THE ACCOUNTING REVIEW
Vol. 70, No. 1
January 1995
pp. 27–47

Knowledge Structure and the Estimation of Conditional Probabilities in Audit Planning

Mark W. Nelson
Robert Libby
Cornell University
Sarah E. Bonner
University of Southern California

ABSTRACT: Experienced auditors tend to structure their knowledge of financial statement errors with audit objective as the primary organizing dimension and transaction cycle as secondary. Yet, many audit tasks are structured in the opposite manner, requiring auditors to assess whether objectives are met for each transaction cycle. Our paper reports the results of an experiment which indicates that this mismatch between knowledge structure and task structure may hinder auditors' ability to draw on previous experiences when making conditional probability judgments and when allocating audit hours to various objectives within cycles. These results suggest one instance where knowledge structures that are often functional may have adverse effects when they do not match the task structure to which they are applied.

Key Words: *Knowledge structure, Error frequencies, Probability judgments, Audit planning.*

Data Availability: *Contact the authors.*

I. INTRODUCTION

Auditors' organization of knowledge in memory (their "knowledge structure") is often viewed as one of the keys to effective decision performance (see Bédard and Chi 1993; Libby and Luft 1993; Libby 1994; Smith and Kida 1991). Experienced auditors tend to

We thank KPMG Peat Marwick for supporting this research with a Research Opportunities in Auditing Grant, and for providing additional support to Mark Nelson through a KPMG Peat Marwick Research Fellowship. This research has benefited from the comments of Stephen Asare, Rob Bloomfield, Karl Hackenbrack, Jim Jiambalvo, Jane Kennedy, Bill Messier, Mark Peecher, Steve Salterio, Rick Tubbs, Bill Wright, two anonymous reviewers, and workshop participants at the Universities of Florida, Iowa and Washington.

Submitted April 1994.
Accepted July 1994.

structure their knowledge of financial statement errors with audit objective as the primary organizing dimension and transaction cycle as secondary (Frederick et al. 1994), yet audit tasks are usually structured with transaction cycle as the primary organizing dimension and audit objective as secondary (Arens and Loebbecke 1991; Ernst & Young 1990; KPMG Peat Marwick 1993). We investigate whether this difference between auditors' knowledge structures for financial statement errors and the structure of audit planning tasks adversely affects auditors' ability to access and use previously experienced error frequencies when making conditional probability judgments and audit planning decisions. This issue is important because it indicates that knowledge structures developed through experience may not always enhance audit decisions, implying a need to weigh both the positive and negative effects of alternative knowledge structures when considering interventions designed to either communicate structure to novice auditors or augment the decisions of experienced auditors.

The audit planning process typically requires auditors to allocate audit effort across tests that are designed to satisfy various audit objectives for each transaction cycle. The probability that an audit objective is violated depends on the transaction cycle in question (e.g., validity errors occur more often in the sales and receivables cycle, while completeness errors occur more often in the acquisitions and payments cycle). Given that auditors assess the quality of accounting systems and internal controls at the transaction cycle level (Arens and Loebbecke 1991), it is natural to condition this probability judgment on transaction cycle. Knowledge of the frequency with which errors have occurred under similar circumstances would be useful when making such conditional probability judgments. However, prior research in psychology indicates that frequency knowledge is difficult to access and apply to a probability judgment that is conditioned on a dimension which is different from the primary dimension used to organize knowledge (Gavanski and Hui 1992; Sherman et al. 1992). Thus, given auditors' objective-dominant knowledge structures, auditors' probability judgments and audit planning decisions may reflect frequency information when conditioned on objective (the dominant dimension), but not when conditioned on cycle (the secondary dimension). Because the allocations of audit effort that result from these conditional probability estimates determine to a large extent whether over- or under-auditing occurs, they have an important influence on audit effectiveness and efficiency.

To examine this issue, we performed an experiment in which auditors acquired error frequency knowledge through experience, estimated probabilities of error of the form P(objective error | cycle) or P(cycle error | objective), and allocated hours of audit effort within various transaction cycles and audit objectives. The auditors also performed sort tasks and estimated unconditional frequencies to provide data to identify the dominant dimension of their knowledge structures and to rule out potential alternative explanations for results.

Results were consistent with expectations. The sort data indicated that objective was the dominant feature in auditors' knowledge structures. Consistent with this objective-dominant knowledge structure, auditors' estimates of conditional frequencies of the form P(cycle error | objective) were influenced by the frequencies presented to them in the experiment, while their estimates of conditional frequencies of the form P(objective error | cycle) were not. Allocations of audit hours likewise were influenced by experimental frequencies when auditors distributed hours across cycles for each objective, but not when they distributed hours across objectives for each cycle. Additional data were used to rule out alternative explanations for the results. The auditors estimated unconditional frequencies equally well for both cycle and objective, indicating that the results cannot be explained by differential frequency learning or by auditors merely failing to condition their probability estimates. Also, analyses based on another group of auditors' estimates of real-world frequencies indicated that the results cannot be explained by interference from pre-existing frequency knowledge.

These results imply that the knowledge structures which auditors develop through experience may hinder application of experienced frequencies to audit decisions in audits organized on a cycle basis. More generally, the results indicate that knowledge structures developed through experience or conveyed by training may either facilitate or inhibit auditors' application of experienced frequencies to audit decisions, depending on the degree to which knowledge structure matches task structure.

The rest of this paper proceeds as follows. Section II describes related literature in accounting and psychology and develops our hypotheses. Section III presents the method we used to test the hypotheses. Section IV presents our results. Section V provides discussion of the results and suggests directions for future research.

II. RELATED LITERATURE AND HYPOTHESIS DEVELOPMENT

The Effect of Knowledge Structure on Conditional Probability Estimation

Previous research in psychology has shown that people estimate conditional probabilities by accessing "natural sample spaces" in their knowledge structures (Gavanski and Hui 1992). A natural sample space is one which is likely to be accessed spontaneously and corresponds to the primary dimension of the knowledge structure a person has in his or her mind. For example, auditing textbooks generally portray financial statement errors as categorized primarily according to transaction cycle, and secondarily according to audit objective (e.g., Arens and Loebbecke 1991). Figure 1 depicts such a structure.[1] For an auditor with this knowledge structure, each of the transaction cycle categories forms a natural sample space. If asked to estimate the probability that a given cycle will produce an error violating the validity objective (as opposed to some other objective), that auditor would find it relatively easy to access a cycle category and compare the relative numbers of validity errors to errors violating other objectives for that cycle. The natural sample spaces provided by this auditor's knowledge structure are consistent with, and therefore facilitate, judging P(validity error | sales and receivable cycle), or judgments of any other conditional probabilities of the form P(objective error | cycle).

In contrast, suppose an auditor has the knowledge structure pictured in figure 2, in which errors are categorized primarily according to audit objective, and secondarily according to transaction cycle. In estimating P(validity error | sales and receivable cycle), the auditor possessing this knowledge structure must access and compare the number of sales and receivables errors that violated the validity objective to the numbers of sales and receivables errors that violated all other objectives, which requires that several primary and secondary categories be identified and accessed, rather than only several secondary categories within one primary category. The natural sample spaces provided by this knowledge structure are not consistent with judgments of P(validity error | sales and receivable cycle) (or judgments of any other conditional

[1] Figures 1 and 2 describe alternative dominance relations among dimensions of a knowledge organization, rather than the fundamental cognitive structures used to achieve those dominance relations. A number of fundamental theories are consistent with the dominance relations pictured in these organizations. For example, exemplar models view categories as comprised of previously encountered instances, with some features more important than others for determining categorization (see, e.g., Estes et al. 1989). Prototype models view categorization as occurring through comparison with a category prototype that typifies the most important features of the category (see, e.g., Rosch 1978). Connectionist models view categorization as determined by the associations between categories and their features, with some features having stronger associations than others (see, e.g., Gluck and Bower 1988). At some level many of these models are indistinguishable (Barsalou 1990; Estes 1986). We make no attempt to distinguish among alternative fundamental cognitive theories in this paper. Instead, we concentrate on the influence of knowledge organizations that are consistent with a variety of fundamental cognitive theories.

FIGURE 1
Illustration of a Cycle-Dominant Knowledge Structure

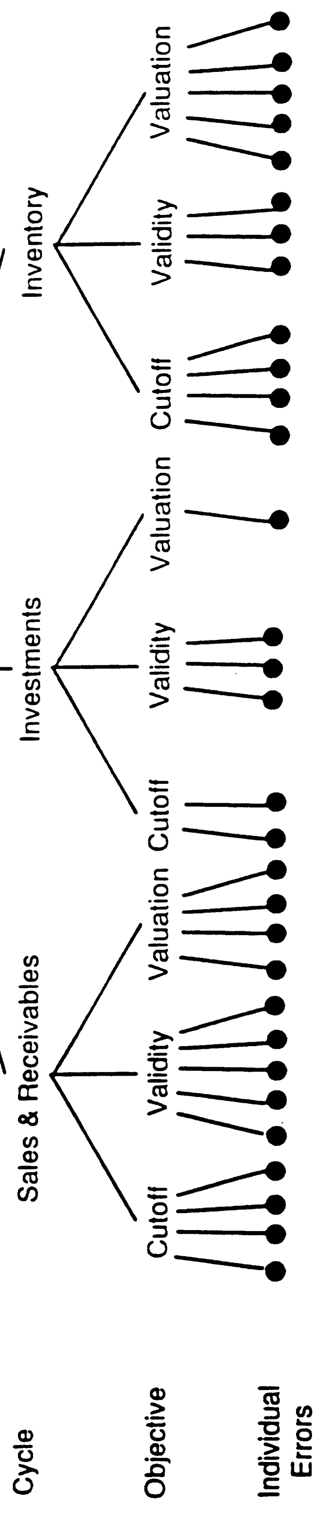
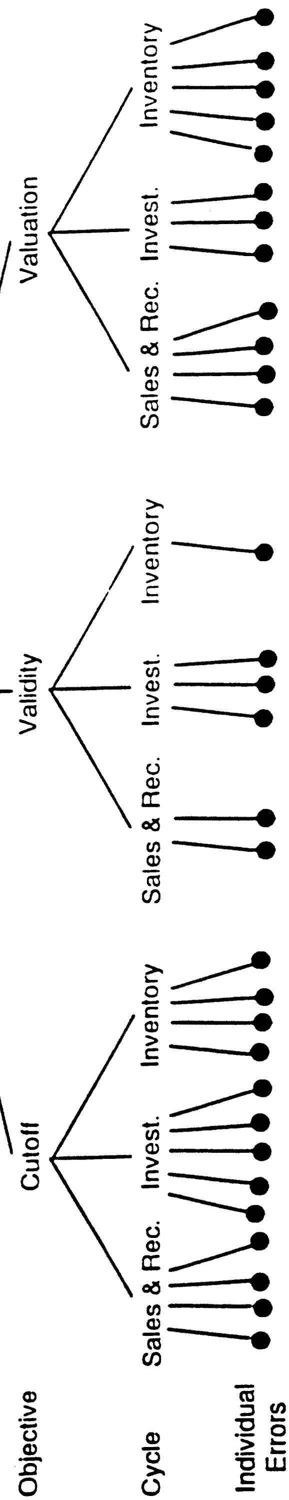


FIGURE 2
Illustration of an Objective-Dominant Knowledge Structure



probabilities of the form P(objective error | cycle)), and thus may hinder retrieval of prior experience for use in the estimation process. Prior psychology research indicates that estimates of conditional probabilities that require accessing such unnatural sample spaces reflect experienced frequencies much less than do estimates that access natural sample spaces (Gavanski and Hui 1992). In fact, subjects making estimates that require them to access unnatural sample spaces often appear to inappropriately access a related natural sample space instead (Sherman et al. 1992).[2] For example, auditors with an objective-dominant knowledge structure might tend to provide P(cycle error | objective) when P(objective error | cycle) was requested.

Note that the knowledge structure depicted in figure 1 (figure 2) is inconsistent (consistent) with judgments of P(cycle error | objective). What determines consistency is whether the knowledge structure organizes primarily on the same dimension that conditions the probability judgment.

Experienced Auditors' Knowledge Structures

Tubbs (1992) demonstrates that audit objective becomes increasingly salient as auditors gain experience, and Frederick et al. (1994) demonstrate that experienced auditors use objective as the primary dimension by which they sort financial statement errors. An objective-dominant organization discriminates between errors based on cause (e.g., validity errors occur because an amount is included in the financial statements that should not be included), which is a naturally-occurring organization in a variety of contexts (Lien and Cheng 1990). This organization may serve a variety of important functions in auditing. For example, its causal orientation may facilitate identification of the source of an error or the control procedure necessary to prevent or detect the error (Tubbs 1992). Also, because many substantive tests are based on general testing techniques that apply to objectives irrespective of cycle (e.g., vouching to examine records for validity errors), an objective-dominant organization may facilitate the design of audit tests (Arens and Loebbecke 1991). Thus, it may by very useful for an objective-dominant structure to develop with experience, because it suits many of the task requirements faced by auditors (Anderson 1990; Murphy and Medin 1985).

However, an objective-dominant knowledge structure is inconsistent with audit decisions that require estimating P(objective error | cycle) from experience, because these decisions require the natural sample spaces that are provided by a cycle-dominant knowledge structure. For example, one audit decision that requires auditors to estimate these conditional probabilities is the allocation in audit planning of the effort to be expended in various audit tests. This decision requires auditors to consider the probability that various audit objectives will be violated and, given that auditors assess the quality of accounting systems and internal controls at the transaction cycle level (Arens and Loebbecke 1991), it is natural to condition this probability judgment on transaction cycle (that is, to judge P(objective error | cycle)).

Knowledge of previously-experienced error frequencies is useful in estimating these conditional probabilities. Prior research indicates that auditors learn error frequencies from experience (Ashton 1991; Butt 1988; Libby 1985; Libby and Frederick 1990; Nelson 1993, 1994), that error frequencies are sometimes considered consciously by auditors when making probability judgments (Waller 1990), and that error frequencies may also influence auditors' judgments in ways that are not conscious, e.g., by influencing the availability of potential hypotheses in memory (Libby 1985). However, no prior studies have examined the influence of frequency information

[2] This behavior is not merely the result of semantic confusion over the meaning of a conditional probability, because subjects asked to estimate conditionals that are consistent with their knowledge structure do not make similar errors.

on judgments of conditional probabilities or the degree to which knowledge structure enhances or diminishes this influence.

This lack of research is a concern because the inconsistency between auditors' objective-dominant knowledge structure and the cycle-dominant structure of audit planning may result in probability estimates that do not reflect experienced frequencies. Note that this inconsistency does not merely imply that auditors' decisions will reflect an unconditional probability estimate instead of a conditional probability estimate (e.g., reflecting P(validity error) rather than P(validity error | sales and receivables cycle)). Rather, the inconsistency between knowledge structure and task structure implies that auditors will attempt to estimate the conditional probability and fail, such that audit planning decisions do not reflect previously experienced frequencies.

Two aspects of the accounting context may lessen the degree to which knowledge structure influences auditors' ability to draw on their experience with error frequencies in estimating conditional probabilities. First, prior research in psychology suggests that human judgment is highly adaptive and responsive to the needs of decision contexts (Anderson 1990). The prior research demonstrating that knowledge structure can hinder estimation of conditional probabilities examined subjects' responses to unfamiliar problems in the laboratory, and so allowed no contextual adaptations to take place. Given the importance of estimating P(objective error | cycle) in the auditing context, experienced auditors may have developed knowledge structures that enable them to apply their frequency knowledge to the estimation process. For example, although prior accounting research indicates that objective is the more dominant dimension in experienced auditors' knowledge structures, the same research demonstrates that auditors sort on either cycle or objective with high accuracy when instructed to do so, suggesting that both dimensions are well developed in auditors' knowledge structures (Frederick et al. 1994). It may be that these two dimensions are close enough in dominance to avoid hindering auditors' judgments of probabilities conditioned on either cycle or objective.

Second, the results of prior psychology studies suggest that the effect of knowledge structure on the estimation of conditionals diminishes with the salience of the features that distinguish categories (Sherman et al. 1992). Prior studies have used either simple pictorial stimuli (e.g., the shapes of alien faces in Gavanski and Hui 1992) or highly discriminable natural categories (e.g., gender, marital status, introvert/extrovert in Sherman et al. 1992; social or occupational categories in Gavanski and Hui 1992). Yet, even for these categories, a decrease in the significance of results was apparent for categories that have distinguishing features that are less salient. For example, Sherman et al. (1992) attribute their finding that results for the gender category were stronger than results for the marital status and introvert/extrovert categories to prior suggestions by Rothbart and Taylor (in press) and Smith and Zarate (1992) that categories related to physical differences are particularly discriminable. The concepts of interest in the auditing context are not defined by obvious characteristics such as physical differences, so it is possible that no influence of knowledge structure on conditional probability estimation will be detectable.

Therefore, to determine the effect of auditors' knowledge structures on their ability to apply prior experience to conditional probability estimates, we tested the following hypothesis:

H1: Auditors' estimates of P(cycle error | objective) are influenced by experimental frequencies more than are auditors' estimates of P(objective error | cycle).

It is also necessary to test the degree to which the effects of knowledge structure on conditional probability estimation extend to audit decisions, because the manner in which audit decisions are made may mitigate the influence of knowledge structure. Specifically, during audit planning an auditor distributes planned audit effort across audit objectives for each transaction cycle. Recall that one explanation for the influence of knowledge structure on estimation of conditional probabilities is that conditioning on the dominant (nondominant) feature of the knowledge structure facilitates (inhibits) the identification of category members and the retrieval of previously experienced instances of those category members. To the degree that providing the names of categories in the audit program performs this identification for auditors, the influence of knowledge structure on planning decisions may decrease. Thus, to determine whether the effects of knowledge structure extend to decisions, the following hypothesis is tested:

H2: Auditors' allocations of audit effort across cycles for each objective are influenced more by experimental frequencies than are auditors' allocations of audit effort across objectives for each cycle.

The following section describes the experiment used to test H1 and H2.

III. METHOD

Subjects

Subjects were 47 auditors from a single Big 6 firm who were enrolled in a staff training session. The auditors' average experience was 3.26 years. The results of Frederick et al. (1994) suggest that this is sufficient time to have developed a knowledge structure in which audit objective serves as the primary basis for sorting financial statement errors. Thus, these auditors are expected to form natural sample spaces for estimating conditional probabilities of financial statement errors on the basis of audit objectives rather than transaction cycles. The auditors were required to participate in the experiment as part of their training activities.

Overview, Design, and Procedures

Overview and Design

Subjects completed all portions of the experiment on Macintosh computers in special breakout rooms configured for computerized instruction. The computerized procedure standardized the timing of frequency presentations, provided immediate feedback during those instruction sequences that required it, and facilitated randomization of items and treatments (to be discussed later) and data recording. Subjects were prohibited from using reference materials or speaking with each other throughout the experiment. An experimenter was available at all times to supervise data collection and answer questions.

Table 1 summarizes the experimental procedures. In the experiment, subjects completed a knowledge structure pretest, observed individual presentations of financial statement errors designed to convey frequencies, and completed a distractor task, conditional probability estimation test, conditional audit decision task, unconditional frequency estimation test, unconditional decision task, and a debriefing questionnaire. The experiment required an average of 41.25 minutes to complete.

The type of conditional probability estimated by the auditors (either P(cycle error | objective) or P(objective error | cycle)) was manipulated between-subjects, because: (1) within-subjects manipulation might encourage subjects to relate their responses to the two types of conditionals, and (2) within-subjects manipulation might increase subjects' fatigue by increasing the number of judgments that each subject had to make. Whether the unconditional frequency estimation test occurred before or after the unconditional decision task was also manipulated between-subjects to provide exploratory data for another study. Subjects were randomly assigned to one of the four combinations of these between-subjects factors and proceeded as follows.

TABLE 1
Sequence of Experimental Procedures

1. Free-sort nine individual errors
2. Presentation of nine individual errors in varying frequencies (total of 49)
3. Distractor task (4 ability questions)
4. Conditional probability questions (2 questions conditioned on either cycle or objective)
5. Conditional audit decision task (3 questions allocating hours across cycles for each objective or across objectives for each cycle)
6. Unconditional frequency estimation task (6 questions estimating frequency of error for each cycle and objective)
7. Unconditional audit decision task (2 questions allocating hours across cycles and across objectives)
8. Directed sort by cycles
9. Directed sort by objectives
10. Demographic questions

Note: Order of 6 and 7 manipulated between subjects; order of 8 and 9 determined randomly.

Materials and Procedure

Sort tasks are commonly used in cognitive psychology to determine how objects are classified.[3] To determine whether our subjects possessed the objective-dominant knowledge structure observed by Frederick et al. (1994), subjects were asked to sort nine financial statement errors into categories on the basis of "how they thought the errors best go together." This "free sort" task thus requires subjects to select some primary organizing dimension. Subjects were required to sort the errors into three groups to hold constant the level of detail in their sorts, thus facilitating data analysis (discussed in section IV). To accomplish the free sort task, subjects clicked on either a "1," "2," or "3" beside each error to indicate how they would group the errors. The order of presentation of the nine errors in the free sort task was randomized and held constant across subjects. Subjects could change their groupings at any time during the sort task, and were prompted before proceeding to reconsider their answers. The free sort task is shown in appendix A.

The nine financial statement errors used in the free sort task and the remainder of the experiment were selected to fully cross three transaction cycles with three audit objectives. The

[3] A less-direct method commonly used in the accounting literature is examining the clustering of free recalls (e.g., Choo and Trotman 1991; Weber 1980). See Chi et al. (1982) for a classic study employing a sort-task methodology, and Bédard and Chi (1993) for a discussion of the advantages of using this technique in auditing.

transaction cycles and audit objectives used were determined by examining practice manuals from several large public accounting firms to identify those transaction cycles and audit objectives for which firms demonstrate the highest agreement in categorizing financial statement errors. The results of Frederick et al. (1994) were then used to identify errors for which there was the highest agreement on both cycle and objective category membership among audit managers affiliated with the firm providing subjects for this study. The transaction cycles, audit objectives and financial statement errors resulting from this selection process are shown in table 2.[4] These cycles, objectives and errors were the same as those used in Bonner et al. (1993).

After completing the free sort task, subjects viewed presentations of the nine individual financial statement errors in varying frequencies, for a total of 49 presentations. A different random order of the 49 presentations was used for each subject. To avoid primacy and recency effects, the random orders were determined with the constraints that: (1) each repetition of an error was separated by at least one other error, and (2) the three highest frequency errors were presented as often in the first half of the sequence as in the second half. The frequencies of individual errors were determined by randomly assigning the three transaction cycles and the three audit objectives shown in table 2 to the rows and columns shown in table 3 (which describes the presentation frequencies) for each subject.[5] Subjects were instructed that they would be presented with a series of errors found during the audits of medium-sized manufacturers by a major accounting firm

[4] The names of the objectives and cycles were the same as those used by the firm providing subjects.

[5] The results of this experiment should not be influenced by any frequency knowledge that auditors possessed prior to the experiment. Prior research indicates that frequency knowledge is time-tagged in memory (Hintzman and Block 1973; Hintzman et al. 1973; Reichardt et al. 1973), such that experimental frequencies can be discriminated from pre-existing frequency knowledge (Butt 1988; Nelson 1993). Also, the random assignment of cycles and objectives to experimental frequencies should ensure that any influence of pre-existing frequency knowledge is spread across treatments. As discussed further in the results section, we gathered data from another group of auditors which confirmed that results were not influenced by subjects' pre-existing frequency knowledge.

TABLE 2
Financial Statement Errors

	Transactions Cycle		
Audit Objective	*Sales*	*Inventory/Purchases*	*Investments*
Proper Cutoff	Next period's sales were included in the current year's revenue and receivables	Raw materials were improperly shown as received after year-end.	Purchases of treasury bills were recorded in the wrong fiscal period.
Validity	Billings to legitimate customers were booked twice.	More finished goods were recorded as received than were actually received.	Fictitious investments were included in the account balance.
Valuation	The bad debt expense and allowance were underestimated.	Obsolete inventory was not written down to net realizable value.	Marketable securities were not reduced to lower of cost or market.

during the last quarter of 1992, and that their task was to remember the errors to the best of their ability. No mention was made of frequency learning, or of the tasks that subjects would perform after the error presentations. The subjects were also told that each error would remain on the computer screen for ten seconds. The structure and content of this frequency presentation task is the same as that used by Bonner et al. (1993), and similar to that used by Butt (1988).[6]

After viewing the individual error presentations, subjects answered four questions from the ability test used by Bonner and Walker (1994) as a distractor task to clear short-term memory. Next, subjects were asked to imagine that they were auditors who were working for the firm from which the list of errors was obtained, and to assume that they had been exposed to all of those audits, such that the errors constituted their experience. Then the subjects performed the conditional probability estimation test.

The conditional probability estimation test required judgments of two conditional probabilities, with the specific type of conditional depending on the assignment of the subject to either P(cycle error | objective) or P(objective error | cycle) at the beginning of the experiment. Subjects assigned to the P(cycle error | objective) task estimated conditionals of the form P(high-frequency cycle | low-frequency objective) and P(low-frequency cycle | high-frequency objective), while subjects assigned to the P(objective error | cycle) task estimated conditionals of the form P(high-frequency objective | low-frequency cycle) and P(low-frequency objective | high-frequency cycle). Only these conditionals were estimated because they differ the most in the conditional probability implied by the frequencies presented to subjects (1/7 for P(low | high), 4/7 for P(high | low)), and thus allow detection of an effect without a large number of estimates. The identity of the high- or low-frequency cycles and objectives referenced in the conditionals depended on the random assignment at the beginning of the experiment of the errors in table 2 to the columns and rows of table 3. Subjects answered each question by moving a pointer from the "zero" endpoint of a sliding scale numbered from zero to 100 to a position that indicated their probability estimate. The order of the two questions (P(high | low) and P(low | high)) was randomized between subjects. An example of a conditional probability estimation test question is shown in appendix B.

[6] The differences between the presentations in this experiment and Butt (1988) are minor; e.g., presentations are by computer rather than slide projector, exposure durations differ slightly.

TABLE 3

Individual and Category-Level Frequencies of Errors

Audit Objective	*Transaction Cycle*			
	A	*B*	*C*	*Total — Objective Category*
I	1	2	4	7
II	2	4	8	14
III	4	8	16	28
Total — Cycle Category	7	14	28	49

After completing the conditional probability estimation test, subjects viewed an instruction screen, and then completed the conditional audit decision task. The conditional audit decision task asked subjects to assume that they were planning the audit of a medium-sized manufacturing company that is similar to the manufacturing companies whose errors they saw previously in the study, and required subjects to perform three different allocations of 100 hours of audit effort.[7] A subject assigned to P(objective error | cycle) (P(cycle error | objective)) for the conditional probability estimation test allocated 100 hours of audit effort across the three audit objectives (transaction cycles) used in the experiment and an "other" category for each of the three cycles (objectives), thus performing three 100 hour allocation tasks. For each allocation task, subjects moved a pointer from the "zero" endpoint on a sliding scale with endpoints of zero and 100 to indicate the hours of effort they would budget for each category, and were not allowed to proceed unless they had allocated all 100 hours. The order of the three allocation tasks and the order of categories within each task were randomized between subjects, with the constraint that the "other" category always appeared fourth. An example of a conditional audit decision task question is shown in appendix C.

After performing the conditional probability estimation test and the conditional audit decision task, subjects answered six unconditional frequency estimation questions (one for each of three cycles and three objectives) and performed two unconditional audit decision tasks (one allocating across cycles, the other across objectives). The unconditional questions elicited marginal category frequencies (7, 14 or 28, see table 3). They were the same frequency tests and audit decision tasks used in Bonner et al. (1993). The unconditional frequency estimation test provided data to determine that subjects had attended to the experimental frequencies. Subjects answered each question by moving a pointer from the "one" endpoint of a sliding scale numbered from one to 34 to a position that indicated their frequency estimate. These endpoints were chosen to range six below and six above the minimum and maximum actual category frequencies to insure that answers were not obvious. The order of the two unconditional frequency estimation tests (cycle or objective) and the order of the three cycle or objective questions within each test were randomized between subjects.[8] The unconditional audit decision task provided exploratory data unrelated to this study. Except with respect to determining whether subjects learned experimental frequencies, the results from the unconditional tasks will not be discussed further in this study.

Following the unconditional frequency test and unconditional audit decision task, subjects performed two additional sort tasks. These sort tasks were identical to the free sort task at the beginning of the study, except one task directed subjects to sort according to transaction cycle, and the other according to audit objective. Data from these "directed sort" tasks were tested to ensure that differences between the "type of category given" treatments could not be attributed to differences in subjects' knowledge of the transaction cycles and audit objectives related to each error. The experiment concluded with subjects answering some demographic questions about age, years of experience in public accounting and auditing, title, client industry emphasis, and whether English was a first or second language.

---

[7] A higher probability of error might influence not only budgeted hours of audit effort, but also the timing, extent and level of staff assigned to audit tests. To the extent that the audit hour allocation task does not encompass these audit decisions, results are biased against supporting H2.

[8] Recall that whether or not the unconditional frequency estimation test preceded the unconditional audit decision task was a between-subjects variable unrelated to this study.

IV. RESULTS

Verification of Objective as Dominant Dimension in Knowledge Structure

To determine whether our subjects possessed the objective-dominant knowledge structure identified by Frederick et al. (1994), subjects' free sorts were compared to the free sort that would have been made had subjects sorted according to objective or cycle. To conduct these comparisons, each subjects' set of nine free sort responses was converted into a nine-by-nine similarity matrix containing "ones" if errors were grouped together and "zeros" if errors were not grouped together. Nine-by-nine similarity matrices were also formed for the predetermined transaction cycle and audit objective categorizations presented in table 2. The fit of each subject's similarity matrix to a cycle and an objective categorization was measured by correlating the vector formed from the lower left triangle of the subject's similarity matrix with the vector formed from the lower left triangle of the predetermined cycle and objective similarity matrices. This measure of the correspondence between two similarity matrices is commonly called a "cophenetic correlation" (Sneath and Sokal 1973).

The average cophenetic correlation between subjects' free sort similarity matrices and the predetermined objective (cycle) similarity matrix is .53 (-.04). The average correlation with objective is significantly higher than the average correlation with cycle ($t=4.22$, two-tailed $p=.0001$), indicating that the audit objective dimension was more dominant than the transaction cycle dimension on average in our auditors' free sorts.[9]

Test of H1

H1 predicts that subjects' conditional probability judgments will be influenced more by experimental frequencies for P(cycle error | objective) than for P(objective error | cycle). In other words, with an objective-dominant knowledge structure, subjects should be better able to discriminate event frequencies when a probability judgment is conditioned on objective than when it is conditioned on cycle. This implies that the difference between P(high frequency cycle error | low frequency objective) and P(low frequency cycle error | high frequency objective) should be greater than the difference between P(high frequency objective error | low frequency cycle) and P(low frequency objective error | high frequency cycle). Recall that each subject completed two conditional probability judgments (P(high | low), P(low | high)) with only one of the two categories (cycle or objective) as "given" (i.e., judgments were either of the form P(objective error | cycle) or P(cycle error | objective)). Thus, the influence of knowledge structure on estimation can be tested in a 2 x 2 ANOVA, with type of conditional probability estimate (P(high | low), P(low | high)) a within-subjects variable and given category type ((. | objective), (. | cycle)) a between-subjects variable. An interaction between these two variables of the form proposed above would support H1. Specifically, the influence of type of conditional probability estimate should be greater when the given category type is "objective."

The mean conditional probability estimates are shown in table 4. The interaction between type of conditional probability estimate and given category type is significant ($t=1.78$; one-tailed $p=.04$), supporting H1. The auditors' conditional probability estimates were influenced by frequency information when the conditional was of the form P(cycle | objective) ($t=2.10$; one-tailed $p=.02$), but not when it was of the form P(objective | cycle) ($t=-0.33$; one-tailed $p=.63$).

---

[9] Thirty-six of the 47 subjects' sorts (77%) were more like the predetermined objective sort than the predetermined cycle sort, and no results changed when only these subjects were included in analyses. The proportion of subjects whose sorts favored the dominant dimension is similar to results found in prior psychology research (Gavanski and Hui 1992).

TABLE 4
Conditional Probability Estimates

Analysis of Conditional Type[1] x
Presented Conditional Probability Contrast

Means (Standard Deviations) [Contrast Weights]

Presented Conditional Probability	Estimates of P(Objective I Cycle)	Estimates of P(Cycle I Objective)
P(low I high) = 4/28 = 14.3%	48.6 (30.7) [-1]	40.6 (22.4) [+1]
P(high I low) = 4/7 = 57.1%	46.3 (31.0) [+1]	56.8 (26.0) [-1]

[1] "Conditional Type" is either P(cycle I objective) or P(objective I cycle).

Test of H2

H2 predicts that subjects' audit effort allocations will be influenced more by experimental frequencies when subjects allocate effort across transaction cycles for each audit objective than when subjects allocate effort across audit objectives for each transaction cycle. Recall that, in the effort allocation task, subjects who had estimated conditional probabilities with cycle as "given" allocated effort across objectives for each cycle, while subjects who had estimated conditional probabilities with objective as "given" allocated effort across cycles for each objective. For example, a subject with cycle as "given" made a total of 12 responses in the effort allocation task, consisting of P(low frequency objective error I cycle), P(medium frequency objective error I cycle), P(high frequency objective error I cycle), and P(other objective errors I cycle) for each of the low, medium, and high frequency cycles that appear as given in the conditional. Dropping allocations to P("other" I .), each subject's remaining nine responses can be classified in a 3 x 3 design, consisting of "given category *frequency*" ((. I low), (. I medium), and (. I high)), and "event category *frequency*" (low I .), (medium I .), and (high I .)). When coupled with the between-subjects manipulation of the *type* of the category given ((. I objective), (. I cycle)), the resulting model is a 2 x 3 x 3 ANOVA. H2 would be supported by a significant given category *type* x linear trend[10] in event category *frequency* contrast. That is, event category frequencies should influence audit effort allocations more when the given category type is "objective." Note that, if subjects are estimating the requested conditional probability, given category *frequency* should not have an effect.[11]

The mean audit effort allocations are shown in table 5. The given category type x linear trend in event category frequency contrast is significant (t=2.24; one-tailed p=.01).[12] The linear trend

[10] See Rosenthal and Rosnow (1988) for a discussion of the use of linear trends and focused contrasts in analysis of variance.

[11] Referring to table 3, regardless of whether the given category frequency is low (seven presentations), medium (14 presentations) or high (28 presentations), the relative event category frequency is 1:2:4. Therefore, the proportion of 100 hours of audit effort allocated to a particular event category should not depend on given category frequency.

[12] A conventional test of the interaction between event category frequency and given category type yields a p-value of .0322 (F=3.57).

was significant when the given category was objective (t=2.53; one-tailed p=.008), but not when the given category was cycle (t=-0.55; one-tailed p=.71), supporting H2. The auditors' audit effort allocations were influenced by frequency information when distributing effort across cycles for each objective, but not when distributing effort across objectives for each cycle.

Ruling Out Alternative Explanations

Knowledge of Both Cycle and Objective

One explanation for a difference in results between the "type of category given" treatments is that subjects might differ in their knowledge of the cycle and objective dimensions. To test for this possibility, the directed sorts that subjects performed at the end of the experiment were compared to the sorts that would have been made had subjects sorted according to objective and cycle. The mean cophenetic correlation between the directed objective (cycle) sort and the predetermined objective (cycle) sort was .71 (.74). These two means are not significantly different (t=0.38; two-tailed p=.71), indicating that our subjects' knowledge of objective and cycle was roughly equivalent. This result indicates that support for H1 and H2 cannot be attributed to differential understanding of membership of errors in objective and cycle categories.

Knowledge of Experimental Frequencies

Another explanation for a difference in results between the "type of category given" treatments is that subjects might differ in the degree to which they acquired knowledge of the experimentally-manipulated frequencies of error in the cycle and objective categories, and then simply answered the conditional probability questions using their estimates of unconditional cycle or objective frequencies. To test this possibility, the degree to which the experimental frequencies are reflected in each subject's estimates of unconditional cycle and objective frequencies were evaluated and compared. The mean unconditional frequency estimates are shown in table 6. The linear trend in presented frequencies is significant (t=5.15; one-tailed p<.0001), indicating that subjects acquired unconditional frequencies from their experience in the experiment. The linear trend in presented frequencies x category-type (cycle or objective)

TABLE 5
Conditional Audit Effort Allocations

Analysis of Conditional Type[1] x
Linear Trend in Presented Event Category Frequencies[2] Contrast

Means (Standard Deviations) [Contrast Weights]

Presented Frequency	*Conditioned on Cycle*	*Conditioned on Objective*
7	33.3 (15.5) [-1]	25.3 (15.1) [+1]
14	27.5 (14.4) [0]	31.5 (13.2) [0]
28	31.0 (16.1) [+1]	37.4 (17.3) [-1]

[1] "Conditional Type" is either P(cycle | objective) or P(objective | cycle).

[2] "Presented Event Category Frequencies" are the frequencies of the events that are conditioned on either objective or cycle categories. For example, assume that the Sales and Receivable cycle had a frequency of 28. In the conditional probability P(Sales Error | Cutoff), Sales is the event being conditioned on Cutoff, so "presented event category frequency" has a value of 28.

contrast is not significant (t=0.00; two-tailed p=0.95), indicating that the degree to which subjects' frequency estimates reflected experimental frequencies did not differ significantly between the cycle and objective categories.[13] These results indicate that support for H1 and H2 cannot be attributed to differential learning of unconditional frequencies. Also, because subjects estimated unconditional cycle and objective frequencies equally well, but differed in the degree to which conditional probabilities reflected frequencies depending on whether the probability was conditioned on cycle or objective, we can conclude that subjects were not merely answering the conditional probability questions with their estimates of unconditional probabilities.

Knowledge of Non-Experimental Frequencies

A final possible explanation for the results is that pre-existing frequency knowledge interfered with the experimental frequencies in some systematic way. Prior research, the design of our experiment, and further analyses all indicate that this is not a plausible explanation. First, as mentioned previously, prior research indicates that frequency knowledge is time-tagged in memory (Hintzman and Block 1973; Hintzman et al. 1973; Reichardt et al. 1973), such that experimental frequencies can be discriminated from pre-existing frequency knowledge (Butt 1988; Nelson 1993, 1994). Second, recall that an element of our design was that the frequencies of the cycle and objective categories (and thus of the errors) were assigned randomly between subjects. Examination of the assignments that occurred indicated that this random assignment was successful—each category appeared as the low, medium, and high frequency category approximately the same number of times. Therefore, any influence of pre-existing frequency knowledge that did occur should have been distributed across treatments, and thus would only have decreased the power of our tests. Third, to insure that pre-existing frequency knowledge did not bias results, we asked a separate group of 46 auditors who were enrolled in the same staff training session to rate the relative frequencies of the three cycle and the three objective categories as part of another study. The auditors were not exposed to the experimental frequencies, so these ratings constituted their estimates of real-world frequencies. Results indicated that valuation errors are considered to occur relatively more frequently than cutoff and validity errors, and that

[13] Whether the unconditional frequency estimation task was completed before or after the unconditional audit decision task did not influence the effect of actual frequency on unconditional frequency estimate (t=0.79; two-tailed p=.43).

TABLE 6

Unconditional Frequency Estimates
Analysis of Category (Cycle v. Objective) x Linear Trend in Presented Frequencies Contrast

Means (Standard Deviations) [Contrast Weights]

Presented Frequency	*Unconditional Objective Estimates*	*Unconditional Cycle Estimates*
7	12.5 (7.1) [-1]	13.0 (7.1) [+1]
14	12.8 (5.9) [0]	13.8 (6.9) [0]
28	18.4 (7.7) [+1]	18.9 (9.0) [-1]

investments errors are considered to occur relatively less frequently than sales and inventory errors. Therefore, hypothesis H1 was re-tested after omitting those subjects whose conditional probability estimates involved either the valuation objective or the investments cycle.[14] Results were unchanged. This result indicates that support for the hypotheses is not driven by interference from pre-existing knowledge of error frequencies.

IV. DISCUSSION

As noted previously, a growing literature in accounting examines the influence of knowledge structure on important cognitive processes like recall and frequency estimation. For the most part this literature concludes that structure enhances audit judgment. For example, in their recent examination of 25 studies of heuristics and biases in auditing, Smith and Kida (1991) document that biases often are mitigated when *knowledgeable* subjects perform *familiar* tasks. This possibility was first raised by Joyce and Biddle (1981a,b). However, Smith and Kida further point out that Frederick and Libby's (1986) results using a highly abstract task suggest that, depending on the relationship between the auditor's knowledge structure and the task, knowledge could either help or hinder auditors' judgments. The current paper demonstrates a case where the auditor's knowledge structure, as developed through experience, actually hinders the auditor's ability to apply knowledge acquired through experience to an important audit judgment. While the estimation of conditional probabilities (required to test our hypothesis 1) is an abstract task, the allocation of audit effort to audit objectives within transaction cycles (used in testing hypothesis 2) is a relatively familiar, intuitive task to most experienced auditors.

Several specific results merit further discussion. The free sort results replicate Frederick et al. (1994) by demonstrating that audit objective dominates transaction cycle in experienced auditors' free sorts of audit errors. Although this structure might be quite useful for a variety of audit tasks, the results of tests of H1 and H2 suggest that it hinders auditors' ability to draw on previously experienced frequencies when estimating conditional probabilities of the form P(objective error | cycle) and when allocating audit hours to various objectives within transaction cycles. Other tests indicate that these results cannot be attributed to interference from pre-existing frequency knowledge or to differences in understanding or frequency learning between cycle and objective categories.

These results extend the results of prior psychology research in three ways. First, we demonstrate that the influence of knowledge structure occurs with respect to categories and stimuli whose features are less saliently defined than those used in prior research (e.g. Gavanski and Hui 1992; Sherman et al. 1992), supporting the generality of the effects identified in prior research. Second, we demonstrate that the influence of knowledge structure occurs for experienced professionals making intuitive, familiar judgments, indicating that prior results are not due to semantic confusion in answering conditional probability questions and suggesting that these effects are not averted through adaptation in meaningful contexts. Third, we demonstrate that the influence of knowledge structure on conditional probability judgments is powerful enough to be observed in decisions, even when the decisions are structured to provide category names for use as retrieval cues.

From an audit practice perspective, these results indicate that the current structure of audit planning judgments may not always be conducive to auditors utilizing their knowledge of the frequencies of previously experienced errors. This does not necessarily imply that the audit

[14] Because H2 required subjects to allocate hours of audit effort across either all cycles or all objectives, it could not be tested for a subset of cycles or objectives.

planning task should be restructured to better match auditors' knowledge organization. It is unlikely that the benefits of this restructuring would exceed the costs, given that the current cycle-dominant structure facilitates the identification of inherent and control risks that vary from cycle to cycle. However, auditors' knowledge organizations could be changed to better match the audit planning task. For example, training approaches could be used to render cycle equally or more dominant than objective in the knowledge structures of experienced auditors and thus better match knowledge structure to audit structure. Likewise, college instruction or early staff training could be modified to communicate better to novice auditors a cycle-dominant organization that matches the structure of the audit decisions to which their future professional experience will eventually be applied. Finally, decision aids might be constructed that provide auditors with retrieval cues for individual errors and combination rules for formation of appropriate conditional probabilities.

APPENDIX A
Sort Task Screen

1 2 3 Next period's sales were included in the current year's revenue and receivables.

1 2 3 More finished goods were recorded than were actually received.

1 2 3 Marketable securities were not reduced to lower of cost or market.

1 2 3 Fictitious investments were included in the account balance.

1 2 3 The bad debt expense and allowance were underestimated.

1 2 3 Raw materials were improperly shown as received after year end.

1 2 3 Obsolete inventory was not written down to net realizable value.

1 2 3 Purchases of treasury bills were recorded in the wrong fiscal period.

1 2 3 Billings to legitimate customers were booked twice.

Classify the errors into 3 groups. Each group should contain the errors that you believe go together. After you are finished, click on "done" to continue.

Help

Done

APPENDIX B
Conditional Probability Estimation Screen

Please base your answer to the following question on the error frequencies to which you were exposed earlier in the study.

Given that you are dealing with an error that occurred in the TRADE RECEIVABLES, SALES AND COLLECTIONS CYCLE, what is the probability that the error violated the VALIDITY OBJECTIVE?

Please respond by moving the slider to the correct probability on the percentage scale below.

0 100

39

continue

APPENDIX C
Conditional Audit Decision Screen

Assume you have a hundred hours to allocate to finding errors that occur in the TRADE RECEIVABLES, SALES AND COLLECTIONS cycle. Please allocate the hours to the following categories of audit objective:

PROPER CUTOFF

11

VALUATION

19

VALIDITY

55

Other

15

Remaining: 0

Help

Done

REFERENCES

Anderson, J. R. 1990. *The Adaptive Character of Thought.* Hillsdale, NJ: Erlbaum.

Arens, A. A., and J. K. Loebbecke. 1991. *Auditing: An Integrated Approach.* 5th ed. Englewood Cliffs, NJ: Prentice Hall.

Ashton, A. H. 1991. Experience and error frequency knowledge as potential determinants of audit expertise. *The Accounting Review* (April): 218–239.

Barsalou, L. W. 1990. On the indistinguishability of exemplar memory and abstraction in category representation. In *Advances in Social Cognition* Vol. 3, edited by T. K. Srull and R. S. Wyer. Hillsdale, NJ: Erlbaum.

Bédard, J., and M. Chi. 1993. Expertise in auditing. *Auditing: A Journal of Practice & Theory* (Supplement): 21–45.

Bonner, S. E., R. Libby, and M. W. Nelson. 1993. Category instruction and the acquisition and use of category-level error frequency knowledge. Working paper, Cornell University, Ithaca, NY.

———, and P. L. Walker. 1994. The effects of instruction and experience on the acquisition of auditor knowledge. *The Accounting Review* (January): 157–178.

Butt, J. 1988. Frequency judgments in an audit-related task. *Journal of Accounting Research* (Autumn): 315–330.

Chi, M. T. H., R. Glaser, and E. Rees. 1982. Expertise in problem solving. In *Advances in the Psychology of Human Intelligence* 1, edited by R. Sternberg. Hillsdale, NJ: Erlbaum.

Choo, F., and K. T. Trotman. 1991. The relationship between knowledge structure and judgments for experienced and inexperienced auditors. *The Accounting Review* (July): 464–485.

Ernst & Young. 1990. *Audit Manual, Vol. 1.* Ernst and Young International, Limited.

Estes, W. K. 1986. Array models for category learning. *Cognitive Psychology* 18: 500–549.

———, J. A. Campbell, N. Hatsopoulos, and J. B. Hurwitz. 1989. Base-rate effects in category learning: A comparison of parallel network and memory storage-retrieval models. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 15: 556–571.

Frederick, D. 1991. Auditors' representation and retrieval of internal control knowledge. *The Accounting Review* (April) 240–258.

———, V. Heiman-Hoffman, and R. Libby. 1994. The structure of auditors' knowledge of financial statement errors. *Auditing: A Journal of Practice & Theory* (Spring): 1–21.

———, and R. Libby. 1986. Expertise and auditors' judgments of conjunctive events. *Journal of Accounting Research* (Autumn): 270–290.

Gavanski, I., and C. Hui. 1992. Natural sample spaces and uncertain belief. *Journal of Personality and Social Psychology* (November): 766–780.

Gluck, M. A., and G. H. Bower. 1988. From conditioning to category learning: an adaptive network model. *Journal of Experimental Psychology: General* 117: 227–247.

Hintzman, D. L., and R. A. Block. 1973. Memory for the spacing of repetitions. *Journal of Experimental Psychology* 99: 70–74.

———, R. A. Block, and J. J. Summers. 1973. Contextual associations and memory for serial position. *Journal of Experimental Psychology* 97: 220–229.

Joyce, E., and G. Biddle. 1981a. Anchoring and adjustment in probabilistic inference in auditing. *Journal of Accounting Research* (Spring): 120–145.

———, and ———. 1981b. Are auditors' judgments sufficiently regressive? *Journal of Accounting Research* (Autumn): 323–349.

KPMG Peat Marwick. 1993. *Audit Program Software.*

Libby, R. 1985. Availability and the generation of hypotheses in analytical review. *Journal of Accounting Research* (Autumn): 648–667.

———. 1994. The role of knowledge and memory in audit judgment. In *Judgment and Decision Making Research in Accounting and Auditing*, edited by R. H. Ashton and A. H. Ashton. Cambridge, MA: Cambridge University Press (forthcoming).

———, and J. Luft. 1993. Determinants of judgment performance in accounting settings: Ability, knowledge, motivation, and environment. *Accounting, Organizations, and Society* (July): 425–450.

Lien, Y., and P. Cheng. 1990. *A Bootstrapping Solution in the Problem of Differentiating Spurious from Genuine Causes.* Paper presented at the 31st Annual Meeting of the Psychonomic Society.

Murphy, G. L., and D. L. Medin. 1985. The role of theories in conceptual coherence. *Psychological Review* (July): 289–316.

Nelson, M. W. 1993. The effects of error frequency and accounting knowledge on error diagnosis in analytical review. *The Accounting Review* (October): 803–824.

Nelson, M. W. 1994. The learning and use of frequency information in auditing. Working paper, Cornell University, Ithaca, NY.

Reichardt, C. S., J. J. Shaughnessy, and J. Zimmerman. 1973. On the independence of judged frequencies for items presented in successive lists. *Memory & Cognition* 1: 149–156.

Rosenthal, R., and R. L. Rosnow. 1988. *Contrast Analysis: Focused Comparisons in the Analysis of Variance.* New York: Cambridge University Press.

Rothbart, M., and M. Taylor. 1992. Category labels and social reality: Do we view social categories as natural kinds? In *Language, Interaction and Social Cognition*, edited by G. Semin and K. Fiedler. London: Sage Publications.

Rosch, E. 1978. Principals of categorization. In *Cognition and Categorization,* edited by E. Rosch and B. B. Lloyd. Hillsdale, NJ: Erlbaum.

Sherman, S. J., M. N. McMullen, and I. Gavanski. 1992. Natural sample spaces and the inversion of conditional judgments. *Journal of Experimental Social Psychology* (September): 401–421.

Smith, E. R., and M. A. Zarate. 1992. Exemplar-based model of social judgment. *Psychological Review* (January): 3–21.

Smith, J. F., and T. Kida. 1991. Heuristics and biases: Expertise and task realism in auditing. *Psychological Bulletin* (May): 472–489.

Sneath, P. H., and R. R. Sokal. 1973. *Numerical Taxonomy.* San Francisco: W. H. Freeman.

Tubbs, R. M. 1992. The effect of experience on the auditor's organization and amount of knowledge. *The Accounting Review* (October): 783–801.

Waller, W. S. 1990. *Auditors' probability assessments: A contructive perspective.* Monograph in process.

Weber, R. 1980. Some characteristics of the free recall of computer controls by EDP auditors. *Journal of Accounting Research*: 214–241.