern in many other disciplines such as law and history. See Mautz and Sharaf (1961, p. 76) for a comparison of evidence approaches in various fields.

3. Although some significant differences exist between information systems in general and audit or measurement systems in particular, this paper emphasizes their similarities. Each system is concerned with the development of data which meets both factual and purposive objectives. Rather than mix terminology, our discussion relies primarily upon measurement concepts and definitions. The basic concepts are developed in Mock (1976) and Mock and Grove (1979).

4. In the statistical sampling literature (Vanasse, 1976; Robertson, 1976) the term "reliability" is used to indicate the confidence level provided of making a correct decision given sample results, i.e., the representativeness of the sample or $1 - \alpha$. This concept of reliability actually provides a measure of decision *risk* and does not conform to the precise meaning adopted in measurement theory, as defined earlier on page 18 (Mock and Grove (1979). The approach presented in this paper corresponds closely with the constructs of measurement theory.

References

Ackoff, R. L., *Scientific Method* (John Wiley, 1962).

AICPA Auditing Standards Board, *Financial Statement Assertions, Related Audit Objectives, and the Design of Substantive Tests.* Exposure Draft (American Institute of Certified Public Accountants, December 31, 1979).

American Accounting Association Committee on Basic Auditing Concepts, "Report of the Committee on Basic Auditing Concepts," *Supplement to the Accounting Review,* 1972.

American Institute of Certified Public Accountants, *Statement on Auditing Standards No. 1* (AICPA, 1973).

Anderson, H. M., J. Giese, and J. Booker, "Some Propositions About Auditing," *Accounting Review* (July 1970).

Ashton, R., "An Experimental Study of Internal Control Judgments," *Journal of Accounting Research* (Spring 1974).

Elliott, R. and J. Rogers, "Relating Statistical Sampling to Audit Objectives," *The Journal of Accountancy* (July 1972), pp. 46-55.

Ijiri, Y., *Theory of Accounting Measurement,* AAA Studies in Accounting Research No. 10 (American Accounting Association, 1975).

Joyce, E., "Expert Judgment in Audit Program Planning," *Studies on Human Information Processing,* Supplement to *Journal of Accounting Research* (1976).
Kerlinger, F., *Foundations of Behavioral Research* (Holt, Rinehart and Winston, 1973).
Kinney, W., "Decision Theory Aspects of Internal Control System Design/Compliance and Substantive Tests," *The Journal of Accounting Research Supplement* (1975), pp. 14-29.
Kissinger, J., "A General Theory of Evidence as the Conceptual Foundation in Auditing Theory: Some Comments and Extensions," *The Accounting Review* (April 1977).
Libby, R. and B. Lewis, "Human Information Processing Research in Accounting: The State of the Art," *Accounting, Organizations and Society* (1977).
Mautz, R., *Fundamentals of Auditing* (John Wiley & Sons, 1964).
Mautz, R. and H. Sharaf, *The Philosophy of Auditing* (American Accounting Association, 1961).
Metcalf Report . . . See United States Congress, Senate.
Mock, T., *Measurement and Accounting Information Criteria,* Studies in Accounting Research No. 13 (American Accounting Association, 1976).
Mock, T. and H. Grove, *Measurement, Accounting, and Organizational Information* (Wiley, 1979).
Mock, T. and J. Turner, *Internal Accounting Control and Auditor Judgment,* Audit Research Monograph No. 3 (American Institute of Certified Public Accountants, 1980).
Moss Report . . . See United States Congress, House.
Peat, Marwick, Mitchell & Co., *Research Opportunities in Auditing* (Peat, Marwick, Mitchell, 1976).
Robertson, J., *Auditing* (Business Publications, Inc., 1976).
Schandl, C., *Theory of Auditing* (Scholars Book Company, 1978).
Scott, W., "A Bayesian Approach to Asset Valuation and Audit Size," *The Journal of Accounting Research* (Autumn 1973), pp. 304-330.
Slovic, P. and S. Lichtenstein, "Comparison of Bayesian and Regression Approaches to the Study of Information Processing in Judgment," *Organizational Behavior and Human Performance* (November 1971).
Sorkin, H. and K. Meuwissen, "A New Technique in Confirmations: The Expanded Field," *The Internal Auditor* (October 1978), pp. 88-92.
Sneed, F., *Parallelism In Two Disciplines* (Arno Press, 1978).
Toba, Y., "A General Theory of Evidence as the Conceptual Foundation in Auditing Theory," *The Accounting Review* (January 1975), pp. 7-24.
Tracy, J., "Bayesian Statistical Methods in Auditing," *The Accounting Review* (January 1969), pp. 90-98.
Turner, J. and T. Mock, "Economic Considerations in Designing Audit Programs," *Journal of Accountancy* (March 1980).
United States Congress, House. Subcommittee on Oversight and Investigation, Chairman Rep. J. Moss, *Federal Regulation and Regulatory Reform* (October 1976).
United States Congress, Senate. Senate Subcommittee on Reports, Accounting and Management, Chairman Sen. L. Metcalf, *The Accounting Establishment* (November 1977).
Vanasse, R., *Statistical Sampling for Auditing and Accounting Decisions: A Simulation* (McGraw-Hill, 1976).
Weber, R., "Auditor Decision Making on Overall System Reliability: Accuracy, Consensus, and the Usefulness of a Simulation Decision Aid," *Journal of Accounting Research* (Autumn 1978).
Windal, F., "Standards of Reliability for Audit Evidence," *The New York Certified Public Accountant* (June 1961).
Wright, A., *An Examination of the Effect of Environmental Cues on Audit Judgments* (Ph.D. dissertation, University of Southern California, 1979).