11-1-1998

Inter-Auditor variability in expert-like task behavior: A model and an application.

Joseph A. Russo Jr

CENTER FOR APPLIED RESEARCH

MONOGRAPH

November 1998

Inter-Auditor Variability In Expert-Like Task Behavior: A Model and an Application

by

Joseph A. Russo, Jr., Ph.D.
Professor of Accounting
Lubin School of Business
Pace University

THE LUBIN SCHOOL OF BUSINESS

PACE UNIVERSITY

INTER-AUDITOR VARIABILITY IN EXPERT-LIKE TASK BEHAVIOR: A MODEL AND AN APPLICATION

by

Joseph A. Russo, Jr., Ph.D.

Joseph A. Russo, Jr., is Professor of Accounting, Lubin School of Business, Pace University.

SUMMARY

This paper is a process-based investigation of the sources of variation in expert-like task behavior between individuals who would otherwise be considered as members of a homogenous group. Auditor expertise is explored by means of a model based on behavior automaticity, a widely recognized indicator of expertise in the performance of a task. The model's three major features are: first, quantitative measures of the effects on task automaticity due to differences between auditors in the accessibility of their knowledge, its rate of utilization in a task, and the particular mix of behaviors that characterize their task strategies; second, a hierarchical presentation of the underlying phenomena that determine how relative expertise in a task is perceived that, in classic reductionist fashion, explains higher level perceptions in terms of lower level phenomena, patterns among which can lead to a richer characterization of relative expertise than is permitted by use of the unidimensional criterion of behavior automaticity; third, identification of specific aspects of the observer-observed system through which assessments of behavior are made, together with guidance for quantifying and incorporating them into expertise evaluations. The submersion of information through aggregation in assessments based on norms and higher level phenomena; strong, weak, and non-complementary phenomena; and the assurances and cautions that are associated with each of these are discussed and illustrated.

To illustrate the application of the model, behavior observation and think-aloud verbal protocol methodologies are employed to collect data on the observable behavior and cognitive processes of four first-year staff auditors during performance of simulated field tasks. The findings reported in this paper show that it is possible, perhaps even likely, that auditors displaying the same degree of expert-like task behavior are actually quite different in terms of the underlying properties of knowledge accessibility and the expression of that knowledge as task strategy. Differences in these properties may offset at higher levels of perception, hiding them from an observer's view. However, consideration of underlying differences have potentially important implications for the careers and professional development of the individuals involved. Finally, the model indicates, and the findings illustrate, that the naive choice of typical metrics for use in evaluations of individual auditor performances submerges information that may be of significance for both the individual and the organization whose success depends to some extent on that individual's performance. This danger exists even if a salient standard is available, but is especially important to consider when individuals are evaluated one against another. The model shows that in the latter instances, choice of the reference auditor is not a significant consideration for most strongly complementary phenomena (*viz*, task automaticity, knowledge accessibility and its subordinate phenomena), but can have a potentially significant impact on evaluations involving weakly complementary and non-complementary effects (primarily strategy-dependent effects.) These findings argue for placing greater emphasis on individuals rather than groups in future expertise research in auditing.

INTRODUCTION

What accounts for our varying perceptions of expertise while observing others perform a task? Is it possible to distinguish varying forms and degrees of expertise among auditors who otherwise display the same level of expert-like behavior? Finally, to what extent are assessments of the relative expertise of two or more auditors affected by differences in the properties[1] of the assessors and the circumstances surrounding the need for such assessments to be made? These are examples of questions that raise issues of importance to managers, supervisors, instructors, and others who find it necessary to assess the quality and potential of those in their charge based on observations of those individuals' behaviors while performing tasks in the field.

This paper is a process-based investigation of the sources of variation in assessments of expertise made by observers of the behaviors of individuals who otherwise would be considered as members of a homogenous group. A model is proposed for use as a basis for quantitative examination of some of the issues raised by the preceding questions. The model takes cognizance of the fact that any assessment process inextricably binds observer and observed, so that in certain instances, there appears to be an arbitrariness to the process. It makes salient those instances in which arbitrariness may be of concern, and identifies from among them those that are most in need of consideration when comparing assessments made by different observers or by the same observer at different times. The model's application is illustrated by means of an analysis of the observed behaviors of four first-year auditors who performed audit-related tasks in simulated audit environments.

The Limitations of "Typicality"

A substantial portion of behavioral research in auditing attempts to ascertain and compare the properties of "typical" members of groups (e.g., "novice auditors" vs. "expert auditors"). To the extent that individuals are members of a homogenous group, the observation of the one is as good as the mean of observations of the group. To the extent that there exists great variability among individuals, then to that extent, the mean of observations of the group is uninformative regarding any particular individual. However, variability in expertise or expert behavior within what are, by conventional criteria, taken to be homogenous groups of auditors has not been systematically studied by behavioral researchers in auditing.[2]

---

1. The term "properties" is used in this paper to denote a set of characteristics, such as choices, modes of behavior, and interactions by which one person, entity, phenomenon, etc. is differentiated from another.

2. Research employing very small numbers of non-randomly selected individuals (usually verbal protocol studies in which subjects are volunteers) can be loosely interpreted as individual studies. For example, in an early study of this type, Biggs & Mock (1983) compared the strategies of two auditors having two years of experience with that of two auditors having three years experience. Their paper presents a detailed tabulation of each subject's performance. Two basically different strategies are identified among these subjects: *directed strategy*, in which an auditor searches

In contrast with what is the usual case in the laboratory, in field situations, during performance of a task, one confronts individuals rather than a group. Our assessments of behaviors are based on observations made of individuals and comparisons across individuals, usually observed two at a time.[3] If individual treatment is based on "typical" assessments, then the potential for sub-optimal outcomes, serious errors, injustices, or other deleterious consequences is present. In these situations, there is a greater need to know more about the properties of the individual than about those of the group. This paper begins the examination of this and other issues raised by the opening questions.

The Nature of Expertise

Expertise is a quality attributed to others whose behaviors are observed during performance of a task (Russo 1997b). The nature of expertise is widely discussed in the behavioral research literature. Generally, this discussion appears to revolve around two central issues. First, different forms of expertise exist, some better suited than others for use in specific contexts. Second, various conceptualizations of expertise emphasize either processes or outcomes. In addition, there appears to be some confusion among researchers between the *what* of this research and the *how* (See Russo 1997b for a complete discussion). Setting these matters aside, this paper focuses on process rather

---

for information only on an as-needed basis, and *systematic strategy*, in which an auditor reviews and attempts to digest all available information before attacking the specific needs of the problem. Biggs and Mock then associate the individuals in their tabulation with the specific strategic pattern each individual displayed. In a later study, Biggs, et al (1988) compared the think-aloud verbal protocols of two audit managers and two audit seniors while performing an analytical review task. They note that, based on an analysis of their information acquisition activities, the seniors were primarily concerned with obtaining information related to the revision of audit programs while managers were more focused on understanding the client and its business, possibly for the purpose of preparing comments for the management letter. On a more general level, Newell and Simon (1972), in their ground-breaking study of human problem solving, advocated flow charting the decision processes of subjects for the purposes of gaining insight into their overall goal-directed processes. However, none of this, and related research, is concerned with a systematic study of specific individuals as actual or potential targets of action. Rather, these studies represent the study of individuals for the purpose of gaining insights into processes that may be present in broader populations. It should also be noted that the research previously cited, and other research like it, focus on cognitive processing given a clearly defined decision-making task rather than the broad process of behavior while preforming tasks in the field (see Russo 1997b).

3. In this regard, there is no need to become bogged down in any debate concerning how many individuals constitute a group, as this will not further the purpose of this research. Suffice it to say that this research is concerned with those situations in which one observes two or more auditors for the purpose of making an assessment of one vs. another. "Typical" measures are of value only in situations in which an observer requires normative data in order to form an expectation. For example, the answer to the question "Does X behave as would a novice auditor in this situation?" requires normative data. In contrast with this example, the issues of concern in situations contemplated in this paper are more correctly represented by questions such as "Which of the two auditors performing this task appears to be the more expert?" and "In what ways do these auditors differ in their apparent expertise?" The only information required in these situations is a criterion (or set of criteria) and a means for mapping observations on that criterion (or each of them) to a relative scale of expertise.

than outcome and, in particular, on the automaticity of task behaviors and what may be inferred about unobservable processes and circumstances given what is observed. In so doing, two basic and related assertions are made for which I believe there exists wide-spread agreement. First, expert behavior is knowledge-driven (see Russo 1998: 12-13, note 23). Second, the attribution of expertise to an individual implies that that individual possesses qualities of knowledge and behavior that differ significantly in terms of one's expectations from those of individuals to whom expertise is not attributed.

The Perception of Expertise

It is generally acknowledged that the degree of automaticity, a term applied to sequences of observable task behaviors that are performed without cognitive mediation, is a distinguishing feature of expert task behavior. Alba and Hutchinson (1987) cite extensive research attesting to the common observation that as one gains familiarity and expertise in a task, one's behaviors become more automatic. That is, with growing expertise, tasks are performed with diminishing effort and *without conscious control.* The increasing automaticity of behavior with experience is a central concept in artificial intelligence and learning theory (e.g., Anderson 1982, 1987; Mayer 1992: 305). Davis & Solomon (1989) employ the term "expert" to describe one whose behavior during performance of a task displays a high degree of automaticity. Bedard (1989) notes that experts exhibit little self-insight into how their decisions are made because most are made subconsciously. The automaticity of task behaviors, therefore, is a significant factor in the perception of the quality of expertise. However, in using automaticity as a criterion of expertise, neither the appropriateness of task behaviors nor the quality of task outcomes is considered. Therefore, the behavior studied can only be described as being relatively expert-*like.* In spite of this limitation, as will be illustrated later in this paper, automaticity of task behavior can serve as a productive focus for a comprehensive and internally consistent, quantitative model for investigating and understanding several significant properties of task expertise.[4]

4. Ericsson & Simon (1983), in the context of think-aloud protocols, discuss omitted varbalization, which if present can overstate automaticity, as due to either (1) uncertainty, confusion, or an inability to evoke knowledge or (2) a rate of cognitive evocation that is too rapid for the verbal encoding process. Although the model they employ differs in some important respects from that used in this research (theirs is a data-processing-oriented model, whereas the research reported in this paper is based on current thinking about brain structure and function, and concepts of evolution and consciousness), none of these circumstances is necessarily inconsistent with the concepts of automaticity and availability of knowledge used in the present research. The major point of difference with Ericsson & Simon lies in their implicit assumption that all mental processing is cognitive and that the failure to verbalize is a limitation inherent in the processing rate of the Human Information Processing System (HIPS), their hypothesized "mental computer." In contrast, the model and methodology on which the present research is based posits there can be no so-called "failures" to verbalize because no cognition in fact takes place. Sub-conscious mental processes are not accessible for purposes of expression. Their existence during performance of a task is an inference made by observers based on the presence of non-cognitively mediated transitions between observed behaviors.

Model Foundations

Russo (1994, 1995a) has proposed a model and methodology that permits inferences to be made about unobservable processes responsible for the automaticity of task behaviors as auditors perform simulated field tasks, and by extension, about a primary input to the formation of observers' perceptions of auditor task expertise. Employing this model and methodology, Russo (1996) reports findings of significant associations between auditors' perceptions of certain characteristics of their task environments, the states of their knowledge bases, and the kinds of cognitions mediating their task behaviors. He also reports evidence strongly suggesting wide inter-auditor variation in both learning and the mode of learning resulting from experience during performance of a task (Russo 1998). This paper builds on this foundation by (1) examining the degree to which variations in auditor knowledge and strategic behavior contribute to variability in perceptions of inter-auditor expert-like task behavior, (2) identifying and quantifying the properties of knowledge and strategy that account for differences between individual auditors in assessments of relative expert-like task behavior and that offer the potential for broader assessments of relative expertise, and (3) identifying and quantifying certain aspects of comparison processes that can contribute to variability in observer assessments of relative expertise in performance of a task.

MODEL

The Iterative Model of Task Behavior

In the field, audit tasks frequently are empirically intense. Such tasks are characterized by a requirement for significant domain and task knowledge, significant information input from the task environment, solution processes requiring significant interaction with the task environment, and outcomes that are represented by transformations of the task environment.[5] During performance of empirically intense tasks, behavior proceeds according to a process in which information is acquired from the task environment (perception behaviors), processed in memory (cognitive and non-cognitive mental processing), and the environment transformed in accordance with the outcome of that processing (execution behaviors).[6] In thus transforming a task environment, environmental stimuli upon which the selection of behaviors in succeeding iterations is based are altered. Iterations continue until the auditor perceives that a solution state exists (Russo 1995a, 1994).

5. Laboratory tasks, on the other hand, tend to be cognitively intense in that they lack the interactive solution processes that are characteristic of empirically intense tasks, and have intellectual commitments, (i.e., judgments or decisions) rather than transformed task environments as solutions. (See Russo 1997b: 7-8 for a discussion of these differences.)

6. Perception behaviors are all those observable behaviors and covert means by which stimuli are received from a task environment. In auditing, these behaviors typically are reading, listening, and observing. Execution behaviors are those whereby an auditor transforms a task environment. In auditing, these behaviors typically include requesting, writing, using mechanical devices such as calculators and telephones, searching, etc.

While perception and execution are observable behaviors, the various mental processes that mediate transitions between observable behaviors are not.[7] However, because knowledge is expressed through observable behavior, it is possible to attain some understanding of the properties of knowledge driving behavior through an analysis of the various forms of knowledge expression.[8] This is the approach taken in the research reported in this paper.

Task Automaticity

The perception of expert-like task behavior is based on the automaticity of behaviors observed during the performance of tasks. Variations across auditors can be due to inter-auditor differences in accessible knowledge and the manner in which that knowledge is expressed.

Transitions between observable behaviors during performance of a task are mediated by episodes of either unreportable (subconscious) mental processes; cognitive processes, which under suitable conditions can be rendered reportable; or a combination of both these processes.[9] The nature

7. All perception behaviors ultimately involve reception of sensory stimuli from the task environment. Such stimuli may be received by overt means, as by reading, attentive listening, and observation. However, most stimuli are received as "covert perceptions," a term denoting passive, unattended reception and interpretation of stimuli. It is the covert reception of stimuli that is primarily and ultimately responsible for the moment-to-moment perception of our environment, our place within it (e.g., the context of behavior in all its dimensions: temporal, physical, social, etc.), and our actions. Were it not for covert perception, the brain would be overwhelmed with the amount of information it would have to retain and process. By substituting continuous, passive input from the environment for actively maintained environmental information, interpreted subconsciously and "on the fly," the brain frees its limited resources for more complex and/or critical processes, thereby gaining an evolutionary advantage. Because of its critical role in establishing context, covert perception can be said to be the primary contributor to what is commonly called short-term memory. See Hobson (1994, esp. Ch 7) for related discussion.

8. It is highly questionable to denominate that which cannot be expressed as "knowledge," or, indeed, to attribute to it any existence at all. Knowledge is information put into action; it is "process," not "thing." Action, in this context, is the production of environmental change brought about by the actor's internal metabolic processes (referred to hereafter as an autonomous process.) Under this concept, arguably, molecules of chemical reagents contain information and chemical reactions display knowledge. Does a stone possess any knowledge? While *we* may extract information from a stone (e.g., its age, chemical composition, structure, all of which may be informative about geologic history), a stone is incapable of any autonomous mode of expression. Hence, again arguably, a stone has no knowledge. A plant will alter its behavior in response to sensory input from its environment (e.g., turn its light receptors toward the light source, thereby maximizing energy input), and to that extent, we may attribute knowledge to a plant. Insects, fish, dogs, apes, and humans all exhibit appropriate responses to environmental inputs, albeit to varying degrees, and for this reason they may be said to evidence the possession of knowledge.

9. An unobservable process is reportable if the experience of it can be expressed verbally. An observable process is reported when it occurs.

of processes underlying a mediating episode is reflected in the knowledge base[10] response (r_k), the kind of process being indicated in the model by the subscript k ($k \in \{1,2,3\}$). When the process mediating a transition between one observable behavior and the next is unreportable, then the transition is said to have occurred automatically, indicated by a response of type r_1. Cognition, however, while a reportable phenomenon, is not normally observable. If present, cognition reveals the process by which knowledge that is not evoked automatically is accessed, i.e., placed into a state in which it can influence subsequent behavior.[11] If the sought knowledge is evoked as a result of cognitive effort indicating a search of memory (e.g., descriptive, nomological, teleological, or normative expressions of thought), the knowledge base response is of type r_2. On the other hand, if the knowledge base is unable to supply sought knowledge (e.g., cognitions expressing uncertainty or confusion), then the knowledge base response is of type r_3.

To the extent that automatic processes mediate transitions from one behavior to the next in a solution sequence,[12] then to that extent, the perceived automaticity of that task behavior increases.[13] To assess this and other attributes of an auditor's knowledge over any extended period of behavior observation, various linear combinations of the sums of knowledge base responses by type for each

10. In this research, the term "knowledge base" is used to represent the totality of the knowledge that an auditor called upon during performance of a task. While it is common usage to refer to the collection of that which drives an auditor's behavior as a "knowledge base," this usage has several serious shortcomings. Knowledge is information in action. The "content" of an auditor's knowledge base can only be ascertained by observing the auditor's response to situations making knowledge demands (Russo 1997a: 411). What is commonly referred to as "knowledge" is actually an action potential that is latent until called upon to initiate and guide behavior. This latent capacity is better referred to as "information" or perhaps "data" rather than "knowledge." Although not explicit in the putative usage of the terms "knowledge base" and "knowledge," I believe this distinction is on some level generally understood. Hence, in deference to common usage, I will continue to use these terms in contexts in which neglect of the precise distinctions mentioned are not troublesome.

11. Cognition, commonly referred to as "thinking," is our experience of a process of sublimed learned behavior. Cognition is not a naturally occurring human phenomenon, but develops out of the self-directed speech observed in children, a behavior that remains vestigially present in adults during moments of great surprise, confusion, or stress. Sublimation of this behavior is brought about by its association with contexts in which it is learned that overt verbalization is neither practiced nor necessary in order to achieve access to unavailable knowledge. For additional discussion, see Reber (1985), Ellis & Siegler (1994: 343-4), Dennett (1991), and Russo (1998, esp. note 15).

12. The term "solution sequence" applies to an auditor's chronologically ordered set of observable actions and reportable cognitive activities during performance of a task.

13. Although subconscious mental processes do not occur instantaneously and, therefore, could potentially occupy as much or more time than some cognitive processes, Russo (1995b) has shown that over a task, the duration of mediating episodes is positively related to the number of cognitions composing the episode (i.e., to the cognitive complexity of the episode.) Later in this paper, cognitive complexity will be related to knowledge-base organization.

category of behavior, $j \in \{1,2,..b\}$, are used. These combinations are denominated by the symbol n_{hj}, $h \in \{s,c,e\}$.[14]

The Determinants of Task Automaticity

The automaticity of task behavior, a, can be measured by the ratio of all automatic knowledge base responses, n_{ej}, to the total instances of task behaviors, n. Symbolically:

$$a \equiv \frac{1}{n} \sum^{j} n_{ej} \tag{1}$$

Let a_j be the automaticity of behavior j, $a_j \equiv n_{ej}/n_j$, and m_j the mix of each behavior, $m_j \equiv n_j/n$. Equation (1) can then be written in terms of behavior automaticity and behavior mix as:

$$a = \sum^{j} a_j m_j \tag{2}$$

Where there are several different kinds of behaviors (i.e., $b > 1$), then automaticity conceptually applies to the mean automaticity of the all behaviors, the weights applied being the mix of each behavior, as shown by equation (2).

Now consider two auditors who are observed performing a task. When making comparisons, one auditor serves as the subject and the other as the base or reference. Let all quantities that apply to the subject auditor be designated by the superscript X, and those that apply to the reference auditor by the superscript N. Any perceptions of their relative expertise in the task will be based on the relative automaticities with which they perform their tasks. The auditor whose task automaticity is greatest will be perceived to be the more expert-like of the two. Let E be the ratio of Auditor X's automaticity to Auditor N's.[15] Then, applying equation (2) to this definition, we may write:

$$E \equiv \frac{a^X}{a^N} = \frac{\sum^{j} a_j^X m_j^X}{\sum^{j} a_j^N m_j^N} \tag{3}$$

14. During any extended period of observation, there occur n mediating episodes. Each episode, $i \in \{1,2,...n\}$, terminates in a behavior of type j, the *target behavior*, and includes at least one response, r_{ikj}. The combinations used in the model are: $n_{sj} = \Sigma^i r_{i1j} + \Sigma^i r_{i2j} + \Sigma^i r_{i3j}$, $n_{cj} = \Sigma^i r_{i1j} + \Sigma^i r_{i2j}$, and $n_{ej} = \Sigma^i r_{i1j}$.

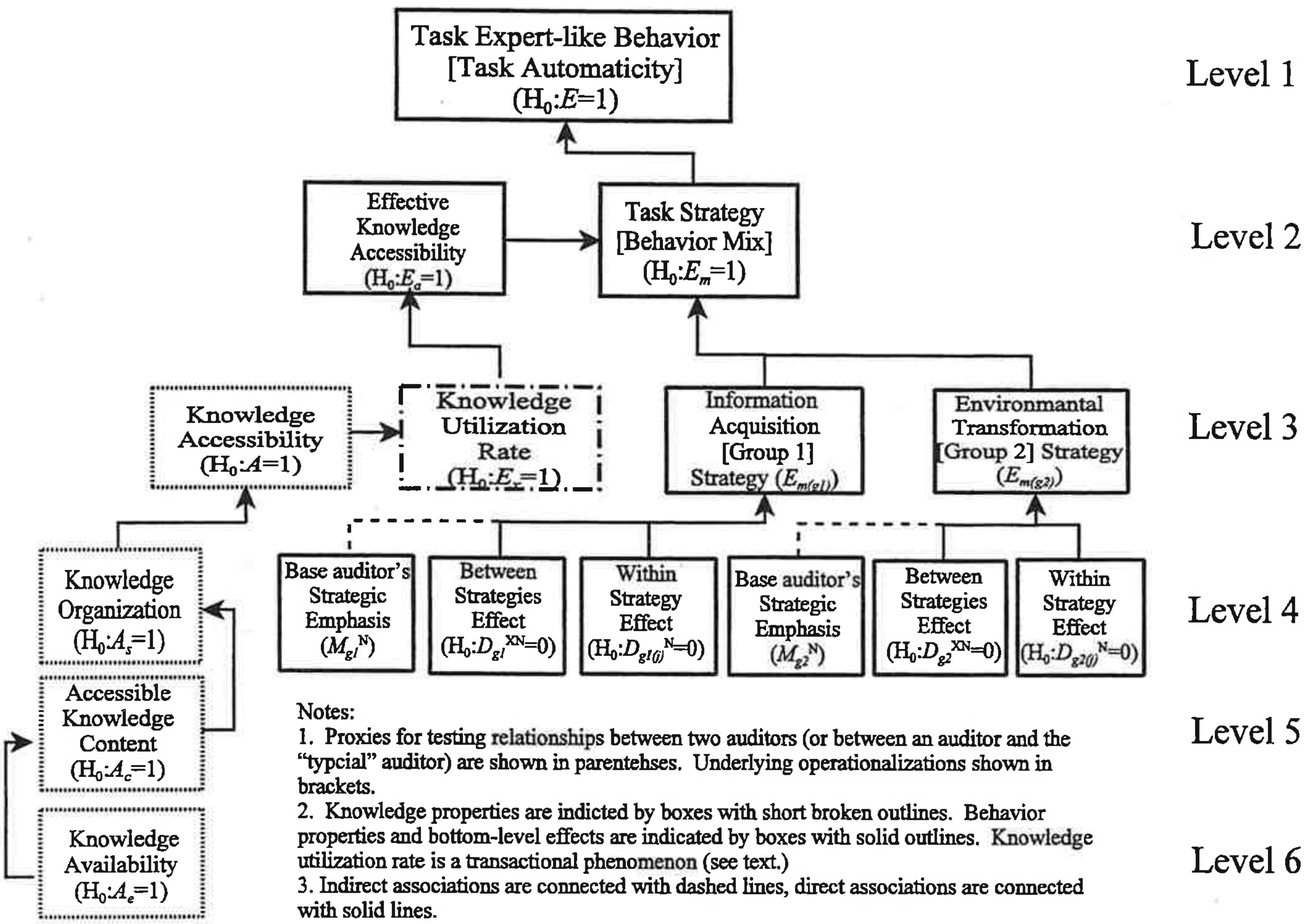
15. In the interests of notational simplicity, superscripts are omitted unless essential to clarity and precision of communication. For example, the precise representation of the independent variable in equation (3) is E^{XN}.

Equation (3) shows that relative task automaticity of auditors X and N is determined by two circumstances of their task situation: the automaticity of each auditor's individual behaviors (a reflection of the accessibility of knowledge, discussed later) and the relative intensities with which each auditor employs various kinds of behaviors during performance of a task (a reflection of strategic knowledge, also discussed later.)[16] Given this result, it is now possible to address questions as to what accounts for the difference between these auditors in observed expert-like behavior (i.e., $E \neq 1$) and whether the absence of a difference is sufficient evidence for concluding that both auditors share a common base of knowledge. Figure 1 provides, in graphic form, an outline of the argument to follow.

16. If purely cognitive activity is excluded (e.g., mental arithmetic, etc.), observed task behavior may be considered *prima facie* the equivalent of Anderson's (1987) concept of "compiled procedural knowledge." However, to make Anderson's concept operational, it is necessary to recognize that it must be more inclusive than simply knowledge of what actions to take, in what sequence they are to be taken, and under what conditions they are to be performed. Procedural knowledge presupposes knowledge of entities, relationships, and processes in the actor's external task environment, since this knowledge is necessary to instantiate any observable action before it can be performed. That is, procedures are purposefully and intentionally directed by the actor at, and, in some way, affect, *something specific and identifiable* in the task environment.

FIGURE 1

HIERARCHY OF PHENOMENA AFFECTING OBSERVER PERCEPTIONS
OF RELATIVE EXPERT-LIKE TASK BEHAVIOR



Notes:
1. Proxies for testing relationships between two auditors (or between an auditor and the "typcial" auditor) are shown in parentehses. Underlying operationalizations shown in brackets.
2. Knowledge properties are indicted by boxes with short broken outlines. Behavior properties and bottom-level effects are indicated by boxes with solid outlines. Knowledge utilization rate is a transactional phenomenon (see text.)
3. Indirect associations are connected with dashed lines, direct associations are connected with solid lines.

Hierarchial Structure

Figure 1 presents a hierarchy of phenomena,[17] each contributing to the assessment of relative task automaticity, and by extension, task expertise.

Levels

With regard to the hierarchical structure itself, Level 1 is the only phenomenon that is generally observable under normal field conditions. Level 2 consists of the two major determinants of task automaticity implied by equation (3). Successively lower levels show the generally unobservable phenomena which contribute to, and can potentially affect, assessments of task automaticity across auditors, tasks, time, or observers. In this paper, the focus is primarily on variations across auditors. The labeling of effects as "bottom-level" indicates only that these effects are not explained by the model in terms of lower level phenomena. Use of this term is not meant to imply that such explanations are not possible or that they may not be provided by further model development.

Functional Relationships

Two kinds of functional relationships are represented in Figure 1. Any arrow that departs from the top of a box, moves vertically to the next higher level, and enters the bottom of a box is equivalent to a mathematical equality (e.g., an "=" sign). The independent variables are named in the boxes at the arrow tails, and the dependent variable is named in the box at the arrowhead. Any arrows entering or leaving a box by the left or right side indicate multiplicative relationships. Additive relationships are indicated by the confluence of arrows representing the independent variables. An arrow represented by a broken line indicates a relationship that is not relevant to the kind of analysis illustrated in this paper but is nevertheless an integral part of the model.

Terminology

Properties and Effects

The model consists of properties and effects. A *property* is an inherent characteristic by which one distinguishes among phenomena, while an *effect* is a change in, or relationship among, the properties of a phenomenon that can be attributed to underlying (more basic, lower level) phenomena. In Figure 1, properties of knowledge are shown in boxes with broken outlines; properties of behaviors and bottom-level effects are shown in boxes with solid outlines. The box

---

17. The term *phenomenon* is used in this research to represent any entity, process, or relationship to which one can attach an identifying symbol, such as a name. For example, automaticity, accessibility, auditor, system of internal control, materiality, etc., are all within the purview of the term.

labeled "Knowledge Utilization Rate" is somewhat unique in that it is a transactional phenomenon by which the property of knowledge accessibility, which is strategy independent, is made effective through strategic expression in a particular task. Effects are either ratios (variables with A or E symbols) or differences (variables with D symbols) in specific properties of the auditors and behaviors being compared. The terms "*model metric*" or "*metric*" are used in this paper to represent properties and effects when a statement refers, or applies with equal validity, to both.

Knowledge Accessibility vs. Behavior Automaticity

The term *accessibility* applies to the ability of an auditor's knowledge base to supply the information demanded by the behaviors employed during performance of a task. Accessibility, therefore, is a property of a knowledge base rather than of behavior. The term *automaticity*, as shown above, relates to the degree to which task behaviors are performed without cognitive mediation. Automaticity, therefore, is a property of behavior rather than knowledge. That is, while "automaticity" relates to a quality of observable behavior, the term "accessibility" relates to the responsiveness of a knowledge base, given a demand for information. As will be discussed in the next section, the automaticity of behavior is actually a reflection of more fundamental knowledge-base properties that together define the accessibility of the information in a knowledge base. Because neither the accessibility of knowledge nor the automaticity of behavior can be assessed except when behavior is observed, what is observed is a jointly determined phenomenon, and for this reason, the two concepts are easily confused. However, for analytical purposes, accessibility and automaticity each convey different information about auditors and their expertise. Consequently, it is important that they be clearly distinguished and treated separately.

To illustrate how these concepts jointly determine task automaticity, consider again equations (1), (2), and (3). Equation (1) describes automaticity because it takes no explicit account of the properties of the knowledge base that make it possible for n_e of n task behaviors to take place without cognitive mediation. In equation (2), on the other hand, the parameter a_j is a property of the knowledge required to support a specific type of behavior, j, **whenever that behavior is performed**. The phrase in bold type is key to the distinction between accessibility of knowledge on the one hand, and automaticity of behavior on the other. The behavior mix, m_j, informs about the proportion of times during performance of a task that a demand was made of the knowledge base for the particular knowledge that behavior requires. However, that demand is independent of the accessibility of the knowledge it elicits when it is made. As shown by equations (2) and (3), the more assessable the knowledge in a knowledge base, the greater the automaticity of task performance.

Decomposition of Relative Task Automaticity (E)

In the discussion of Equation (3), it was shown that task automaticity is determined by two circumstances of a task situation. The first of these, the automaticity of an auditor's individual behaviors, was described as being a reflection of knowledge accessibility, and the second, the intensity with which an auditor employs various kinds of behaviors during performance of the task, was described as a reflection of task strategy. The relationship of these circumstances to relative task automaticity is shown as Level 2 phenomena in Figure 1. In this and succeeding sections, these ideas are refined so that they can serve as a basis for further model development. I begin by defining the level 2 effects:

> *Effective knowledge accessibility effect* (E_a):[18] the measure, given by equation (4), of the effect on the base auditor's task automaticity of differences between auditors in the accessibility of information in their knowledge bases.

$$E_a \equiv \frac{\sum^j a_j^X m_j^N}{\sum^j a_j^N m_j^N} \tag{4}$$

> *Task behavior mix effect* (E_m): the measure of the effect on the base auditor's task automaticity of differences between auditors in the mix of behaviors used during performance of the task. In subsequent sections of this paper, the task behavior mix effect will be related to differences in strategic knowledge. Consequently, E_m will also be referred to as the *task strategy effect*. The task behavior mix effect is given by equation (5).

$$E_m \equiv \frac{\sum^j a_j^X m_j^X}{\sum^j a_j^X m_j^N} \tag{5}$$

Since the two effects above determine relative task automaticity, we have:

$$E = E_a E_m \tag{6}$$

which produces the same result as equation (3).

---

18. This terminology, while somewhat awkward, is nevertheless precise. The word "effective" is used in the sense of a net or resultant of competing forces rather than in the more common sense of a successful, efficient, or proper performance. The final word, "effect," is used in the sense of induced change or "impact upon," as previously defined.

Decomposition of the Effective Knowledge Accessibility Effect (E_a)

Inspection of equation (4) shows that the effective knowledge accessibility effect depends upon the relative accessibility of the knowledge possessed by auditors X and N, and demands made on that knowledge by the base auditor's behaviors. To separate these underlying components, let us define the *mean knowledge accessibility* of Auditor P's ($P \in \{X, N\}$) knowledge as:

$$c^P \equiv \frac{1}{b} \sum^j a_j^P \tag{7}$$

where b is the number of different types of recognized behaviors that can be observed. Then the *knowledge accessibility ratio*, which measures the relative accessibility of Auditor X's task knowledge compared with that of Auditor N is:

$$A = \frac{c^X}{c^N} = \frac{\sum^j a_j^X}{\sum^j a_j^N} \tag{8}$$

Regardless of its accessibility, it is how knowledge is used that determines the automaticity of task behavior. The *knowledge utilization rate* is defined as the ratio of behavior automaticity, $\sum^j a_j m_j$, to mean knowledge accessibility, $(1/b)\sum^j a_j$, given a specific mix of behaviors. This rate describes how knowledge accessibility, an inherent property of a knowledge base, is transformed into a measure of behavior automaticity, an inherent property of behavior, by the interaction of accessible knowledge with the particular mix of knowledge demands made by behaviors as a task is performed. When auditors employ behaviors that take advantage of their most accessible knowledge, their knowledge utilization rates increase. On the other hand, when auditors employ behaviors that demand knowledge that is not as accessible, their knowledge utilization rates fall. Since Auditor N serves in this model as the reference for making comparisons, the knowledge utilization rate of Auditor P, given Auditor N's behavior mix, is:

$$k^{PN} \equiv \frac{b \sum^j a_j^P m_j^N}{\sum^j a_j^P} \tag{9}$$

If Auditor N had the same knowledge base properties as has Auditor X, N's knowledge utilization rate, from equation (9), would be $\frac{b \sum^j a_j^X m_j^N}{\sum^j a_j^X}$. However, Auditor N's actual knowledge utilization rate is $\frac{b\sum^j a_j^N m_j^N}{\sum^j a_j^N}$. Therefore, the effect of the difference between the knowledge properties of auditors *X* and *N* on Auditor *N*'s behavior mix (i.e., their relative rates of knowledge utilization) is:

$$E_x = \frac{k^{XN}}{k^{NN}} = \frac{\sum^j a_j^X m_j^N \sum^j a_j^N}{\sum^j a_j^N m_j^N \sum^j a_j^X} \tag{10}$$

Combining the preceding results, we have:

$$E_a = AE_x = \frac{\sum^j a_j^X m_j^N}{\sum^j a_j^N m_j^N} \tag{11}$$

Decomposition of the Knowledge Accessibility Ratio (A)

The model posits that three lower level properties of a knowledge base determine its knowledge accessibility: organization, accessible content, and availability. These lower level properties are now defined.

> *Knowledge organization* is concerned with the amount of cognitive searching of a knowledge base that is required before an observable behavior is evoked. Knowledge organization is negatively related to the mean complexity of mediating episodes, i.e., the mean over all behavior types of the complexities of mediating episodes associated with each type of target behavior, i.e., $(1/b)\sum^j n_{sj}/n_j$ (Russo 1998: 10-12.) The *knowledge organization ratio* (A_s) compares the complexity of Auditor *X*'s mediating episodes with that of Auditor *N*.
>
> *Accessible knowledge content* refers to the capacity of a knowledge base to respond positively to a demand for information. A response is positive if a search of a knowledge base is successful, i.e., sought knowledge is evoked either by automatic processes or cognitive processes indicating that knowledge is present in the knowledge base.[19] Accessible knowledge content is positively related to the ratio of

19. The adjective "positive" applied to knowledge-base content is used to distinguish a state in which a knowledge base is able to supply the information demanded from a state in which there is a failure to meet the information demand (e.g., "negative knowledge content")

the mean of the frequencies of positive knowledge base responses associated with each target behavior to the mean complexity of mediating episodes, i.e., $(\sum^j n_{cj}/n_j)/(\sum^j n_{sj}/n_j)$ (Russo, 1998, esp. Tables 1 and 3). The *accessible knowledge content ratio* (A_c) compares Auditor *X*'s accessible knowledge content with that of Auditor *N*.

Available knowledge content refers to the capacity of a knowledge base to automatically supply information on demand. Available knowledge content is positively related to the ratio of the mean of the frequencies of automatic knowledge base responses associated with each target behavior to the mean of frequencies indicating positive knowledge content, i.e., $(\sum^j n_{ej}/n_j)/(\sum^j n_{cj}/n_j)$. The *available knowledge content ratio* (A_e) compares Auditor *X*'s available knowledge content with that of Auditor *N*.[20]

Symbolically, these definitions are presented as equations (12a), (12b), and (12c).

$$A_s \equiv \frac{\sum^j n_{sj}^X/n_j^X}{\sum^j n_{sj}^N/n_j^N} \ (a) \quad A_c \equiv \frac{\sum^j n_{cj}^X/n_j^X \sum^j n_{sj}^N/n_j^N}{\sum^j n_{cj}^N/n_j^N \sum^j n_{sj}^X/n_j^X} \ (b) \quad A_e \equiv \frac{\sum a_j^X \sum^j n_{cj}^N/n_j^N}{\sum a_j^N \sum^j n_{cj}^X/n_j^X} \ (c) \qquad \textbf{(12)}$$

where the subscripts qualify the variables as follows: *s* indicates all responses, cognitive and non-cognitive, *c* indicates positive knowledge base responses, and *e* indicates only automatic responses. The relationship of the components of (12) to the knowledge accessibility ratio is given by (13).

$$A = A_s A_c A_e \qquad \textbf{(13)}$$

The multiplicative relationships among the knowledge accessibility ratios and task strategy effects are indicated in Figure 1 by the in-line sequence leading to relative task automaticity. As can be inferred from the above definitions and equation (13), the properties of knowledge are hierarchically related and form a natural progression from most comprehensive (A_s) to most narrow (A_e). This progression is shown in Figure 1 by placing each property at hierarchically appropriate levels. Under a null hypothesis that both auditors possess a common knowledge base and task strategy, all of the preceding ratios will equate to unity.

---

20. Note the distinction between accessible knowledge content (n_c/n_s), available knowledge content (n_e/n_c) and automaticity (n_e/n) related to the *task* vs. accessible knowledge content (n_{cj}/n_{sj}), available knowledge content (n_{ej}/n_{cj}), and automaticity (n_{ej}/n_j) related to a specific *target behavior*. The subscript *j* qualifies a variable as relating to a specific behavior. In addition, the subscripts *s*, *c*, and *e* qualify a variable as a sum of specific kinds of knowledge base responses (see previous notes.)

An Operational Notion of Strategy

Strategy is a term that can be comprehended in either of two senses. In the *ex-ante* sense, strategy is a characterization of prospective behavior, while in the *ex-post* sense, it is a characterization of observed behavior. Since this paper is concerned with the analysis of observed behavior, I confine my discussion of strategy to the *ex-post* sense of the term. The model proposes that *strategy is an observer's (one's own) characterization of the purposes and intentionality of another's (one's own) behavior*. Thus, while the particular behaviors that take place, and their temporal sequence, exist independent of any observer,[21] the perception of strategy is an *observer-dependent phenomenon* in that it necessarily reflects the observer's own knowledge and observational skills.[22] While there is an *a priori* and context-free ordinal scale for mapping task

---

21. The observer here is an ordinary, fallible human observer, as distinguished from observations made by an omniscient being. Making the distinction between observer types is useful for certain theoretical and analytical purposes (See Russo, 1994: 89-90.)

22. The characterization of a sequence of behaviors as strategy has been described as a feat of insight. Newell and Simon (1972), in their study of human problem solving, advocate flow charting the decision processes of subjects for the purposes of gaining insight into their overall goal-directed processes. However, insight remains a distinctly personal phenomenon which, in any given instance and for any given observer, may never be acquired. Further, even if acquired, the insight gained may differ across individuals. For example, following Newell and Simon's suggestion, Biggs & Mock (1983) compare the behaviors of two auditors having two years of experience with that of two auditors having three years experience and identify two basically different strategies among these subjects: *directed strategy*, in which an auditor searches for information only on an as-needed bases, and *systematic strategy*, in which an auditor reviews and attempts to digest all available information before attacking the specific needs of the problem. In a later study, Biggs, et al (1988) compare the think-aloud verbal protocols of two audit managers and two audit seniors while performing an analytical review task. They note that, based on an analysis of their information acquisition activities, the seniors were primarily concerned with obtaining information related to the revision of audit programs while managers were more focused on understanding the client and its business, possibly for the purpose of preparing comments for the management letter. Although it is not the case in the two studies mentioned, might not *both* characterizations be applied with equal validity to the *same* set of data?

Strategic characterization requires insight on the part of the observer (see above.) Lonergain (1978), in the Preface to his extensive discussion of the topic, illustrates the individuality of insight in the following way:

> In the ideal detective story the reader is given all the clues yet fails to spot the criminal. He may avert to each clue as it arises. He needs no further clues to solve the mystery. Yet he can remain in the dark for the simple reason that reaching the solution is not the mere apprehension of any clue, not the mere memory of all, but a quite distinct activity of organizing intelligence that places the full set of clues in a unique, explanatory perspective. ...(Insight is) not any act of attention or advertence or memory but the supervening act of understanding. (x)

To this, I add: Whose understanding? Answer: The observer's. If the observer does not gain the connection of incoming cues to his/her other knowledge, that is, fails to perceive a unification of otherwise disconnected and diverse fragments of information, then there is no insight. That knowledge which insight unifies is uniquely the observer's.

automaticity to expertise (*viz*, greater automaticity implies greater expertise), no such scale exists for quantifying strategy.[23] Consequently, in this research, I focus instead on the cues from which an observer formulates the perception of strategy. For a given observer recognizing a given set of behaviors, the strategic characterization of behavior is determined primarily by the mix and sequence of the behaviors observed and recognized. Later in this paper, I stipulate and hold constant the observer and the set of behaviors recognized and focus mainly on behavior mix.

Strategy and Behavior Mix

It is clear from equations (5) and (6) that perception of the relative automaticity with which two auditors perform an empirically intense task is affected by the mix of behaviors employed by each auditor. The magnitude of this effect is measured most directly by the behavior mix ratio, E_m. This section develops the relationships necessary to account for and interpret this ratio.

One can gain a broad sense of an auditor's task strategy by considering the purpose and intentionality of the behaviors used during performance of a task. Assuming that auditors are purposeful and intentional in their choices of task behaviors, then the various behaviors that take place during performance of a task may be grouped according to their similarity of purpose and intentionality.[24] Each such group will be referred to as an *intentional group* or, where usage is unambiguous, simply as a *group*. An *intentional mix* is the proportion of an auditor's total task behaviors to which an observer attributes the same or similar purpose and intentionality. Assume

---

23. Were such a scale available, then there would be a "best" or "most expert" way to perform a task. While one may argue that in certain constrained circumstances, an optimal strategy may be put forth, such circumstances are not representative of those surrounding the work of auditors nor are they typical of the environments within which auditors function. For example, such a set of circumstances prevails during performance of a task on an assembly line, in which individuals having a specific skill perform a repetitive, manual operation within a controlled and highly standardized physical, technological, and social environment, and in which a specific, quantifiable measure of effectiveness exists and is specified (e.g., zero variance from standard cost).

24. Both purposefulness and intentionality are attributions to another made by observers of the other's actions. Russo's model of auditor behavior during empirically intense audit tasks holds that auditors are purposeful and intentional in their choice of task behaviors. Purposeful means that the behavior observed is consistent with the auditor's perception of the objective of that behavior. Intentional means that the behavior and its targets are consistent with the auditor's beliefs and opinions at the moment the behavior is performed. A concept of intrinsic purposefulness and intentionality is, in the context of this research, not useful. To illustrate, while a flathead screwdriver, as opposed to a philips-head screwdriver (both common household tools), may be attributed intrinsic purpose and intentionality, neither would be correct attributions to make when it is observed that an individual is using it to pry open the lid of a paint can. The ostensible purpose here is to gain leverage, not to join two pieces of wood, and the intentional target of the behavior is the lid of the can, not a screw. For additional discussion of these assumptions, see Russo (1997b: note 26; 1998: 12-13).

that b different kinds of observable behaviors can be assigned to two groups, $g \in \{g1,g2\}$,[25] based on an observer's attributions of purpose and intentionality to each. Let us define the "*g1 mix*" as $m_{g1} \equiv \sum^{g1} m_j$. A similar mix is also defined for behaviors categorized as *g2*. These intentional mix proportions measure the extent to which behaviors of particular type were employed in a task relative to behaviors of all types. For example, suppose *g1* behaviors are those associated with the acquisition of information from the task environment (e.g., reading, inquiry, etc.), and *g2* are other non-information acquisition behaviors associated with altering the task environment (e.g., writing, calculating, etc.), then if the *g1* mix of one auditor were greater than that of another, one may say that, compared with the second auditor, the first was more focused on acquiring information, while the second was more focused on producing the transformations that, in an empirically intense task, will be ultimately required for a task solution. Thus, although the task behavior mix effect (E_m) allows one to state that two auditors did in fact follow different strategies in performing their tasks, the intentional mix allows one to characterize a quality of that difference in terms that are meaningful as strategic characterizations.[26] This quality, when observed about a particular auditor's task behavior, will be referred to as that auditor's *strategic emphasis*, the nature of that emphasis being indicated by reference to the related intentional group (e.g., an "information acquisition emphasis").

Analysis of the Behavior Mix Effect (E_m)

The question to be addressed in the following discussion is: Given knowledge accessibility and the possibility that auditors may, to the extent reflected by their strategic emphases, follow different strategies, how and to what extent are these differences responsible for inter-auditor variations in task automaticity? It must be emphasized that, in making comparisons of this kind, we are not so much interested in characterizing strategic behavior as we are in recognizing differences and changes in behavior that signal a difference or change in strategy. Further, in this paper, our interest in those differences is primarily from the point of view of their effects as contributors to variations in assessments of task automaticity. The analysis of strategy, *per se*, is a matter left for a future paper.

25. When used as a subscript, g indicates either *g1* or *g2*. When used as the range of summation, g indicates that the summation is made over the range of behaviors assigned to the intentional group indicated by the subscript of the independent variable. For example, assume $g1 \equiv \{j \in \{1, 2\}\}$. Then $\sum^{g1} n_j = n_1 + n_2$.

26. Note that while a behavior mix effect that differs from unity is sufficient evidence of a difference in strategy, it is not necessary evidence. This condition arises because, although two auditors might perform the same behaviors over their tasks and with the same intensities, they may do so in a different temporal sequence. In such an instance, although both the behavior mix effect and the intentional mix proportions will be the same, these indications will not reveal the fact of a strategic difference. Since temporal differences do not affect the perception of task automaticity, they are not considered in this paper.

Preliminary Matters

Before discussing the metrics needed to answer the question posed in the preceding paragraph, certain preliminary matters must be introduced.

Maintained condition: Inspection of the mix effect (equation (5), repeated as equation (14) for convenience) shows that it is the ratio of Auditor X's observed task automaticity to what that task automaticity would be if (a) Auditor X's behavior mix were the same as that of Auditor N with (b) no change in X's knowledge base properties. The denominator of this ratio will be referred to as the *base automaticity* of Auditor N. In the discussion which follows, it will be necessary to remember that, unless otherwise stated, condition (b) in the previous sentence is maintained throughout.

$$E_m = \frac{\sum^j a_j^X m_j^X}{\sum^j a_j^X m_j^N} \tag{14}$$

Group automaticity: Paralleling the conceptualization of task automaticity given previously, the automaticity with which the behaviors of an intentional group are performed is measured by the ratio of all group automatic knowledge base responses, n_{eg}, to the total instances of group task behaviors, n_g. It can be shown that, for the auditor designated as X, this ratio can be expressed in a form more convenient for use in subsequent analysis as equation (15).[27] The *expected automaticity* of any auditor P's intentional mix is the group automaticity weighted by P's intentional mix, i.e., $a_g^X m_g^P$.

$$a_g^X = \frac{\sum^g a_j^X m_j^X}{m_g^X} \tag{15}$$

Strategic emphasis effect: This effect measures the proportion of P's task automaticity that is represented by the automaticity of behaviors forming a particular intentional mix under the maintained condition. This proportion is given by equation (16).

$$M_g^P = \frac{\sum^g a_j^X m_j^P}{\sum^j a_j^X m_j^P} \tag{16}$$

27. $a_g = n_{eg}/n_g = \sum^g n_{ej}/\sum^g n_j = \sum^g n_j a_j/\sum^g n m_j = \sum^g n m_j a_j/\sum^g n m_j = \sum^g a_j m_j/\sum^g m_j = \sum^g a_j m_j/m_g$. While this demonstration can be advanced for either Auditor X or N, because of the maintained condition mentioned earlier in the text, the group automaticity of auditor N is not particularly relevant in this discussion.

Components of the Task Behavior Mix Effect (E_m)

Given that task behaviors have been intentionally classified, as previously described, the task behavior mix effect may be expressed as the sum of two components:

$$E_m = E_{m(g1)} + E_{m(g2)} \tag{17}$$

where $E_{m(g1)}$ and $E_{m(g2)}$, referred to hereafter as *group intentional mix effects*, represent the contribution to the task behavior mix effect of differences between auditors X and N in their utilization of $g1$ and $g2$ behaviors, respectively. Applying this definition, and given equations (17) and (14), for any group, g,

$$E_{m(g)} = \frac{\sum^{g} a_j^X m_j^X}{\sum^{j} a_j^X m_j^N} \tag{18}$$

Partitioning of Group Intentional Mix Effects ($E_{m(g)}$)

Equation (17), which represents the contribution of inter-auditor differences in group task strategy to total relative task automaticity, is related to the underlying strategic emphases of both auditors. For this reason, group intentional mix effects will sometimes be referred to as *group strategy effects*. The difference between the group intentional mix effect ($E_{m(g)}$) and the effect of Auditor N's strategic emphasis (M_g^N) can be partitioned into a component due to differences in strategic emphasis between auditors and a component due to differences in the manner in which each auditor pursued the strategic objectives of each intentional group. The first of these will be termed the *between-strategies effect* (D_g^{XN}), and the second, the *within-strategies effect* ($D_{g(j)}^N$).[28]

Between-strategies effects measure the contribution to group intentional mix effects due to differences between auditors in strategic emphasis, and is related to $m_g^X - m_g^N$, as shown by equation (19).

$$D_g^{XN} = \frac{a_g^X (m_g^X - m_g^N)}{\sum^{j} a_j^X m_j^N} \tag{19}$$

28. Auditor N is the model's basis for comparison. Both D_g^{XX} and $D_{g(j)}^X$ are always zero.

Within-strategy effects measure the contribution to group intentional mix effects due to differences between auditors in the *conditional mix* of behaviors performed to implement group intentional strategy (i.e., the within-group mix of each behavior, given by the ratio m_j^P/m_g^P). The within-strategy effects are directly related to $(m_j^X/m_g^X)-(m_j^N/m_g^N)$, as shown by equation (20).[29]

$$D_{g(j)}^N = \frac{m_g^N}{\sum^j a_j^X m_j^N} \sum^g a_j^X \left(\frac{m_j^X}{m_g^X} - \frac{m_j^N}{m_g^N} \right) \tag{20}$$

Equation (21) summarizes the preceding discussion.

$$E_{m(g)} = M_g^N + D_g^{XN} + D_{g(j)}^N \tag{21}$$

Under a null hypothesis that both auditors followed (1) the same broad strategy (indicated by their intentional mix proportions), and, for each intentional group, (2) the same intentional strategy (indicated by the conditional mix of behaviors within intentional groups), then both sources of variance would have a value of zero. Each auditor would then display the same group intentional mix effects, equal to the effect of their strategic emphasis, i.e., $E_{m(g)} = M_g^X = M_g^N$.

The three independent variables of equation (21) constitute lower level components of group strategy effects. Hence, in Figure 1, they are shown at a lower hierarchical level. The base auditor's strategic emphasis effect (M_g^N) is not directly involved in auditor comparisons, its contribution being reflected in the between-strategies effect. This indirect relationship is shown by means of the broken lines in Figure 1. The additive nature of their relationship is indicated in Figure 1 by the confluence of multiple lower level effects to form a higher level effect.

29. Equation (20), while an inelegant presentation, is nevertheless used in the text because it provides a conceptually clear indication of the nature of within-strategy variances and their contribution to group intentional mix effects. The within-strategies effects can also be considered as the effect of any difference between the expected automaticity of Auditor *N*'s intentional mix and its observed automaticity, as shown by the following equation: $D_{g(j)}^N = \frac{a_g^X m_g^N - \sum^g a_j^X m_j^N}{\sum^j a_j^X m_j^N}$. This equation is computationally simpler than is equation (20) but does not immediately foster intuitive recognition of $D_{g(j)}^N$ as arising from within-group mix differences.

Directional Effects

Directional effects are the differential consequences for automaticity assessments arising out of an observer's use of alternative reference bases and strategic perspectives in making comparisons. In any comparative study, the choice of a base to which particular instances of behavior are referenced is a critical issue for any meaningful communication and synthesis of findings. Because of the requirement that auditors be designated as either Subject (X) or base (N), and the availability of the option to study automaticity from the perspective of one intentional group or the other (i.e., g_1 vs. g_2 or g_2 vs. g_1), implementation of the proposed model necessarily incorporates directional effects as an unavoidable mathematical consequence.

There are two pragmatic reasons for studying directional effects. First, directional choice appears to be an arbitrary aspect of the model. Second, an understanding of directional effects can potentially result in saving considerable analytical effort in implementing the model. Each of these reasons is discussed in the following sections.

Apparent Model Arbitrariness

The directional choices mentioned above appear to introduce an element of arbitrariness into the model's analytical process, and for this reason, they demand investigation. In the sections that follow, I discuss how directional choices affect model metrics, and in particular, given a set of metrics, what directional inferences may and may not be validly made concerning relationships in the reverse direction.[30]

---

30. Observer related effects are potential sources of inter-auditor variation in assessments of expert-like behavior that arise from a number of sources. For a discussion of methodologically induced observer error, see Cooper, Heron, & Heward (1987) and Russo (1994: 191). Two forms of observer related effects that are particularly relevant to this paper are "between-observer" variation and "base" variation, referred to in this paper as "directional effects." The first of these, between-observer variation, arises because strategic characterization of behavior is observer dependent, as discussed earlier in this paper. Therefore, different observers of the same behavior can arrive at different characterizations. This source of variation can be minimized by use of the same observer for all cases, and by standardized behavior observation methodology, protocol coding, and observer training. The second observer related effect is actually a procedural and logical error that arises whenever a given observer is inconsistent in use of a base for making comparisons. For certain purposes, and at various times, an observer making comparisons between two cases, A and B, may draw comparisons from A to B, while for other purposes and at other times, the same observer may draw comparisons from B to A. For example, if the measures of the properties of A were {2,1} and those of B were {4,2}, then if a given observer reports the comparative measures {½ ,2}, a base error has been committed. Without knowledge of the direction in which the observer drew the comparisons, (i.e., the base that the observer has adopted) one is unable to gain a correct interpretation of the relationship between A and B. Correct reports are either {½, ½} if B is chosen as the base, or {2,2} if A is chosen. However, note that although both reports may be procedurally correct, logically both can still lead to erroneous interpretation if the receiver of the report does not know the base used by the sender.

Complementarity

The property of complementarity concerns the directional relationship between a measure of an effect (its magnitude) and its expected value in either of the two cases described below. If the signum (i.e., the algebraic sense, plus or minus) of the difference changes when the direction of comparison changes, then the effect is complementary.

Case 1: XN-Complementarity. In Case 1 complementarity, the assignment of auditors as X and N is reversed. More precisely, if the observed measure of an effect is on one side of its hypothesized mean (e.g., $T^{AB} > H_0\,(T^{AB})$)[31] and that effect possesses the property of XN-complementarity, then when the auditor assignments are reversed, the measure of that effect will fall on the side opposite to its former position relative to its hypothesized mean (i.e., $T^{BA} < H_0\,(T^{BA})$).

If a model metric exhibits XN-complementarity, then the signum of the difference between each complementary measure and its expected value under the null hypothesis is the negation of the other. To illustrate, consider the between-strategies effect, given assignments of auditors A and B to X and N, respectively. Applying equation (19), and bearing in mind that the expected value of a difference statistic (any D metric) around its mean is zero, we have:

$$D_g^{AB} = \frac{a_g^A(m_g^A - m_g^B)}{\sum^j a_j^A m_j^B}, \quad D_g^{BA} = \frac{a_g^B(m_g^B - m_g^A)}{\sum^j a_j^B m_j^A} \qquad \textbf{(22)}$$

None of the variables in the preceding functions can be negative. Therefore, the signum of each effect is determined solely by the signum of the difference between the g-conditional mix of auditors A and B. Whatever the signum of (m_g^A - m_g^B), it is clear that when the order of subtraction is reversed, so too will the signum be reversed. Consequently, between-strategies effects exhibit XN-complementarity. By a similar process, it can be determined that the following model effects are XN-complementary: E, E_a, A, A_s, A_c, A_e and D_g^{XN}.[32]

However, as can be surmised from inspection of equation (20), within-strategy effects are not XN-complementary. That equation includes a weighted summation. Reversal of the X and N

---

31. Throughout this paper, the symbol T will be used to represent any arbitrary model metric.

32. The process is more subtle with ratios (E and A variables) than with differences (D variables). To illustrate, A^{XN}-1 can be written as $\dfrac{\sum^j (a_j^A - a_j^B)}{\sum^j a_j^A}$. In this form, it is clear that A (relative knowledge accessibility) is XN-complementary.

assignments, while resulting in a reversal of the signum of the individual differences within the bracketed portion of the equation, also alters the set of accessibility weights applied to those differences, making the signum of the summation term indeterminate.

Case 2: g-complementarity. With regard to Case 2, given an assignment of auditors as X and N and that the directional relationship between the measure of a model effect and its hypothesized mean is examined from the perspective of one intentional group (either $g1$ or $g2$), if the effect is g-complementary, then the opposite directional relationship between that metric and its hypothesized mean obtains when examined from the prospective of the other intentional group. Case 2 keeps the assignments to X and N constant, but reverses the group subscript from that of one group to that of the other. For example, if $T_{g1}^{AB} > H_0\,(T_{g1}^{AB})$, then if T is a g-complementary effect, then $T_{g2}^{AB} < H_0\,(T_{g2}^{AB})$. To illustrate, consider, again, between-strategy effects, in which auditors A and B have been assigned to X and N, respectively. By construction, intentional groups $g1$ and $g2$ are complements in the sense that $\{j\} = \{g_1\} \cup \{g_2\}$, where $\{g_2\} = \{\overline{g_1}\}$. Therefore, $m_{g2} = 1 - m_{g1}$. Again, from equation (19), and applying the preceding relationship between m_{g1} and m_{g2}, we obtain:

$$D_{g1}^{AB} = \frac{a_{g1}^{A}(m_{g1}^{A} - m_{g1}^{B})}{\sum^{j} a_j^A m_j^B}, \quad D_{g2}^{AB} = \frac{a_{g2}^{A}(m_{g2}^{A} - m_{g2}^{B})}{\sum^{j} a_j^A m_j^B} = \frac{-a_{g2}^{A}(m_{g1}^{A} - m_{g1}^{B})}{\sum^{j} a_j^A m_j^B} \tag{23}$$

Since the expected value of any D effect is zero, the change in sign shows that the between-strategies effect is g-complementary.

Conclusions regarding the g-complementarity of within-strategy effects are best approached by examining equation (24).

$$\sum^{g} D_{g(j)}^{N} = D_{g1(j)}^{N} + D_{g2(j)}^{N} \tag{24}$$

Each of the terms of equation (24) is itself a difference and may be either positive or negative. Any combination of positive and negative effects can exist and still satisfy the equation. However, regardless of the signum of each individual term, and given an assignment of auditors, the total on the left side of the equation is a constant. Therefore, within-strategy effects are not g-complementary.

Strong and Weak Complementarity

The model compares the task behaviors of a Subject auditor, designated by the superscript X, against the parallel behaviors of a base auditor, designated by the superscript N. The strength of complementarity, i.e., strong or weak, speaks to the consistency of a finding of significance when the direction of comparison is reversed. If an effect has strong complementarity, and if it is significant when comparisons are made in one direction, it is *always* significant when that direction is reversed. On the other hand, if an effect is complementary, but the consistency of a finding (significant or not significant) cannot be assured when the direction of comparison is reversed, then the effect's complementarity is termed weak. More specifically, given that (1) Auditor A is designated as X and Auditor B as N (denoted by superscript AB), and (2) a finding of significance (a deviation from a hypothesized value that cannot be attributed to chance) for any effect, T^{AB}, then if T^{AB} is a strong complementary effect, T^{BA} will be significant also.[33]

If a model effect exhibits strong complementarity, then the product of the measure of that effect under an initial set of assignments and the measure of the same effect with the assignments reversed evaluates to unity.[34] To illustrate, consider task automaticity, E^{AB}, where Auditor A is assigned as X with Auditor B as the comparative base, N. Then from equation (3), the following relationships hold:

$$E^{AB}=\frac{\sum^{j} a_j^A m_j^A}{\sum^{j} a_j^B m_j^B}, \quad E^{BA}=\frac{\sum^{j} a_j^B m_j^B}{\sum^{j} a_j^A m_j^A} \tag{25}$$

from which it is readily determined that $E^{AB}E^{BA} = 1$. Therefore, task automaticity is a strong complementary effect and one can conclude that if there is a significant finding when Auditor A is compared with Auditor B as a base, there will also be a significant finding if Auditor B is compared using Auditor A as a base. By means of this process, it can be determined that the following are strong complementary effects: E, A, A_s, A_c, and A_e. While it was previously determined that between-strategy effects (D_g^{XN}) are both XN- and g-complementary, applying the above test to these effects shows that they are not strongly complementary.[35]

---

33. The same relationships hold if T^{AB} is not found to be significant.

34. Assume that $T=T(X-N)$ exists over the domain of real numbers ($\Re$) and that $f(T)$ is a real transformation of T over its entire range. Let T be evaluated in the direction of A to B, i.e., $T^{AB}=T(A-B)$. Let the range $[v_l \cdots v_h] \subset \Re$ and $[v_l \cdots v_h] \ni T^{AB}$. Then $[f(v_l) \cdots f(v_h)] \ni f(T^{AB})$. Now, if $T^{BA}=f(T^{AB})=T(B-A)$, then, $[f(v_l) \cdots f(v_h)] \ni T^{BA}$. If $[v_l \cdots v_h]$ delineates the range beyond which the probability of observing any particular value of T is too small to be due to chance, then both T^{AB} and T^{BA} are significant, e.g., they are strong complementary effects. For all strong complementary effects in this model, $f(T) = 1/T$. Therefore, the product of T and $f(T)$ is unity.

35. Strong complementarity can be easily demonstrated by considering the demonstration of the conditions of complementarity in a previous note. The function $f(T) = 1/T$ reverses the sense of the inequalities making up the limits of the range of significance. For example, if $T \geq v_h$ specifies an upper limit significance, then $T \leq f(v_h)$ specifies a lower limit significance. More generally, assume a variable, $T^{XN} = T(X-N)$ and the following : (1) (T^{XN} - EV(T^{XN})) is real and ranges over $-\infty$ to $+\infty$, (2) a function $f(T)$ such that (EV(T^{AB}) = EV(T^{BA})), and (3) P($T^{BA} \mid T^{AB}$) =P($f(T^{AB})) \mid T^{AB}$) and P($T^{AB} \mid T^{BA}$) =P($f(T^{BA}) \mid T^{BA}$) are both equal to 1. There exists only two possible sequences of comparisons that can be made. Either we initially compare A with B, followed by comparing B with A, or we initially compare B with A, followed by comparing A with B. In the first possibility, we obtain the metric set $\{T^{AB}, T^{BA}\}$, in that order, while in the second we

Table 1 provides a summary of the preceding discussion.

TABLE 1
SUMMARY OF OBSERVER DIRECTIONAL EFFECTS

	Task Automaticity									
		Knowledge Accessibility Effect						Strategic Knowledge Effect		
			Knowledge Properties						Between strategies	Within-strategy
			Accessibility				Knowledge Utilization			
				Organization	Content	Availability				
	E	E_a	A	A_s	A_c	A_e	E_x	E_m	D_g^{XN}	$D_{g(j)}^{N}$
Case 1 - XN-COMPLEMENTARITY										
Weak	Y	Y	Y	Y	Y	Y	N	N	Y	N
Strong	Y	N	Y	Y	Y	Y	N	N	N	N
Case 2: g-COMPLEMENTARITY									Y	N

Notes: If an effect is complementary, then:
Case 1 (XN-complementarity) - the signum of the difference between the measure of the effect and its expected value is reversed if the assignment of auditors for comparison purposes is reversed. If an effect is complementary strong, then, given an assignment of auditors for purposes of comparison, a significant (not significant) finding will remain significant (not significant) if that assignment is reversed.

Case 2 (g-complementarity) - the signum of the difference between the measure of the effect and its expected value when examined from a given an intentional perspective is reversed if the metric is examined from the perspective of the complementary intentional group. g-Complementary effects are weak.

---

obtain $\{T^{BA}, T^{AB}\}$, in that order. Since these two sets exhaust the set of possible outcomes, the sum of the joint probabilities of observing each set is $P(T^{AB}|T^{BA})P(T^{BA}|T^{AB}) + P(T^{BA}|T^{AB})P(T^{AB}|T^{BA}) = 1$. Given the conditional probabilities mentioned in (3), this expression reduces to $P(T^{BA}|T^{AB}) = 1 - P(T^{AB}|T^{BA})$. Now for all strongly complementary effects in this model, $f(T^{AB}) = 1/T^{BA}$, a function satisfying all three assumptions above. Where the model permits variables to be either weak or non-complementary, *viz,* $E_a, E_x, E_m, D_g^{XN}, D_{g(j)}^{N}$, assumption (3) above cannot be made. Hence, given a finding concerning one of the pair of directional mates of such effects, no inference can be made on that basis concerning the significance of the other.

Interpretive Implications

As has been stated, the primary focus of this paper is on the variability of assessments of expert-like behavior across auditors. However, these assessments are the products of observer-observed systems. Hence, the directional choices made by different observers, or by the same observer at different times, raises possible concerns as to whether the reported assessments may differ solely as a result of those choices, and if so, what implications can be drawn. Fortunately, the model provides considerable assurances on these points.

For strong complementary observations, all conclusions drawn as to significance will be the same, regardless of directional differences and methodological or observer inconsistencies in direction. Similarly, even in the face of such differences and inconsistencies, weak complementary effects will maintain their qualitative relationships to expected means, although the significance of those relationships must be established for each direction as a separate matter. Thus, regarding conclusions based upon complementary effects, differences in assessments of task automaticity are easily explained. Consequently, where these effects are an important focus of study, choices made and methodological controls, at least with respect to direction, are not major concerns.

In contrast to the preceding, within-strategies effects are not complementary; inconsistencies in auditor assignments and strategic perspective can produce significant differences in both findings and interpretation. If these effects are an important focus of study, care in making directional and perspective choices and strict methodological controls will be required in order to minimize the analytical and interpretive difficulties these effects present.

In summary, the potential danger from ambiguity in the model's directional choices is more apparent than real. Seven of the ten model metrics listed in Table 1 are complementary, and of these, all but two are strongly complementary, thereby assuring that for most analytical purposes, observer directional choices will have little, if any, effect on the significance of findings and interpretations. As to the few for which this assurance is not absolute, the maxim "forewarned is forearmed" precisely draws attention to possible differences in interpretation which, if significant to the purposes of an investigation, can be addressed with little additional effort.

Reduction in Analytical Effort

The second reason mentioned for investigating directional effects is that, through appropriate use, the analytical effort required to apply the model can be greatly reduced. As shown in the preceding discussion, the model assures that most XN-complementary effects are strong. Therefore, it is not necessary that these be evaluated for significance in both directions. In addition, under the null hypothesis that auditors X and N share common knowledge base properties and strategic behavior, the distribution of each complementary model metric is identical in both directions. Therefore, the savings in effort extend not only to computations of the particular effects involved, but perhaps more importantly, they extend to the bootstrap simulations of the related probability distributions, discussed later, that are required for the interpretation of findings. The latter entails vastly more time and resources than does evaluation of the effects themselves.

Deviations from Complementarity: Asymmetry

Suppose that **n** auditors are evaluated one against the other in terms of a particular effect, T, having an expected value of T_0 under a null hypothesis that both share common knowledge properties and strategic behaviors. The various assignments of auditors as Subject (X) and reference (N) can be represented in an **n**-by-**n** array by the intersection of correspondingly ordered row and column positions. Each of the $\mathbf{n}^2$-$\mathbf{n}$ possible pairs of auditor assignments is then tested for significance and the quality of the outcome indicated by placing a symbol in the box representing the intersection of each paired auditor assignment. If T is a strong complementary effect, then that property will be indicated by a symmetric distribution of symbols around the main diagonal of the array. The term *symmetry* will be used to indicate the appearance of any effect that presents in this manner. Symmetry can also be applied to plots of the signa of effects, regardless of the outcomes from tests of significance. If the effects are complementary, they always will be symetrical as to signum, but not necessarily as to significance, around the main diagonal.

It is mathematically impossible to observe asymmetry in an effect the model specifies is strongly complementary or to observe a lack of complementarity in an effect the model specifies is complementary. If such observations are obtained, they indicate possible methodological, analytical, or mathematical error. Where significant effects of this kind are observed and different observers are involved, one should suspect inter-observer inconsistency in the behavior observation and/or coding methodology.

The preceding not withstanding, it *is* possible to observe the negation of these conditions. That is, it is possible to observe significant symmetry and complementarity in any effect for which such occurrences are *not* normal model expectations; the absence of mathematical assurance does not forbid such effects from displaying symmetry or complementarity. Therefore, while the terms "non-complementary" and "asymmetric" will be employed to identify the class of effects for which such assurances are absent, they are not to be understood in the sense that such effects can never be observed to be symmetric or complementary. Outcomes that present patterns of complementarity and symmetry that are not expected, based upon the model, are referred to as *deviant* outcomes or patterns. Effects which may present deviant patterns are those labeled "N" in Table 1.

The issue of deviant outcomes is relevant only to those effects that the model cannot assure are symmetric or complementary and for which the probability of occurrence is too low to be due to chance. Where deviant observations are made, interesting questions are posed as to what such conditions reveal about the processes underlying both the expert-like behaviors of the auditors involved and the assessment process itself. Generally, given appropriate methodological controls and an assignment of Subject and base auditors (auditors X and N respectively, as discussed previously), observation of a deviant effect can result from a pervasive similarity in knowledge properties and strategy (an empirical matter based upon observation), the level of confidence adopted by the observer for purposes of assessing significant differences, or a combination of both of these causes. As to the first of these causes, the extent of the similarity that must be present to produce a deviant observation is most easily ascertained by inspecting the deviant effect's determining function and comparing the variances of that function's arguments. The variance of any dependent variable is a positive function of the variances of its independent variables.[36] Under the null hypothesis, both auditors in a pair (i.e., auditors X and N) have the same knowledge properties and behaviors, estimated by the distributions of responses in their combined solution sequence during performance of a task. Therefore, the expected values and variances of each of the arguments relating to a single auditor, and all arguments that form functions of the products of metrics taken from both auditors (e.g., $\sum^j a_j^X m_j^N, \sum^g a_j^X m_j^N$, etc.), are, pair-wise, equal and unaffected by the direction of auditor assignment.[37] Therefore, under the null, the distribution of any non-complementary effect is identical regardless of direction.

As to the second cause mentioned, consider the case in which the actual knowledge properties and behavior mixes of auditors X and N differ from each other. Under such circumstances, the expected values and variances of the arguments in the determining functions of all weak- and non-complementary effects, and the distributions of the effects themselves, will differ depending on the direction of the auditor assignment. Without the assurance of strong complementarity provided by the model, the power of any test for significance to detect a difference between hypothesized and actual means will be asymmetric with direction. All else being equal, the greater the disparity in the effect's mean and variance when the direction of comparison is reversed, the greater the chance of observing asymmetric findings in significance outcomes. For the same reason, all else being equal, the higher the confidence limit established for detecting significant differences, the greater the chance of observing asymmetric effects.[38] In this respect, observer choices about confidence levels

36. Within the context of this model and its application, covariances which can complicate or pose significant challenge to this statement are not likely to be significant.

37. These pair-wise expectations are $a_j^A = a_j^B, a_g^A = a_g^B, m_j^A = m_j^B, m_g^A = m_g^B$, the composite functions $\sum^j a_j^A m_j^B = \sum^j a_j^B m_j^A, \sum^g a_j^A m_j^B = \sum^g a_j^B m_j^A$ and their associated variances.

38. Each auditor, A and B, independently produces a unique solution sequence $\{a, m\}^A$ and $\{a, m\}^B$, respectively. The actual distributions of effects T^{AB} and T^{BA} are the product of the elements of superset $\{\{a, m\}^A, \{a, m\}^B\}$, but selected and applied in a manner consistent with the auditor assignments in each case. The model assures that if T is a complementary effect, then T^{AB} and T^{BA} will fall on opposite sides of the mean T_0 as determined under the null hypothesis. Given an empirically observed superset, as above, each T^{AB} has associated with it a unique T^{BA}. If T is complementary strong, T^{BA} is a transform of T^{AB}, and no further empirical information is required to determine the latter value. Strong complementarity assures that a given value of T^{AB} will always map to the same value of T^{BA}. However, if T is complementary weak, then T^{BA} must be computed from additional empirical data which must be obtained from the superset. Since there are many possible sets $\{a, m\}^B$ that can be associated with a given set $\{a, m\}^A$, that data, and in consequence, T^{BA}, are subject to random variation. Thus, a given value of T^{AB} will map to a distribution of values of

used to test for significant effects, rather than real inter-auditor differences, can be responsible for variability in assessments of relative expert-like task behavior. Depending on the research issue at hand and the cost-benefit expectations involved, the preceding considerations may be considered as cues for initiating further analysis of the deviant results obtained.[39]

RESEARCH QUESTIONS AND HYPOTHESES

The focus of this paper, as discussed in the Introduction, is an examination of the limits of typicality, i.e., the extent to which individuals who would otherwise be considered as members of a homogenous group differ from one another with respect to their expert-like behavior during performance of a field task. In the previous sections, it is argued that comparative task automaticity is an acceptable measure for assessing expert-like task behavior. This assessment is based on *E*, the ratio of the observed task automaticity of one auditor, *X*, to another, *N*. The model presented in this paper identifies a hierarchy of effects, summarized in Figure 1, contributing to *E*. In this section, with the aid of the preceding model, I propose some specific investigations relevant to issues raised in the Introduction.

In the illustration to be presented later in this paper, the matter of typicality is approached from two perspectives. The first examines the representativeness of group metrics when applied to specific individuals who are members of that group. The second perspective is the extent of variation, if any, between pairs of individual group members. For each of the hierarchically determinative measures in Figure 1, the indicated null hypothesis is based on the knowledge properties and behavior mix of the typical auditor in the group, or the mean auditor of the pair, as appropriate to each of the perspectives mentioned. It is important to note that this examination is limited to variation only between individuals who are actual participants in a group. Therefore, sample selection error, which reflects the degree to which group participants are representative of

T^{BA}, the particular value on any specific occasion being determined by Auditor B's observed solution sequence in that task at that time. Thus, $a_j^A \neq a_j^B$ and $m_j^A \neq m_j^B$ for at least some *j*. Then it does *not* necessarily follow that $EV(a_j^A)=EV(a_j^B)$, $VAR(a_j^A)=VAR(a_j^B)$, $EV(\sum^J a_j^A m_j^B)=EV(\sum^J a_j^B m_j^A)$ or $EV(\sum^g a_j^A m_j^B)=EV(\sum^g a_j^B m_j^A)$. Further, given these conditions, $VAR(\sum^J a_j^A m_j^B) \neq VAR(\sum^J a_j^B m_j^A)$ and $VAR(\sum^g a_j^A m_j^B) \neq VAR(\sum^g a_j^B m_j^A)$, from which the conclusions stated in the text follow.

39. The major benefit to be gained from the discussion above is avoidance of a rush to conclude that observation of XN-symmetry when it is not expected implies any specific relationship between auditors *X* and *N* regarding their knowledge or behavior. Since there are an infinite number of sets of $\{a_j m_j\}^P$ that can produce the same measure of base automaticity, it is incorrect to conclude, solely on the observation of XN-symmetry among between-strategies effects, that the auditors being compared are very similar in either knowledge or behavior. Their respective strategic emphases may make equally significant and complementary contributions to their relative task automaticity, but they may be quite different in their knowledge base properties, their strategic implementations, or both. These, then, are the kinds of matters that are to be subjected to cost-benefit analysis, should they be of particular relevance to the purposes for which the behavior assessment is being performed.

specific individuals within the extended class from which group members are drawn, is not considered.

Variability From Typical Measures

For purposes of the first perspective, the metrics of the hypothesized typical auditor are the mean metrics of all auditors participating in the experiment. Variability from typical measures is a type of analysis most relevant to those situations in which a group norm or expectation is available against which a particular individual is to be compared. In this research, the term "typical" is used in the sense in which the term "normal" is employed in common clinical methodology.[40] In that methodology, relevant norms, developed on the basis of suitably determined sample means, are the starting points for evaluations and decisions affecting specific cases, each of which, in the context of this paper, is a particular individual. In assessing the status of an individual, a clinician will usually encounter deviations from prescribed means which must be evaluated against estimates of the degree of variation that is considered "normal" for each metric, given the population from which the individual is drawn. To the extent that deviations are significant and many, then to that extent, evaluations made and actions taken based on outcomes anticipated from "normal" or "typical" assumptions become increasingly risky and prone to sub-optimal outcome or failure. That variation exists among members of otherwise homogenous populations is well known and expected, as is the fact that the risk discussed above is positively related to the magnitude of that variation. However, in the absence of a more specific identification of the sources of variation, such knowledge is of little comfort to both the clinician and the party at risk. However, where a salient norm is involved, the objective of making comparisons is normally to bring the individual into conformity with the norm rather than *visa versa*. Consequently, in these kinds of comparisons, the complications of directional effects are not likely to be present.

Variability Between Pairs of Auditors

Pair-wise comparison is most relevant in those situations in which a comparative evaluation is required of two auditors who are observed during performance of a task. In such instances, each auditor's behavior acts as a control for comparison with the other. From this perspective, the indicated null hypotheses are based on the mean of the knowledge properties and behavior mix of the auditors forming the pair. This form of analysis is of particular interest to individual auditors because in the absence of a widely recognized norm, there is no generally recognized or salient basis

40. The term "clinician" is used to designate an individual responsible for observations, evaluations, and decision making ("clinical activity") primarily affecting a specific individual or a small group of individuals, and generally carried on with the best interests of that individual or group in mind. However, the term can also be broadly considered to include the bests interests of the organization, mission, etc. for which such activity is carried on. A "clinic" or "clinical setting" is a place where clinical activity takes place. Managers and instructors are, at times, clinicians in this sense.

for comparison.[41] Consequently, directional choices when a salient norm is absent can be a potentially significant source of variability in assessments of expertise.

Sources of Error and Tests for Significance

The sources of error of concern in tests of significance, given the experimental data such as it is, are (1) non-systematic (i.e., random) coding error and (2) non-systematic over/under recognition of the cognitive and automatic components of mediating episodes. The model's functional relationships do not lend themselves to description by any of the commonly used probability distributions. Consequently, in order to test any of the hypothesized model metrics for significance, the probability distributions associated with each were generated by bootstrapping each Subject's task behavior and knowledge base response 10,000 times. The bootstrap simulations involved the distribution of knowledge base responses, by auditor, within target behavior, within mediating episode, at each level of episode complexity. After each iteration, each knowledge base response and behavior sequence was compiled as required by the model, the relevant metrics computed, and the probability distributions updated.[42] [43] The means of the resulting bootstrapped probability distributions were computed and found to evaluate to the appropriate hypothesized means (mean overall error for the metrics tested in this paper: +.0053, range .0168 to .0005, std. deviation: .0057.)

---

41. While the direction of comparison when a norm is involved is generally taken as running from the individual, *X*, to the norm, *N*, it is often less clear or consistent when a salient norm is absent. There are multiple degrees of freedom available in making comparisons between two individuals, each contributing to inter-auditor variability in assessments of expertise. The major possibilities are categorized as follows:

1. The choice of which audifor is designated as *X* and which as *N*. This choice may be:
   a. consistent for all effects compared, or
   b. run in one direction for some effects and in the opposite direction for others.
2. Each of the above degrees of freedom exist both
   a. within the same observer, and
   b. between observers.
3. Each of the preceding degrees of freedom may vary with the nature of the demand necessitating that comparisons be made (e.g., a group norm applied to a group vs. to an individual, etc.)

42. In comparative tests for significance, such as is required here, systematic error in the observation and coding of behavior is not addressed. This type of risk, though not entirely eliminated by the dual coding of protocols, is nonetheless minimized by that procedure. The sources of error of concern in tests of significance are non-systematic (i.e., random) coding error and non-systematic over/under recognition of the cognitive and automatic elements of mediating episodes. For purposes of testing whether the difference from hypothesized means of the metrics obtained from this experiment could be the result of such sources of error, the simulations hypothesize that, for each assignment of auditors as *X* and *N*, the distributions of episode complexity and knowledge- base response is that of their combined solution sequence. Empirical probability distributions were created for both simple and complex mediating episodes.

43. Biddle, et al. (1990) discuss the use of computer intensive methods in auditing.

EXPERIMENT AND BEHAVIOR OBSERVATION METHODOLOGY

The details of the experiment performed and behavior observation methodology employed to obtain data on auditor behavior during empirically intense tasks are too lengthy and complex to be covered here. The following paragraphs present only a brief summary. For a more complete discussion, see Russo (1994, 1995a).

Subjects, Task, and Procedure

Both inexperienced auditor-Subjects and a non-financial-statement related audit task were chosen for this experiment in order to assure observation of a novice problem solving process. The Subjects were four first-year auditors from the professional staff of a Big Six auditing firm. All Subjects were volunteers who had sat for and passed some, but not all, parts of the CPA examination and all had no prior exposure to the subject matter of the task.

The task in this experiment was a review of the Statement of Operating Expenses of a new office building in which the client is a tenant, rendered pursuant to the rent escalation provisions of the client's lease. To acquaint the Subjects with the terminology, administrative, and computational procedures associated with operating expense rent escalations, on the day before the experiment, each was given background material and two samples of completed review reports to study. However, none of this material provided any information on examination or reporting procedures, the landlord's procedures, or the existence and nature of any documents used in the preparation of the statement rendered to the client. Therefore, such a task, to the extent that it differs from that of the usual financial statement audit, would be unfamiliar to the Subjects who participated in this experiment.

Each Subject performed the task on a different day. The task was performed in a simulated business office in which each auditor-Subject was presented with the equipment and supplies normally available in audit environments and the ability to communicate (via intercom) with and receive documents (via a mail slot) from other parties present in the task environment (e.g., client personnel, the audit partner, etc.). During performance of the task, each Subject was free to contact any party in the task environment and to request any documents or explanations required. Although the researcher played the roles of others in the task environment, no face-to-face or verbal contact took place between Subject and experimenter. Responses to requests for explanations were communicated to the Subject via a video display at the Subject's desk.

Behavior Observation Methodology

Synchronized video-taped and think-aloud verbal protocols were used to capture both the observable behaviors and cognitions of the auditor-Subjects during their performance of the task. The experimental protocols were independently coded in terms of the behaviors and cognitions described in Appendix A by the researcher and a first-year doctoral student trained by the researcher. Kappa (Cohen 1960), a widely used measure of the agreement between independent coders, ranged from .78 to .72 over a total of approximately 8 hours of behavior observation. These levels of Kappa are significant at $p < .0000$.

Operationalization of Knowledge-Base Responses

The nature of a knowledge-base response to demands made by a task can be ascertained by examining the composition of episodes mediating transitions between observable behaviors during task performance. The analysis is performed by Subject, and within Subjects, by behavior. Each behavior type is treated as a target. The mediating episodes associated with each group are then analyzed in terms of the kinds of knowledge base responses (r_k) they include. Uncertainty cognitions preceding target behaviors are sufficient evidence that knowledge is absent from a knowledge base. Such a response is coded as type $k = 3$. Analysis and planning cognitions mediating transitions to behaviors are sufficient evidence of cognitive effort in locating knowledge, and hence of accessible, but unavailable knowledge. Such responses are coded as type $k = 2$. Finally, sequences in which a target behavior is preceded by another observable behavior is necessary and sufficient evidence of automatic access to all the knowledge demanded by that behavior transition. These responses are coded as type $k =1$. Definitions of behavior and cognitive response categories are provided in Appendix A.

Operationalization of Intentional Groups

Two intentional groups are used in this study. Group *g1* includes behaviors whose purpose is the acquisition of information from the task environment. These behaviors are targeted at workpapers, source documents, and individuals in the task environment to whom inquiries are addressed. Group *g2* includes behaviors whose purpose is to transform the task environment and that are targeted at workpapers, documents, and devices (e.g., calculator) present in the task environment. These asignments are summarized in Appendix B.

EXPERIMENTAL DATA AND FINDINGS

The data collected from the simulation experiments are voluminous. Consequently only highly condensed data, sufficient for readers to confirm their understanding of the model and reported model metrics, are presented as Table 2. Appendix C provides detailed tables of observed metrics for all

TABLE 2											
OBSERVED RESPONSES											
AUDITOR 1						AUDITOR 2					
Behavior	j	n_j	n_{sj}	n_{cj}	n_{ej}	Behavior	j	n_j	n_{sj}	n_{cj}	n_{ej}
Reading	1	117	170	101	31	Reading	1	102	134	93	42
Requesting	2	29	42	29	16	Requesting	2	41	47	40	29
Calculating	3	11	14	11	3	Calculating	3	13	18	13	7
Writing	4	21	28	22	12	Writing	4	42	48	42	20
Other	5	10	16	10	5	Other	5	20	24	20	7
AUDITOR 3						AUDITOR 4					
Behavior	j	n_j	n_{sj}	n_{cj}	n_{ej}	Behavior	j	n_j	n_{sj}	n_{cj}	n_{ej}
Reading	1	93	117	76	35	Reading	1	59	103	78	22
Requesting	2	32	43	36	22	Requesting	2	27	40	35	18
Calculating	3	34	47	38	12	Calculating	3	14	18	17	5
Writing	4	55	72	50	17	Writing	4	42	79	64	8
Other	5	18	20	16	9	Other	5	6	7	4	3

Notes: n indicates a sum of individual target behavior occurrences (subscript j) or knowledge-base responses (r_1, r_2, r_3). *The subscript s indicates the types of knowledge-base responses included in the sum* $(r_1 + r_2 + r_3)$; *the subscript c indicates that only analysis, planning, and automatic knowledge-base responses are included* $(r_1 + r_2)$; *the subscript e indicates that only automatic knowledge-base responses are included* (r_1). *See Appendix A for definitions of behaviors.*

hypotheses, including the results of tests of significance. Finally, Appendix D provides extracts of relevant probabilities obtained from bootstrapped simulations under the various null hypotheses.[44] Because of the broad scope of the model presented in this paper and the complexity of the relationships to be examined in the discussion to follow, the usual tabular presentation of experimental findings is inadequate for purposes of rapidly gaining insight into the variations in expert-like behavior among the auditors participating in this experiment. Consequently, Figures 2 and 3 present the findings in a form I will refer to as a "hypogram."

44. A more complete report of the experimental data and boot-strap tables is available upon written request. Please contact the author.

FIGURE 2

PHENOMENA AFFECTING RELATIVE EXPERT-LIKE TASK BEHAVIOR EACH AUDITOR vs. "TYPICAL AUDITOR"

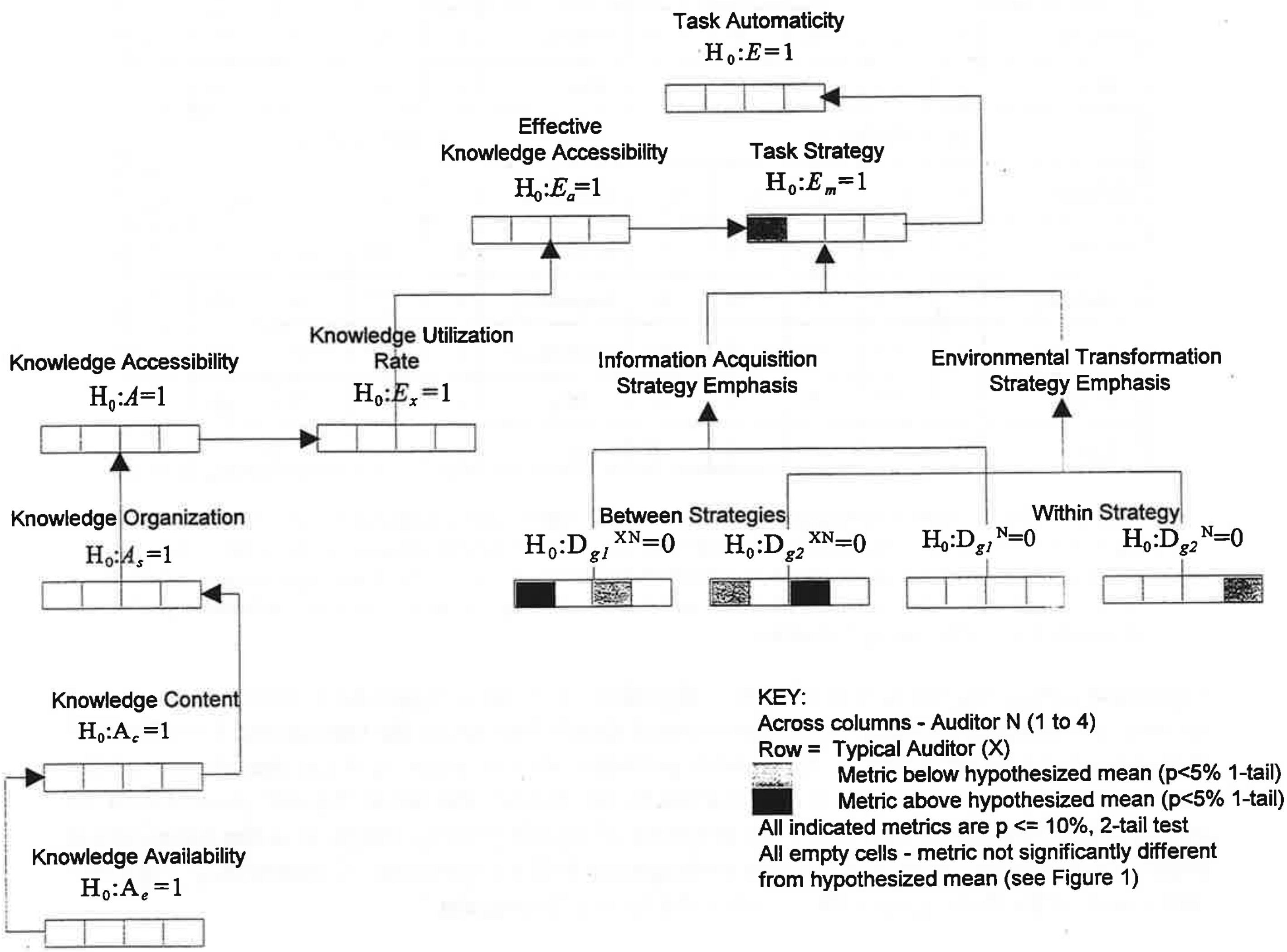
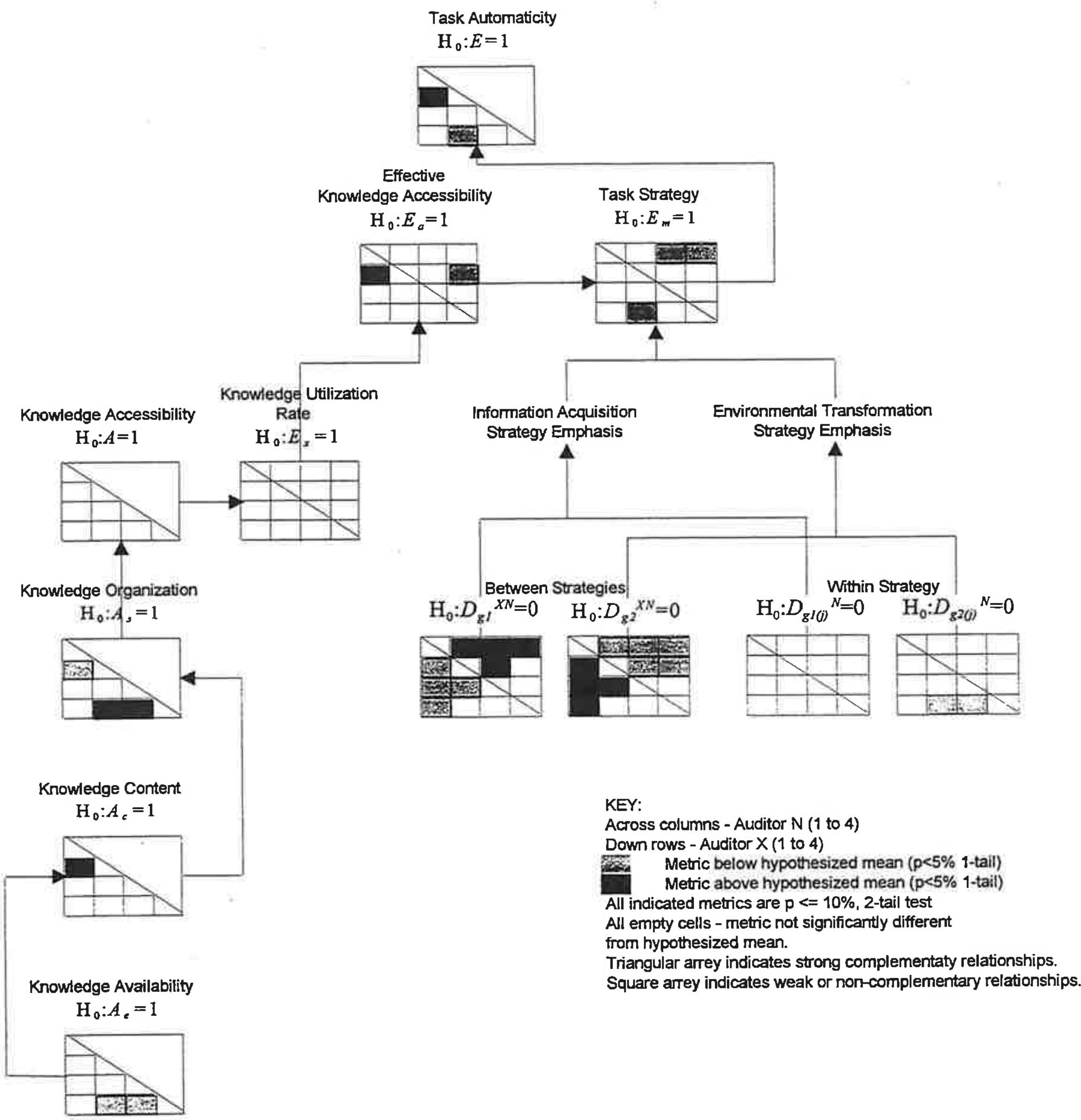


FIGURE 3

PHENOMENA AFFECTING RELATIVE EXPERT-LIKE TASK BEHAVIOR
AUDITOR vs. AUDITOR



Hypogram Presentation

A *hypogram* (*hypo*thesis dia*gram*) conveys in graphic form (1) the structural relationships among constructs employed in a model, and (2) the results of multiple tests of significance. This form of presentation conveys a large volume of information in a manner that takes advantage of the power of visual perception to make salient *patterns* among data which for certain purposes are more important than are the data themselves. A hypogram consists of the following components: (1) an organizational or logical structure that relates each construct employed by a model for which hypotheses are proposed; (2) for each such construct, a grid representing all comparisons made and tested; (3) an indication of the outcome of each such test.

In the present instance, the structure of each hypogram parallels that of the hierarchy presented as Figure 1, with some slight modifications to enhance the objectives of hypogram presentation. First, group between-strategies and within strategy-effects have been placed in close physical proximity to each other to facilitate visual comparison of g-complementary outcomes. Second, the base auditor's strategic emphasis, which, as discussed above, is only indirectly related to the hypotheses concerning strategic emphasis, is omitted. All statistical tests for significance are two-tailed at a confidence interval of 90%, based on the bootstrapped tables shown in Appendix D.

Summary of Findings

Two series of tests are proposed in the previous section, all related to the basic issue of typicality and inter-auditor variation in expert-like task behavior. The outcomes from the first series, in which each auditor is compared with the typical auditor, is presented as Figure 2. The outcomes of the second series, in which each auditor is compared with each other auditor, is presented as Figure 3. In Figure 3, strong complementary effects are represented by the below-diagonal triangular portion of the four-by-four outcomes grid.[45] Weak complementary and non-complementary findings are represented by a complete grid.

The disparity in comparative assessments is readily apparent in the salience and pervasiveness of shaded boxes in these hypograms. When compared against typical task behavior, each auditor's apparent task expertise does not vary significantly from what would be expected based on the performance of the group as a whole. However, when each auditor is compared pair-wise against each other auditor, numerous significant and complementary differences are uncovered. Even allowing for the fact that the number of comparisons tested in each hypogram differs (four comparisons against one standard in Figure 2 vs. six, and in some cases twelve, different pairings in Figure 3), the findings show that, conditional on the nature of the comparisons made, considerable

45. In extending findings to auditor comparisons made in the reverse direction, it is necessary to reverse the signum of the outcome reported in the portion of the grid presented.

variation in assessments of task expertise can and does exist. Considered individually, these four auditors are not a homogenous group.

Two interesting features of this analysis are to be noted: one is a similarity and the other, a contrast. As to the similarity, significant differences exist in the effects related to task strategy in both the auditor vs. typical and auditor vs. auditor comparisons. Most prominent of these are significant between-strategies effects that occur in both analyses. As to the contrast, while no significant differences are found in the knowledge properties of each auditor compared with the properties of the typical auditor, there are many significant differences when each auditor is compared with each other auditor. These differences occur at lower levels in the hierarchy and in such a pattern that they offset one another, so that there is no over-all difference in the higher level metrics of knowledge accessibility, A, or knowledge utilization, E_x.

Although there are many specific findings that can be discussed, I would like to focus on two in particular that have bearing on the questions raised in the Introduction. First, by identifying and interpreting patterns of significant effects in the behavior observation data, it is possible to distinguish levels of expertise among auditors who, by conventional standards and even by observations of task automaticity at higher levels, would be considered as functionally equivalent in terms of performance, accomplishment, and potential. For example, in reference to Figure 3, auditors 1 and 2 differ significantly in knowledge organization (the amount of searching required to locate knowledge in memory) and content (the successful outcome of a knowledge search), but not in availability (the automatic access to successfully located knowledge.). Note that Auditor 2 required less searching of memory and had greater success in evoking required knowledge than did Auditor 1. These differences are consistent with an interpretation that Auditor 2 is more knowledgeable in this task than is Auditor 1. The preceding conclusion is further supported by observing that, compared with Auditor 2, Auditor 1's task automaticity is significantly increased by an emphasis on information acquisition rather than environmental transformation strategy, a relationship that is not inconsistent with an interpretation of a lower level of task understanding on the part of Auditor 1.[46] Findings such as these suggest a potential to differentiate the *quality* of these auditors' expertise rather than merely its *appearance* by employing a broader set of criteria than the unidimensional assessment provided by observed task automaticity.[47]

---

46. Stronger discrimination on this point would involve analysis of the temporal dimension of between-strategies effects, a matter not covered in this paper and whose elucidation is left for a future paper.

47. To illustrate, in Figure 3, a significant difference is noted in task automaticity between Auditors 2 and 1, with Auditor 2 exhibiting the greater automaticity. However, this finding does not indicate that the greater expert-like behavior of Auditor 2 is partially due to that auditor's more efficient knowledge organization and greater positive knowledge content, as discussed in the text. It is this additional information which, in part, opens up the potential to make the transition from "greater *expert-like* task behavior" to "greater *expertise* in the task." The latter entails a more richly based characterization than does the former.

Second, although by conventional criteria, the auditors participating in this study are a homogenous group — all of similar age, educational background, and work experience — the findings show that, depending on the standards against which each is compared, there are significant differences among them that potentially can have significant impact on evaluations affecting their careers and professional development. It is clear from a comparison of Figures 2 and 3 that the evaluations made of relative expertise based on observations of auditor behavior are strongly conditional on the circumstances surrounding the choice of a standard for making such assessments. Evaluations made against a typical standard or norm risk failing to elicit much information that may be relevant with respect to the expert performance of individuals. Aggregation of data, a procedure whose major benefit for statistical metrics is minimization of random noise,[48] in some cases actually subverts that procedure's intent by submerging potentially valuable information in typical metrics and analyses based on high level phenomena. As shown by the findings reported here, assessments of comparative auditor expertise is one of those cases. This loss of information occurs because the variations in the properties that contribute to perceptions of task automaticity are not noise but real data. The fact that these variations may offset in producing typical measures and norms and in higher level phenomena, as they can be seen to do in Figures 2 and 3, is not justification for concluding that they can be ignored; they have explanatory and predictive value. These hypograms also show the potential for error in naively assuming that finding a significant relationship (or an absence of one) at a high level necessarily implies the presence (or absence) of significant underlying contributory relationships at lower levels (e.g., note the relationships below effective knowledge accessibility in Figure 3.)

Evaluating Directional Effects

The hypogram presented as Figure 3 is constructed with individual auditors as reference bases. With the exception of the strategy effects, the same conclusions regarding task automaticity and the properties of knowledge accessibility would have been drawn had the comparisons been made with Subject and base auditor assignments reversed because all such effects are strongly complementary. Consequently, the omitted above-diagonal outcomes would be the mirror image of those shown, but of opposite signum. However, note the almost perfect symmetry of the between-strategies effects in Figure 3. Since between-strategies effects are only weakly complementary, if this analysis were to be carried out with auditor assignments reversed or at confidence levels other than that used in this analysis (i.e., 90 percent on a two-tail test), the findings with respect to these effects might differ. Also, since within-strategy effects are non-complementary, findings regarding these effects may also differ. For this reason, the full four-by-four grids for weak and non-complementary effects are included in Figure 3.

48. Random noise is generally assumed to be normally distributed, with a mean of zero and constant variance.

DISCUSSION AND IMPLICATIONS FOR FUTURE RESEARCH

This paper explores auditor expertise by means of a model based on behavior automaticity, a widely recognized indicator of expertise in the performance of a task. It builds on and extends previous work by Russo (see references) in three principle ways. First, the model provides quantitative measures of the effects on task automaticity due to differences between auditors in the accessibility of their knowledge, its rate of utilization in a task, and the particular mix of behaviors that characterize their task strategies. Second, the underlying phenomena that determine how relative expertise in a task is perceived are presented as a hierarchy that, in classic reductionist fashion, explains higher level perceptions in terms of lower level phenomena, patterns among which can lead to a richer characterization of relative expertise than is permitted by use of the unidimensional criterion of behavior automaticity. Finally, specific aspects of the observer-observed system through which assessments of behavior are made are identified, together with guidance for quantifying and incorporating them into expertise evaluations. In particular, the submersion of information through aggregation in assessments based on norms and higher level phenomena; strong, weak, and non-complementary phenomena; and the assurances and cautions that are associated with each of these are discussed and illustrated.

The research presented herein differs from the usual practice in behavioral auditing research in a number of respects. First, this research is focused upon the behaviors of individual auditors rather than upon any specific group of auditors. Second, behavior observation methodology is employed in conjunction with widely used concurrent verbal protocol methodology. Third, the properties of knowledge and strategic behavior examined are context independent in that they are not conditional on the substantive content of the task. Finally, a new methodology, hypogram presentation, is introduced as a means of rapidly and intuitively making salient the outcomes of a hierarchically structured array of hypotheses.

Conclusions

In the Introduction several questions are posed that motivate this research. Answers to these questions can now be addressed in specific terms. Our varying perceptions of auditor task expertise are the products of observer-observed systems. As to the observer portion of this system, the very notion of expertise is observer dependent. Russo (1997b) points out that the study of expertise is a comparative exercise. The model presented in this paper instantiates that conceptualization and shows how issues of observer choice as to (1) recognized behaviors, (2) intentional and purposeful categorization of behaviors, and (3) reference standards are involved in making comparative assessments of task behavior. In addition, observers must search for and recognize patterns in observed phenomena which are then subject to interpretation. As to the observed portion of the system, the model provides a road map for quantitatively exploring expertise that guides observers by identifying places to look and pointing out relationships of interest that can reveal a broader vista for assessing the qualities of expertise than is revealed by a unidimensional assessment based simply on relative task automaticity. The major checkpoints on this road map are knowledge accessibility, knowledge utilization, and strategic behavior.

The findings reported in this paper show that it is possible, perhaps even likely, that auditors displaying the same degree of expert-like task behavior are actually quite different in terms of the underlying properties of knowledge accessibility and the expression of that knowledge as task strategy. Differences in these properties may offset at higher levels of perception, hiding them from an observer's view. However, consideration of the underlying differences have potentially important implications for the careers (e.g., task assignments) and professional development (e.g., training) of the individuals involved. Finally, the model indicates, and the findings illustrate, that the naive choice of typical metrics for use in evaluations of individual auditor performances submerges information that may be of significance for both the individual and the organization whose success depends to some extent on that individual's performance. This danger exists even if a salient standard is available, but is especially important to consider when individuals are evaluated one against another. The model shows that in the latter instances, choice of the reference auditor is not a significant consideration for most strongly complementary phenomena (*viz*, task automaticity, knowledge accessibility and its subordinate level phenomena), but can have a potentially significant impact on evaluations involving weakly complementary and non-complementary effects (primarily strategy-dependent effects.)

Implications for Future Research

In a review of expertise research, Bouman & Bradley (1997: 120) conclude that a great need exists to know more about the process by which expertise is acquired, and point to the acquisition of expertise as an attractive research opportunity. The model presented in this paper can serve as the foundation for an integrated pursuit of that opportunity. As a comprehensive exposition of the phenomenon of task behavior, it integrates both the cognitive and behavioral aspects of the problem solving process, the separation of which Russo (1997b) argues hinders progress in expertise research. The model's quantitative nature, individual focus, and relative context independence make it well suited to service as a research paradigm. For example, the effects of experience on learning can be studied by examining induced changes in knowledge-base properties; strategic behavior and task evolution can be studied through temporal analysis of phenomena related to intentional mix. Finally, in addition to its potential for illuminating problem solving processes and the acquisition of expertise, the model can be used as a foundation upon which to build computer simulations of auditor task behavior that can be valuable adjuncts for training and research. For example, model simulations can assist in evaluations of artificial intelligence models and the behavior of expert systems.

REFERENCES

Alba, J. W., and J. W. Hutchinson. 1987. Dimensions of consumer expertise. *Journal of Consumer Research* 13: 411-454.

Anderson, J. R. 1987. Skill acquisition: Compilation of weak-method problem solutions. *Psychological Review* 94 (2): 192-210.

Anderson, J. R. 1982. Acquisition of Cognitive Skill. *Psychological Review* 89 (4): 369-406.

Barlow, D., and M. Hersen. 1984. *Single Case Experimental Designs: Strategies for Studying Behavior Change (2nd ed.)*. Needham Heights, MA: Allyn & Bacon

Bedard, J. 1989. Expertise in auditing: Myth or reality. *Accounting Organizations and Society* 14 (2): 113-131.

Biddle, G. C., C. M. Bruton, and A. F. Siegel. 1990. Computer-intensive methods in auditing: Bootstrap difference and ratio estimation. *Auditing: A Journal of Practice and Theory* 9 (3): 92-114.

Biggs, S. F., T. J. Mock, and P. R. Watkins. 1988. Auditor's use of analytical review in audit program design. *The Accounting Review* LXIII (1): 148-161.

Biggs, S. F., and T. J. Mock. 1983. An investigation of auditor decision processes in the evaluation of internal controls and audit scope decisions. *Journal of Accounting Research* 21 (1): 234-255.

Cohen, J., 1960. A coefficient of agreement for nominal scales. *Educational and Psychological Measurement* XX (1): 37-46.

Cooper, J. O, T. E. Heron, and W. L. Heward. 1987. *Applied Behavior Analysis*. Columbus, OH: Merrill Publishing Company.

Davis, J. S., and I. Solomon. 1989. Experience, expertise, and expert-performance research in public accounting. *Journal of Accounting Literature* 8: 150-164.

Dennett, D. C. 1991. *Consciousness Explained*. New York: Little, Brown, and Company.

Ericsson, K. A., and H. A. Simon. 1993. *Protocol Analysis (Revised ed.)*. Cambridge, MA: The MIT Press.

Ellis, A., and R. S. Siegler. 1994. Development of problem solving. In Sternberg, R. J. (ed.). *Thinking and Problem Solving* (2nd ed). New York: Academic Press.

Ericsson, K. A., and H. A. Simon. 1993. *Protocol Analysis: Verbal Reports as Data*. Cambridge, MA: The MIT Press.

Hobson, J. A. 1994. *The Chemistry of Conscious States*. Toronto, Canada. Little, Brown, and Company (Canada) Limited.

Lonergan, B. J. F. 1978. *Insight: A Study of Human Understanding*. San Francisco, CA. Harper & Row.

Mayer, R. E. 1992. *Thinking, Problem Solving, and Cognition* (2nd ed.). New York: W. H. Freedman.

Newell, A., and H. A. Simon. 1972. *Human Problem Solving*. Englewood Cliffs NJ: Prentics-Hall.

Reber, A. S. 1985. *Dictionary of Psychology*. New York: Penguin Books USA.

Russo, J. A. 1998. Assessments of progress toward more expert-like behavior while performing audit tasks in the field. Paper presented at the 45th International Atlantic Economic Conference, Rome, Italy, March 1998. Sponsored by the International Atlantic Economic Society, Edwardsville, IL.

——. 1997a. Learning during audit tasks in the field: A model, methodology, and experiment. *International Advances in Economic Research*. 3 (4): 409-422.

——. 1997b. Behavioral research and auditor expertise: Time to broaden the focus. Paper presented at the Ninth Asian-Pacific Conference on International Accounting Issues held in Bangkok, Thailand, November 1997. Sponsored by The Sid Craig School of Business, California State University, Fresno.

——. 1996. Associations between knowledge states and environmental perceptions during performance of unfamiliar audit tasks. *Proceedings of the Eighth Asian-Pacific Conference on International Accounting Issues*. Vancouver, BC, Canada. Oct, 13-16. The Sid Craig School of Business, California State University, Fresno. 158-162.

——. 1995a. A model and methodology for studying the problem solving behavior of auditors in simulated auditing environments. *American Accounting Association Northeast Regional Meeting Collected Abstracts and Papers*. April 20-22. University of Hartford, Department of Accounting and Public Administration. 60-69.

——. 1995b. The availability of procedural knowledge in novice auditors. Paper presented at the 41st International Atlantic Economic Conference, Paris, France, October 1995. Sponsored by the International Atlantic Economic Society, Edwardsville, IL.

——. 1994 *An Investigation of Auditor Problem Solving Behavior In An Unfamiliar Task Situation*. Ph.D. Dissertation. Rutgers University, Newark. Ann Arbor: UMI International.

APPENDICES

APPENDIX A
RECOGNIZED OBSERVABLE TASK BEHAVIORS AND KNOWLEDGE BASE RESPONSES
(Target behaviors are shown **bold**)

OBSERVABLE BEHAVIOR		DESCRIPTION
Reading	*j=1*	In this experiment, reading documents in the task environment and answers to questions received via the CRT
Requesting	*j=2*	Requesting documents, information, and explanations from individuals in the task environment
Calculating	*j=3*	Either verifying a calculation or performing an original calculation
Writing	*j=4*	Writing a memo or workpaper (other than margin notes or underlining while reading or preparing the engagement report)
Other	*j=5*	Cross-referencing, indexing, or comparing documents; writing the draft of the report; organizing the engagement folder or the work area; searching the work area for a document, and discarding a document.
KNOWLEDGE BASE RESPONSE		
Automatic knowledge evocation	*k=1*	Any transition between two adjacent observable behaviors that is not mediated by cognition (*k=2* or *k=3*)
Analysis and planning	*k=2*	Subject states an objective or action he/she considers taking, states an assumption or draws a conclusion about the state of the task environment, summarizes for himself/herself personal knowledge of some aspect of the task environment
Uncertainty and confusion	*k=3*	Subject states a question or expresses uncertainty about specific entities, relationships, or processes in the task environment; expresses uncertainty about the task strategy, objectives, or how to proceed in the task; or expresses a general state of confusion.

<table>
<tr><th colspan="3">APPENDIX B</th></tr>
<tr><th colspan="3">BEHAVIOR ASSIGNMENTS TO INTENTIONAL GROUPS</th></tr>
<tr><th>GROUP</th><th>STRATEGIC CHARACTERIZATION</th><th>BEHAVIORS</th></tr>
<tr><td>g1</td><td>Information acquisition</td><td>Reading ($j = 1$)
Requesting ($j = 2$)</td></tr>
<tr><td>g2</td><td>Environmental transformation</td><td>Calculating ($j = 3$)
Writing ($j = 4$)
Other ($j = 5$)</td></tr>
</table>

APPENDIX C
TABLE C-1
EXPERIMENTAL OUTCOMES
EACH AUDITOR vs. "TYPICAL" AUDITOR

Auditor	Relative Task Automaticity (E)	Knowledge Accessibility Effect (E_a)	Strategic Knowledge Effect (E_m)
1	.867	.989	.911*
2	1.172	1.137	1.002
3	.996	1.019	.975
4	.921	.955	.967

KNOWLEDGE UTILIZATION AND PROPERTIES

Auditor	Knowledge Utilization (E_x)	Knowledge Accessibility (A)	Knowledge Organization (A_s)	Knowledge Content (A_c)	Knowledge Availability (A_e)
1	.963	.989	1.053	.904	1.039
2	1.029	1.137	.916	1.034	1.200
3	1.003	1.019	.948	.993	1.082
4	.997	.955	1.120	1.042	.818

COMPONENTS OF TASK STRATEGY EFFECT

Auditor	Between Strategies Effects		Within Strategy Effects	
	Information Acquisition Emphasis (D_{g1}^{XN})	Environmental Transformation Emphasis (D_{g2}^{XN})	Information Acquisition Strategy ($D_{g1(j)}^{N}$)	Environmental Transformation Strategy ($D_{g2(j)}^{N}$)
1	.116*	-.171*	-.028	-.006
2	.020	-.019	.011	-.011
3	-.106*	.082*	.001	-.001
4	-.065	.036	.027	-.031*

* H_0 rejected, $p \leq .10$, 2-tails

APPENDIX C
TABLE C-2
TASK AUTOMATICITY
OBSERVED METRICS: AUDITOR vs. AUDITOR

Auditor Pair	Task Automaticity (E)		Knowledge Accessabilty (E_a)		Strategic Knowledge (E_m)	
1 vs. 2	0.740	*	0.831		0.891	
1 vs. 3	0.870		0.969		0.899	*
1 vs. 4	0.942		1.095		0.860	*
2 vs. 1	1.352	*	1.316	*	1.028	
2 vs. 3	1.176		1.175		1.000	
2 vs. 4	1.273	*	1.304	*	0.976	
3 vs. 1	1.150		1.185		0.970	
3 vs. 2	0.850		0.896		0.948	
3 vs. 4	1.082		1.101		0.982	
4 vs. 1	1.062		1.134		0.937	
4 vs. 2	0.786	*	0.868		0.937	*
4 vs. 3	0.924		0.922		1.002	

Note: * Significant difference, p<.10 (2-tail)

APPENDIX C
TABLE C-3
KNOWLEDGE PROPERTIES AND UTILIZATION
OBSERVED METRICS: AUDITOR vs. AUDITOR

Auditor Pair	Organization (A_s)		Content (A_c)		Availability (A_e)		Accessibility (A)		Knowledge Utilization (E_x)	
1 vs. 2	1.149	*	0.875	*	0.866		0.871		0.954	
1 vs. 3	1.110		0.911		0.960		0.971		0.998	
1 vs. 4	0.940		0.867		1.270		1.035		1.058	
2 vs. 1	0.871	*	1.143	*	1.155		1.150		1.144	
2 vs. 3	0.966		1.041		1.109		1.115		1.054	
2 vs. 4	0.818	*	0.992		1.467	*	1.190		1.096	
3 vs. 1	0.901		1.098		1.041		1.030		1.150	
3 vs. 2	1.035		0.960		0.902		0.896		1.000	
3 vs. 4	0.847	*	0.952		1.322	*	1.066		1.033	
4 vs. 1	1.064		1.153		0.788		0.967		1.173	
4 vs. 2	1.222	*	1.009		0.682	*	0.841		0.997	
4 vs. 3	1.181	*	1.050		0.756	*	0.937		0.984	

Note: * Significant difference, $p < 10\%$ (2-tail)

APPENDIX C
TABLE C-4
STRATEGIC KNOWLEDGE DIFFERENCES
OBSERVED METRICS: AUDITOR vs. AUDITOR

Auditor Pair	Information Acquisition				Environmental Transformation			
	Between Strategies (D_{g1}^{XN})		Within Strategy ($D_{g1(j)}^{N}$)		Between Strategies (D_{g2}^{XN})		Within Strategy ($D_{g1(j)}^{N}$)	
1 vs. 2	0.097	*	-0.041		-0.144	*	-0.021	
1 vs. 3	0.193	*	-0.022		-0.286	*	0.014	
1 vs. 4	0.152	*	-0.046		-0.225	*	-0.021	
2 vs. 1	-0.128	*	0.043		0.117	*	-0.004	
2 vs. 3	0.121	*	0.010		-0.110	*	-0.021	
2 vs. 4	0.075		-0.009		-0.069	*	-0.021	
3 vs. 1	-0.257	*	0.033		0.200	*	-0.006	
3 vs. 2	-0.124	*	-0.015		0.096	*	-0.010	
3 vs. 4	0.046		-0.025		0.038		0.018	
4 vs. 1	-0.225	*	0.065		0.125	*	0.028	
4 vs. 2	0.086		0.013		0.048		-0.037	*
4 vs. 3	0.052		0.024		-0.029		-0.046	*

Note: * Significant difference, p<.10 (2-tail)

APPENDIX D

TABLE D-1

SELECTED POINTS ON BOOTSTRAPPED PROBABILITY DISTRIBUTIONS

ANY AUDITOR vs. TYPICAL AUDITOR

(Based on 10,000 simulations)

Cumuative Probability	Task Automaticity and Major Components			Knowledge Utilization and Accessibility					Strategic Knowledge			
									Between Strategies		Within Strategy	
	E	E_a	E_m	E_x	A	A_s	A_c	A_e	D_{g1}^{XN}	D_{g2}^{XN}	$D_{g1(j)}^{N}$	$D_{g2(j)}^{N}$
5%	.818	.819	.946	.863	.786	.879	.898	.757	-.085	-.075	-.048	-.024
95%	1.223	1.218	1.057	1.159	1.263	1.139	1.112	1.324	.089	.074	.049	.024

APPENDIX D
TABLE D-2
SELECTED POINTS ON BOOTSTRAPPED PROBABILITY DISTRIBUTIONS
AUDITOR vs. AUDITOR
(Based on 10,000 simulations)

Auditor Pair and Cumulative Probability		Task Automaticity and Major Components			Knowledge Utilization and Accessibility					Strategic Knowledge			
										Between Strategies		Within Strategy	
		E	E_a	E_m	E_x	A	A_s	A_c	A_e	D_{g1}^{XN}	D_{g2}^{XN}	$D_{g1(j)}^{N}$	$D_{g2(j)}^{N}$
1 vs. 2 2 vs. 1	5%	0.826	0.826	0.948	0.848	0.786	0.879	0.878	0.757	-0.073	-0.084	-0.047	-0.024
	95%	1.209	1.204	1.056	1.176	1.266	1.143	1.140	1.313	0.073	0.084	0.048	0.025
1 vs. 3 3 vs. 1	5%	0.817	0.818	0.946	0.864	0.796	0.878	0.889	0.764	-0.080	-0.079	-0.048	-0.025
	95%	1.221	1.218	1.057	1.152	1.262	1.139	1.127	1.304	0.076	0.083	0.049	0.026
1 vs. 4 4 vs. 1	5%	0.787	0.787	0.933	0.799	0.727	0.851	0.864	0.696	-0.088	-0.083	-0.057	-0.039
	95%	1.267	1.256	1.073	1.254	1.366	1.179	1.159	1.421	0.090	0.080	0.064	0.039
2 vs. 3 3 vs. 2	5%	0.836	0.843	0.954	0.898	0.823	0.898	0.910	0.797	-0.082	-0.070	-0.043	-0.023
	95%	1.188	1.183	1.048	1.117	1.216	1.117	1.097	1.255	0.084	0.069	0.043	0.024
2 vs. 4 4 vs. 4	5%	0.822	0.827	0.943	0.854	0.785	0.878	0.900	0.749	-0.095	-0.071	-0.046	-0.025
	95%	1.218	1.213	1.061	1.169	1.279	1.138	1.113	1.334	0.095	0.072	0.048	0.025
3 vs. 4 4 vs. 3	5%	0.811	0.811	0.936	0.867	0.784	0.883	0.909	0.760	-0.098	-0.071	-0.047	-0.038
	95%	1.232	1.219	1.070	1.150	1.264	1.131	1.099	1.303	0.099	0.070	0.049	0.039

Listed below are some of the most recent publications issued by the Center for Applied Research. Apply to the Director, Center for Applied Research, for single copies. Associate Membership in the Center is also available ($25 annually) which entitles the subscriber to free copies of all new Center publications.

WORKING PAPERS

165 Susanne O'Callaghan, John P. Walker, and Jeri B. Ricketts
Conservative Response Bias in Assessing Internal Controls Over Operations, Financial Reporting, and Compliance Under SAS Nos. 55 and 78 Guidelines: September 1997

166 Roy J. Girasa
Hong Kong 1997: Constitutional and Legal Aspects of Its Transition: September 1997

167 Mary Ellen Oliverio and Bernard H. Newman
Investigation of an Alleged Audit Failure: The McKesson and Robbins Case: October 1997

168 Ira J. Morrow
Instrumentation for the Evaluation of Performance in Behavioral Simulations: November 1997

169 Joseph A. Russo, Jr.
Assessments of Progress Toward More Expert-Like Behavior While Performing Audit Tasks in the Field: December 1997

170 Gerald Wohl
Teaching the Statement of Cash Flows (Direct Method) in Intermediate Accounting: A Direct Approach: February 1998

171 Thomas J. Webster
American Depository Receipts, Listing, and Market Efficiency: Three Case Studies: May 1998

172 Kwang-Hyun Chung
Utility Firms' Accounting for Income Taxes: Preference and Adoption Decisions Under Regulatory Environment: May 1998

173 Robert A. Isaak
Cultural Resistance to Globalization: June 1998

174 Raymond H. Lopez, and Surendra K. Kaushik
Profitability of Credit Unions and Commercial Banks in the 1990s: A Comparative Analysis: June 1998

175 Alan B. Eisner, Quintus Jett, and Helaine J. Korn
Does Cyber David Have Paper Goliath on the Run? An Examination of Newcomer and Incumbent Entrants into a Subfield of the Magazine and Newspaper Publishing Industry: June 1998

CASE STUDIES

20 Robert R. Cangemi and Raymond H. Lòpez
The Battle for Paramount - A (1993): April 1998

21 Robert R. Cangemi and Raymond H. Lopez
The Battle for Paramount - B (1993): April 1998

22 Robert R. Cangemi and Raymond H. Lopez
The Battle for Paramount - C (1993): April 1998

MONOGRAPHS

Joseph A. Russo, Jr.
Inter-Auditor Variability in Expert-Like Task Behavior: A Model and an Application: November 1998

REPRINTS

109 Narendra C. Bhandari
Enron-Dabhol Power Project: Process and Problems: October 1997
(*Journal of the International Academy for Case Studies*, vol. 3, no. 2, 1997)

110 Joseph M. Pastore, Jr.
A Trustee's Core Obligation: Ensuring Economic Performance and Educational Integrity: December 1997
(*Trusteeship*, vol. 5, no. 5, Sept./Oct. 1997)

111 Susanne O'Callaghan, John Paul Walker, and J. Timothy Sale
Over and Under Reliance on Internal Controls: Neural Networks Versus External Auditors: February 1998
(*Artificial Intelligence in Accounting and Auditing*, vol. 4, 1998)

112 Robert Dennehy, David M. Boje, Grace Ann Rosile, and Debra J. Summers
Restorying Reengineering: Some Deconstructions and Postmodern Alternatives: April 1998
(Communication Research, vol. 24, no. 6, December 1997)

113 Samir M. El-Gazzar
Predisclosure Information and Institutional Ownership: A Cross-Sectional Examination of Market Revaluations During Earnings Announcement Periods: April 1998
(The Accounting Review, vol. 73, no. 1, January 1998)

114 Robert Isaak and Alexander Keck
The Limits of EMAS: From the Risk-Reducing Certification of German Enterprises to Risk-Provoking, System-Transforming Ecopreneurship: June 1998
(UmweltWirtschafts Forum, vol. 5, no. 3, September 1997*)*

115 Robert Isaak
Globalization and Green Entrepreneurship: September 1998
(Greener Management International: The Journal of Corporate Environmental Strategy, vol. 18, Summer 1997*)*

116 Daniel Tinkelman, Claire Eckstein, and Peter Kyviakidis
Resolving Audit Issues in a High-Tech Environment: A Case Study: September 1998
(Issues in Accounting Education, vol. 13, no. 3, August 1998*)*

117 Samir M. El-Gazzar, and Philip M. Finn
Economic Implications of Proposed Accounting for Stock-Based Compensation: November 1998 (*Atlantic Economic Journal*, vol. 26, no. 3, September 1998)

CENTER FOR INTERNATIONAL BUSINESS STUDIES

Working Papers

5 Stephen Blank and Anne Taillandier
Atlantic Interdependencies and a Transatlantic Free Trade Agreement: January 1998