1. Introduction

Bonner and Lewis investigate the determinants of auditor expertise by examining the roles of various types of knowledge and ability in auditor performance. They conclude that differences in knowledge and ability better explain variance in task performance than does experience. The paper is an innovative step in developing an understanding of auditor expertise. Bonner and Lewis are the first auditing researchers to elaborate what elements might determine expert performance in a number of different auditing tasks and to provide an estimate of the relative importance of those elements in each task. Their paper prompted a broad discussion of auditor expertise and its measurement by conference participants. This paper summarizes and extends that discussion by focusing on the two main issues raised: (1) What is expertise in auditing? and (2) How can an auditing expert be identified?

2. What Is Expertise in Auditing?

Much of the conference discussion focused on the question of what constitutes an auditing expert. Bonner and Lewis define expertise as task-specific superior performance and criticize the use of experience as an operational measure of expertise in auditing research. This definition of expertise is incomplete, in the sense that it does not include any notion of how an individual might develop this ability to achieve "task-specific superior performance." Most theoretical and experimental discussions of expertise include the axiom, "practice makes perfect," to account for the development of expertise (e.g., Frensch and Sternberg [1989]). Thus, these definitions of expertise include the idea that expertise is the acquired-through-practice ability to perform well in a task domain: that is, experience is a prerequisite for acquiring expertise.

Bonner and Lewis define superior performance as task performance. Superior performance may be defined in a number of other ways, some of which may be of greater relevance to auditing than task performance. In auditing, performance might also be defined in terms of overall job performance or cognitive performance. Superior overall job performance might best be measured by title or position (Schmidt et al. [1988]), although it is probably highly correlated with experience since seniority is often a consideration in promotion decisions. However, as pointed out by conference participants, job performance as measured by title or position is not necessarily highly correlated with superior technical expertise if, as suggested by Davis and Solomon [1989], some auditors are promoted on their revenue-generating ability. If superior performance is
defined as cognitive performance, then superior memory or reaction times might be used to assess expert performance. If superior performance is defined as task performance, as in Bonner and Lewis, then is one an expert auditor if he or she performs well at one auditing task (such as internal control evaluation) but not at other auditing tasks? Or is auditing expertise task specific, so that some auditors are experts at internal control evaluation, some at ratio analysis, etc.? The multiple task approach taken by Bonner and Lewis enables them to examine the performance of the same auditors on a number of different auditing tasks as well as the aggregate performance of a given auditor on multiple tasks. Cross-sectional performance differences are a function of the nature of the task (Voss et al. [1986]). Thus, in investigating expertise, it is important to examine one expert's performance across several tasks. (Bonner and Lewis do not analyze these issues, though their data could be used to do so.) Moreover, a multiple task approach could be used to ascertain what features distinguish expert performance from novice performance, and whether there is more variation in two experts performing the same task than there is in one expert performing two different tasks.

Conference participants suggested an alternative framework to that used by Bonner and Lewis for conceptualizing the role of various factors in performance: think of performance as the output and identify the inputs and their relations. In one possible model of the process, ability and experience can be described as the inputs and task performance as the output; the inputs and outputs are linked by knowledge and strategies. Given this model of the process, as one conference participant performance when knowledge variables are introduced. Experience represents the opportunity to acquire knowledge; knowledge variables are a measurement of knowledge actually acquired.

Although a model of auditing task performance is certainly desirable, several features of the auditing domain create problems for the development of such a model. For example, auditing, unlike some other domains such as chess or bridge in which expertise has been studied, involves many different tasks without much repetition. In addition, as Davis and Solomon [1989] suggest, auditors do not often receive feedback, so they must use algorithms such as focused sampling to learn from their environment (Billman and Heit [1988]).

2.1 GENERAL VERSUS SPECIFIC EXPERTISE

Bonner and Lewis indicate that they are looking at specific expertise, which they identify by task-related experience, training, and knowledge, as opposed to general auditing expertise. Differences between the two kinds of expertise are important since the use of experience to measure expertise appears more appropriate in examining general, rather than specific, auditing expertise. The reason is that the nature and type of experience are more important for developing task-specific expertise than for developing general auditing expertise. In a multifaceted domain such as auditing, the auditor is more likely to gain expertise in a specific subdomain without gaining exposure to auditing in general.2

1 Strategies are goal-directed processes which may involve automatic nonstrategic operations but which are potentially available to consciousness (Schneider and Weinert [1990, p. 289]).
2 People often develop expertise in very specific domains. For example, Chi and Koeske [1983] describe a child who was an expert on one specific class of dinosaurs but a novice in regard to other classes.
In a study of political expertise, McGraw and Pinney [1990] outline three possible relations that could exist between general and specific expertise: (1) positive (development of specific expertise across more than one task leads the individual to develop general expertise); (2) negative (using scarce cognitive resources in developing specific expertise reduces the individual's ability to develop general expertise); or (3) independent (expertise in a specific task is unrelated to the level of general expertise). According to McGraw and Pinney, the available evidence suggests that specific and general expertise operate independently; the former is likely to be developed by direct experience, whereas the latter is likely to be developed from education, training, and other indirect sources. The influence of how expertise is developed on the future use of knowledge is greater when that knowledge is gained through direct experience rather than by some other means.

General expertise and specific expertise can also be differentiated on the basis of how often they are used. Specific expertise is used only in the contexts that match the task, while general expertise is likely to be invoked constantly in response to problems for which no proceduralized knowledge exists. This difference in use should lead to processing differences; the general expert will have well-developed strategies for organizing and representing knowledge, while the task-specific expert has proceduralized knowledge to deal with specific tasks.

McGraw and Pinney also suggest that general and task-specific expertise have unique effects on a number of information-processing variables in the political realm which affect the political experts' performance. While the auditing domain might involve a completely different relation between general and specific expertise, these results suggest the need for care in specifying the type of expertise when investigating the relation of expertise to memory or judgment factors. The auditing expertise literature has not yet addressed this point, with some studies examining general expertise (Frederick [1989]) and some specific expertise (Bonner [1990]). The relation between experience and expertise depends on the notion of expertise being used. Different types of expertise will have different cognitive characteristics.

2.2 COSTS AND BENEFITS OF EXPERTISE

To support the notion that experience may not be a good operational measure of expertise, Bonner and Lewis identify the mixed results obtained in auditing research with respect to the relation between experience and performance. Their discussion is based on the idea that the benefit of expertise is superior performance and that experts will always outperform novices. However, as Adelson [1984] and Frensch and Sternberg [1989] point out, the development of a large structured knowledge base which enables the expert to exhibit superior performance also has associated costs which may on occasion lead to situations where novices outperform experts. Thus, in situations where novices outperform experts, it is possible that novices are expert at a certain task, as Bonner and Lewis suggest, or that some characteristic of the task interferes with experts' performance.

---

3 Many of the studies cited by Bonner and Lewis were pointed out by conference participants to be studies of the relation of experience to performance. These studies are not dealing with expertise per se.
2.3 THE RELATION OF KNOWLEDGE AND ABILITY

Bonner and Lewis present evidence from their review of the literature that ability does interact with knowledge and that ability does not interact with knowledge in terms of performance. They suggest that the ambiguity of this relation justifies the inclusion of the ability variable in their study. However, the very nature of the ability variable is brought into question by recent evidence that high-ability individuals have knowledge in many domains, whereas low-ability individuals have knowledge in only a few domains (Sternberg and Wagner [1986]). Since standard psychometric tests sample knowledge from many domains, it is not surprising that low-ability individuals who are not familiar with these domains process information less efficiently and effectively than high-ability individuals.

Because of this measurement error such tests underestimate the skills of low-ability individuals in the domains with which they are familiar. Schneider, Korkel, and Weiner [1989] studied the interaction of ability and knowledge in a series of studies investigating soccer expertise. They found that low-ability experts outperformed high-ability novices on all performance measures. Additionally, high- and low-ability experts could not be differentiated from each other, although they definitely outperformed novices. These results suggest that when task-specific knowledge is required, measurement of general ability will not be a good predictor of performance. This relation reflects a more complex interaction of ability and knowledge in auditing expertise than suggested by Bonner and Lewis's hypothesis of a main effect for ability.

3. How Can an Auditing Expert Be Identified?

The standard paradigm for expertise research is the contrastive method, which consists of identifying individuals who differ on a basic characteristic, measured by test scores, training and experience, or some other criterion. Once individuals are defined according to this basic characteristic, an experiment is conducted to determine whether these individuals differ with regard to performance on a comparison task. Bonner and Lewis use knowledge and ability tests and experience to separate individuals into expert and novice groups, and then select four auditing tasks to be used as the comparison tasks.

3.1 SELECTION OF COMPARISON TASK

Conference participants raised many design questions related to the selection of the auditing tasks and the breakdown of knowledge and ability required to solve them. Specifically, there appears to be no systematic variation in terms of knowledge and abilities across the tasks nor do the tasks obviously vary in degree of complexity, although the results suggest some differences in degree of difficulty across tasks. In addition, no clear connection between these four tasks and the concept of the auditing expert is provided. Are these tasks in combination representative of the auditing experts repertoire? Conference participants indicated that task 1 in particular is usually done with the assistance of a decision aid; such tasks may therefore routinely be made online rather than by memory. As Park and Hastie [1986] indicate, on-line decisions involve a different set of information-processing strategies than do memory-based decisions. This distinction has important consequences for the study of expertise, since forcing subjects to do by memory tasks typically done on-line could attenuate the differences in performance between novices and experts.
3.2 IDENTIFYING THE BASIC CHARACTERISTIC

To identify the basic knowledge and ability factors required for performance in the four auditing tasks, Bonner and Lewis used a combination of tests, self-reports, and specific experience measures. They indicate that they did not validate the knowledge and ability tests, although certain parts of the tests were extensively pretested. Voss et al. [1986] in their review of the contrastive approach to examining individual differences, identify the serious consequences of improperly or inadequately validating test materials used to measure a basic characteristic. Conference participants noted that a failure to validate leaves unanswered such questions as whether the items in the test measured the same factor (i.e., general business knowledge) and whether the test scores behave as one might predict. For example, are the higher scores by students on the ability test consistent with how one might predict such test scores would behave? Another measurement problem faced by psychometric studies of expertise such as Bonner and Lewis is that each characteristic measured (such as general domain knowledge) is measured in several ways, including test scores. Profile analysis might prove useful in allowing evaluation of all the measures of a characteristic and would provide the basis for measuring the reliability of the profile represented by the various measures.

3.3 DATA COLLECTION AND ANALYSIS

During the conference discussion, a number of issues were raised with respect to data collection. Since independent raters were not used by Bonner and Lewis to code the open-ended data, no measure of coding reliability was provided. With respect to both the knowledge tests and the auditing tasks, conference participants questioned whether the implicit weighting scheme in the grading was consistent with the scheme an expert might apply. For example, identifying the right financial statement errors but selecting the wrong auditing procedures might constitute better performance than identifying the wrong financial statement errors but selecting the right auditing procedures. This seems particularly significant when examining table 1 of Bonner and Lewis, which shows that for three of the four tasks performance was, at best, barely above the 50% mark. Again, this raises the question, when is performance expert performance?

Bonner and Lewis analyzed the raw scores of the knowledge and ability tests. No weights were assigned for different knowledge, yet task demands implicitly will lead to differential weighting of the knowledge items. Raw test scores have many disadvantages, including difficulties with interpretation, problems with meeting normality requirements for analysis, and lack of comparability across tests. One method for overcoming these data problems is to transform the data into percentiles, standardized scores, or normalized scores. Adjustments may also be calculated for guessing and omissions to improve the comparability and consistency of the scores. These procedures for fine-tuning the scaling and scoring procedures used in development of the data would aid the interpretability of the findings and comparability across studies.

Additional conference discussion concerned the interpretation of table 2, which reports the proportion of "experts" for each task. Expert performance is defined as that score which represents a complete and correct response as determined by the authors. The arbitrary nature of the cutoff makes comparability across tasks impossible. Also, the apparent differences in the overall difficulty of the tasks and in the degree to which expertise is required by the task further complicate the interpretation of the table.

If job performance and audit problem solving are related, then the significant difference in general problem-solving ability between senior managers and audit seniors confirms that survivorship is based on superior ability. On the other hand, if job performance and audit problem
solving are not highly correlated, then a significant ability difference might be the result of a poorly developed test. The fact that students, who represent the population from which auditors are drawn, have higher ability scores than seniors is consistent with the second conjecture. The other issue raised with respect to ability is the question of why there is no interaction between ability (a surrogate for general cognitive strategies) and knowledge reported by Bonner and Lewis, since these interactions would seem more likely to predict performance than an ability main effect.

Conference participants found puzzling the range of performance in terms of raw scores across the four auditing tasks. On task 1 nearly everybody is an expert, while in the other tasks many subjects' levels of performance are low. Is the differential performance on these tasks due to variations in task difficulty or to variations in the degree to which expertise is required to perform the task? The lack of comparability across tasks makes it difficult to assess which explanation is appropriate. Analysis of these across-task differences in the performance of experts may reveal much about the nature of auditor expertise. The data collected by Bonner and Lewis could be used to examine these differences through multivariate analysis.

The research question of Bonner and Lewis is motivated by the presumption that experience-based classification of auditors into expert and novice groups leads to classification error. However, the results reported in tables 3 through 6 do not address this question. A direct approach to measuring classification error would involve comparing the amount of error when experience is used as the only basis for classification and when a combination of experience, knowledge, and ability is used for classification. Discriminant analysis or some other classification procedure would be an appropriate technique for this analysis.

REFERENCES

PARK, B, AND R. HASTIE. "The Relationship Between Memory and Judgment Depends on Whether the Judgment Task Is Memory Based or On-Line," Psychological Review 1-July 1986): 258-68,

