NOTICE

AVIS

ISBN 0-315-55451-7

THE UNIVERSITY OF ALBERTA

Determination of Probable Cause by Auditors: A Study of the Omission Effect in Fault Trees

BY

Richard Douglas Rennie ©

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES AND

RESEARCH IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR

THE DEGREE OF DOCTOR OF PHILOSOPHY

IN

ACCOUNTING

FACULTY OF BUSINESS

EDMONTON, ALBERTA

FALL 1989

THE UNIVERSITY OF ALBERTA

RELEASE FORM

NAME OF AUTHOR: Richard Douglas Rennie

TITLE OF THESIS: Determination of Probable Cause by Auditors: A Study of the Omission Effect in Fault Trees

DEGREE: Doctor of Philosophy

YEAR THIS DEGREE GRANTED: 1989

Richard D. Rennie

Edmonton, Alberta.

Date: August 22, 1989

THE UNIVERSITY OF ALBERTA

FACULTY OF GRADUATE STUDIES AND RESEARCH

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies and Research for acceptance, a thesis entitled: Determination of Probable Cause by Auditors: A Study of the Omission Effect in Fault Trees submitted by Richard Douglas Rennie in partial fulfilment of the requirements for the degree of Doctor of Philosophy in Accounting.

....
(Supervisor)

....

....

....

....

....

....

Date: Aug. 11, 1989

ABSTRACT

This dissertation provides experimental evidence concerning the cognitive processes by which auditors make assessments of the probable cause for the occurrence of an event. Previous studies in psychology using fault trees have shown that when subjects (including experts) are asked to estimate the probabilities of a partial list of possible causes plus all causes not listed, they underestimate the probabilities of all other causes not listed, relative to subjects who assessed a more complete list. This is called an omission effect. Prior studies have suggested that the omission effect may be due to availability of the listed causes or to output interference via part-list cuing.

The results of the first study indicated a significant omission effect for both expert and novice auditors. However, the omission effect was significantly less for the experts. There was no difference in omission effects between lists of causes which were of high and low availability (based on prior recall), suggesting that availability does not explain the omission effect.

The second study examined other explanations for the omission effect. When subjects were told that they would be sequentially assessing either three or six possible causes there was an omission effect. This result is consistent with anchoring-and-adjustment. That is, when subjects are asked to assess the probabilities of a list of possible causes, they may anchor on a probability estimate that is the reciprocal of the number of listed causes. They then increase or decrease their estimate from the anchor based on their perception of the underlying frequency. When large adjustments are needed (e.g. for low frequency causes), the adjustment may not be sufficient. This study also demonstrated an omission effect in the absence of output interference. Thus, output interference is not a necessary condition for the omission effect.

Although some expertise effects were found, experts also displayed omission effects. If similar omission effects also exist in actual audit judgments, these results suggest that auditors must attempt to obtain complete lists of possible causes prior to assessment of their relative probabilities. List completeness might be achieved by accumulating frequency data related to specific audit situations.

ACKNOWLEDGEMENT

I would like to express my thanks to my dissertation committee for their guidance in the completion of this research and to the external examiners for the time and effort they have expended on my behalf.

I would also like to thank the Chartered Accountants and students who participanted in this study and especially Ms. Orlanda Endicott and the Institute of Chartered Accountants of Saskatchewan for assistance in arranging participants for the study and in the running of the experiments.

Financial assistance during the tenure of my doctoral studies, for which I am grateful, was provided by the Social Sciences and Humanities Research Council of Canada, the University of Alberta, the University of Regina, and the Institute of Chartered Accountants of Saskatchewan.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

I. INTRODUCTION

Judgments of the likely (or probable) cause among alternative possible causes are an integral part of the professional judgment process in auditing. Consider the following hypothetical conversation that an audit manager might have with one of his staff assistants in a common audit judgment situation:

Manager: "I see from these draft financial statements that the current ratio has increased over last year. What do you think caused this?"

Assistant: "I spoke to the controller. She thought that this increase was likely caused by the higher levels of receivables this year due to slower collections. The chief accountant thought that the increase might be caused by errors in transferring balances when the new computer system was installed."

Manager: "Can you think of any other possible reasons? Maybe the marketable securities haven't been adjusted down to market values since the October crash. For the receivables, I think you should investigate the slow collections first.

> That sounds more likely. There may be a problem with bad debt expense. The computer conversion is less likely since we looked at the transfer in the interim audit."

...

The purpose of this research is to examine how auditors make assessments of probable cause. An experimental approach to the study of auditors' cognitive processes is adopted since, as noted by Nisbett and Wilson (1977), people are generally inaccurate at giving causal reasons for their behaviors. Gibbins (1984) also proposes that fully conscious professional judgment by auditors is infrequent. Therefore cognitive processes must be inferred through controlled experimentation. The need for more research into the nature of expert judgment in auditing was suggested by Kinney (1988).

This research is consistent with the first stage of a program of research into expert judgment by auditors outlined by Waller and Jiambalvo (1984). Waller and Jiambalvo see a need for three types of studies: (1) laboratory experiments comparing auditors' judgments with predictions of normative models; (2) laboratory studies involving the analysis of concurrent protocols that attempt to simulate actual audit situations; and (3) studies that attempt to examine expert knowledge in natural settings.

This thesis is concerned with auditors' judgments of probable cause made in fault tree and similar representations. A fault tree is a hierarchical grouping of possible causes and subcauses for an event, together with their associated probabilities. Prior research in psychology (e.g. Fischhoff, Slovic and Lichtenstein, 1978) has shown that both experts and students underestimated the probability of missing branches of a fault tree. This underestimation is referred to here as an omission effect.

Probable Cause in Auditing

Kinney (1975a) presented a single person decision-theory model by which an auditor may combine evidence culminating in an opinion concerning a set of financial statements. The model assumed that there are two states of the world regarding a client's reported account balance under investigation: (s_1) the account is not, or (s_2) is materially misstated. The auditor may choose the actions of either (a_1) accepting or (a_2) rejecting the reported balance. From experience the auditor is assumed to be able to specify a probability distribution around the two states: $P(s_1)$ and $P(s_2)$ (where $P(s_2) = 1 - P(s_1)$). If the auditor takes action a_1 and the true state is s_1 or if he takes action a_2 and the true state is s_2 then he has made correct decisions.

The state-action pairs (s_1,a_2) and (s_2,a_1) involve errors with costs C_1 and C_2 respectively. The cost (C_1) of rejecting a correct balance (s_1,a_2) involves such items as the cost of unnecessary additional audit procedures or loss of client goodwill. The cost (C_2) of accepting an incorrect balance (s_2,a_1) includes expected potential lawsuits and loss of professional reputation.

The auditor's decision problem is to choose the optimal action a^* so as to minimize the expected loss given that action:

$$E(\text{Cost} \mid a^*) = \min \{ C_1 \times P(s_1) \; ; \; C_2 \times P(s_2) \}$$

If the auditor chooses a_2 and conditionally rejects the reported balance on the basis of current knowledge he has a second stage decision problem wherein he may choose to obtain more evidence through additional sampling (at a cost $[C_3 + C_4 \times n]$ where C_3 is a fixed cost component and C_4 is variable depending upon the size (n) of the sample). The auditor must then balance the additional information against this additional cost. Kinney assumes that the standard deviation of the audited value in the population is the same under either state and can be estimated through experience. Then, appealing to the central limit theorem (with large n) the normal distribution of the true population balance can be approximated and a critical region (K) can be constructed for rejection of the reported balance given the mean of the sample. For a given sample size n the auditor may choose a critical region $K^*|n$ to minimize expected total cost. The auditor may find the optimal sample size n^* iteratively by: (1) varying n; (2) finding $K^*|n$; (3) computing $E(Cost|a^*,n,K^*)$ over all n to find $E(Cost|a^*,n^*)$.

In a related paper, Kinney (1975b) applied this model to the combination of evidence from the evaluation of internal control, compliance testing and substantive tests (including analytical review and tests of details) in a sequential manner. Although Kinney's network model may be criticized for leaving out or combining some of the decisions related to the audit decision process (e.g. see Arens, Loebbecke and Lemon, 1984, pp. 352 - 353), Kinney's model is useful for illustrating some of the relationships among the major types of audit tests. Kinney's network and the decision points (d_i) are shown below:

```
Internal Control                        Substantive
   Evaluation                              Tests                Opinion
_____________________                 ________________        ________

                                        Analytical
Preliminary  Compliance --->             Review   -------->
Evaluation -->  Tests   d2                  |      d3
          |       |                         |                 Opinion
          | d1     ---------->  Tests of
            ------------------>  Details   ------->
```

After making a preliminary evaluation of internal control, the auditor must decide (d_1) whether or not to rely on the controls and test them for compliance or to proceed to tests of details. If compliance testing is undertaken the auditor must decide (d_2) whether or not to perform analytical review procedures. If analytical review procedures are conducted the auditor must decide (d_3) whether to do additional tests of details.

The decision-theory model may be extended to apply to this network by the auditor estimating $P(s_1)$ and the expected related costs at each decision point (d_1, d_2, d_3). The auditor is thus performing a Bayesian updating of the estimated probability that the financial statements are not materially misstated, given new sample information. This estimated probability and the related costs of rendering an opinion at this stage, or performing additional procedures, form the inputs to the decision problem as outlined above.

Kinney's model and other Bayesian models (e.g. see Sorensen (1969), Leslie, Teitlebaum and Anderson (1980)) have been criticized as being too simplistic or impracticable. Moriarity (1975, p. 33) notes that at the substantive testing stage of an audit investigation, the auditor does not merely tally errors in order to estimate the relevant population parameters as implied by Kinney's model. Instead, the auditor is looking for unanticipated errors. Upon finding one, the auditor will attempt to determine its probable cause in order to ascertain whether such an error is likely to recur in the population. Arens, Loebbecke and Lemon (1984, p. 439), C.I.C.A. (1980, pp. 38-39), Grobstein, Loeb and Neary (1985, p. 93) and Johnson and Jaenicke (1980, p. 120) all simlarly refer to the need to determine the cause of audit errors which have been discovered.

Determination of the probable cause for a control deviation in compliance testing may allow the auditor to more efficiently conduct the audit. That is, if there is a systematic reason for certain deviations, the auditor may choose to change the nature of some of the planned testing procedures. For example, assume that an unacceptable level of control deviations was found and that most of these deviations related to a period of time when the person normally performing the control function was on vacation. An overall unacceptable level of deviations would normally mean that the auditor cannot rely on that control, and must instead perform more time-consuming substantive tests (i.e. detailed testing of transactions). In this particular situation, however, the auditor could reason that the probable cause of the high level of control deviations found in the vacation period was the use of a temporary replacement person in the control position. The auditor might therefore choose to do sufficient substantive testing to determine the extent of monetary errors during the vacation period and rely on internal control for the balance of the year (thereby reducing total substantive testing). Since substantive testing is more expensive than compliance testing, (Arens, Loebbecke and Lemon, 1984, p. 340) this latter approach, which relies on probable cause, is likely to be more efficient.

The example above illustrated the use of probable cause in terms of reliance of internal control, in order to reduce the amount of substantive testing to be done and to increase the efficiency of the audit. Probable cause assessment may similarly be applied to the investigation of fluctuations found in analytical review, since analytical review is the least costly of all forms of audit testing. One could conceive of many other uses of probable cause because within each major classification of audit tests (i.e. compliance tests, tests of details, analytical review) there are often alternative forms and combinations of tests which may be conducted to provide the desired level of assurance. These alternative test procedures will have differing costs. Thus the use of probable cause to stratify the population of transactions is likely to have efficiency implications for the audit.

This research is concerned with the cognitive processes underlying expert auditors' assessments of probable causes. Thus the concept of expertise in audit judgments is a major concern. In these experiments audit partners and managers were used as expert subjects, while chartered accountancy students who had recently begun careers as auditors were used as the novice subjects. Therefore experience has been used as a surrogate for expertise. As a result, any failure to find expertise effects may be due to use of experience rather than a more refined definition of expertise. However, the use of audit partners and managers as experts is consistent with their being held out to the public as experts. There is therefore some external validity for the use of experience as a surrogate for expertise.

Both the novice and expert auditors used as subjects in the experiments came from 14 different offices. This should have had the effect of diversifying any firm-specific expertise or heteregeneity of prior beliefs across the various experimental conditions so that a more general conception of auditor expertise could be examined.

. . .

This thesis consists of two papers. The first paper examines probable cause assessments in auditing, and provides experimental evidence of: (1) an omission effect when auditors assessed an incomplete listing of possible causes and; (2) an expertise effect in audit partners' and managers' judgments of probable causes in fault trees. The first paper also provides evidence that availability (in the sense of ease of recall of items in the shorter list of possible causes) (Tversky and Kahneman, 1973) is not the mediating factor in the omission effect. Fischhoff, Slovic and Lichtenstein (1978) had speculated that the omission effect was due to the operation of an availability heuristic. The second paper explores various other explanations for the omission effect and concludes that the assignment of probabilities in fault trees is consistent with an anchoring-and-adjustment process. Following the second paper is a summary chapter which provides concluding remarks concerning the findings reported in the two papers and their relevance to the field of auditor judgment. The appendix provides a summary of the experimental design and the statistical method used to assess the significance of the results found.

This dissertation is written in the paper (rather than traditional) format. The first paper relates the omission effect to auditing and demonstrates the omission effect in auditor subjects. As well, it discusses expertise effects in auditing. It is intended that the first paper, which contributes primarily to an accounting or auditing audience, be submitted for publication to an accounting journal. The second paper has an appeal to a broader judgment and decision-making audience since the second paper provides an explanation for the omission effect. It is intended that a version of the second paper, together with the results from Rennie and Johnson (1988) be submitted for publication to a psychology journal. The division of this study into two papers is somewhat arbitrary, as both papers can be viewed as making potential contributions to the knowledge of auditor judgments and of judgment and decision-making more generally. As a result of using the paper format for presentation of this thesis, there is a considerable overlap between the literature reviews of the first and second papers. For this I apologize to the reader of this thesis and ask that the reader view the two papers as independent studies within the literatures for which they are intended.

REFERENCES

Arens, A. A., Loebbecke, J. K. & Lemon, W. M. (1984). Auditing: An integrated approach, 3rd Canadian edition. Scarborough, Canada: Prentice-Hall.

C.I.C.A. (1980). Extent of audit testing: A research study. Toronto, Canada: The Canadian Institute of Chartered Accountants.

Fischhoff, B., Slovic, P. & Lichtenstein, S. (1978). Fault trees: Sensitivity of estimated failure probabilities to problem representation. Journal of Experimental Psychology: Human Perception and Performance, 4, 330-344.

Gibbins, M. (1984). Propositions about the psychology of professional judgment in public accounting. Journal of Accounting Research, 22, 103-125.

Grobstein, M., Loeb, S. E. & Neary, R. D. (1985). Auditing: A risk analysis approach. Homewood IL: Irwin.

Johnson, K. P. & Jaenicke, H. R. (1980). Evaluating internal control. New York: Wiley.

Kinney, W. R., Jr. (1975a). A decision theory approach to the sampling problem in auditing. Journal of Accounting Research, 13, 117-132.

Kinney, W. R., Jr. (1975b). Decision theory aspects of internal control system design / compliance and substantive tests. Journal of Accounting Research, 13, supplement, 14-29.

Kinney, W. R., Jr. (1988). Attestation research opportunities: 1987. Contemporary Accounting Research, 4, 416-425.

Leslie, D. A., Teitlebaum, A. D. & Anderson, R.J. (1980). Dollar unit sampling: A practical guide for auditors. London: Pitman.

Moriarity, S. (1975). Discussion of decision theory aspects of internal control system design / compliance and substantive tests. Journal of Accounting Research, 13, supplement, 30-34.

Nisbett, R. E. & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. Psychological Review, 84, 231-259.

Rennie, R. D. & Johnson, R.D. (1988, October). Auditors' judgments of probable causes: Effects of availability, experience, focusing and omission. Presentation at ORSA/TIMS, Denver.

Sorensen, J. E. (1969). Bayesian analysis in auditing. The Accounting Review, 44, 555-561.

Tversky, A. & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. Cognitive Psychology, 4, 207-232.

Waller, W. S. & Jiambalvo, J. (1984). The use of normative models in human information processing research in accounting. Journal of Accounting Literature, 3, 201-226.

II. AUDITORS' ASSESSMENTS OF PROBABLE CAUSES IN FAULT TREES

This paper examined auditors' judgments of the probable cause for the occurrence of an event. Hierarchical lists of possible causes together with their probabilities of occurrence are called fault trees (Fischhoff, Slovic and Lichtenstein, 1978). Through studying the effect on auditors' judgments of the omission of possible causes, this study contributes to our understanding of cognitive processing and expertise in auditing. Libby (1981, p. 103) suggests that "...when combined with probability theory, fault trees ... show promise for providing more exact estimates of the probability of error."

Judgments of probable cause are an integral part of the audit process. For example, Anderson (1977, p. 325) notes that auditors must determine the nature and cause of "critical compliance deviations." Gibbins (1984) proposes that auditors reason causally. Kida (1984) found that auditors relied on data that provided a causal explanation to a greater extent than equally diagnostic noncausal data. Burgstahler and Jiambalvo (1986) found that auditors tended to isolate errors and not extend them to the population where the underlying cause was viewed as nonrecurring or unique. Waller and Felix (1989) presented a framework and evidence to support the use of causal reasoning in auditing. Thus the assessment of probable cause of errors appears to be a normal part of the audit process. This is consistent with Libby's (1981) review which indicated that judged causal relations are important for explanations of observed events.

Fischhoff, Slovic and Lichtenstein (1978) studied the probability judgments of expert mechanics and students in a listing of possible causes (fault tree) for a car not starting. They found that for both experts and novices the probability of "all other" possible causes in a short list was underestimated relative to that of the same items in a longer, more complete list. The "credibility" hypothesis, that expert mechanics would be more likely to ascertain what items were missing from the incomplete list, was not supported.

Because none of the long lists nor the short lists in Fischhoff et al.'s (1978) study and in this study can possibly have contained a complete listing of all possible causes, this study refrains from using the established (and possibly misleading) terminology of "full tree" and "pruned tree." In addition, the listing of possible causes is not called a fault tree because in this study it contains only items from a single level of the tree, and not a hierarchical listing of nested causes. Because Fischhoff et al. found similar omission effects in both single and hierarchical levels of fault trees, this study does not address the issue of whether the omission effect is different for different levels within a fault tree.

Fischhoff et al. (1978) speculated that this underestimation or omission effect in the incomplete listing was due to the lack of availability (Tversky and Kahneman, 1973, 1974) of the unspecified possible causes in the incomplete list; that is, "out of sight, out of mind." They did not, however, directly manipulate availability.

Libby (1985) examined the role of availability in the hypotheses generated by auditors in an analytical review task. He found that frequency and recency of experience affected ease of recall. Thus there was a relationship shown between audit experience and perceived frequency of actual errors. However, Libby's study also indicated possible bias (due to recency and salience) which availability may have on judgments of probable cause. As suggested by Waller and Felix (1984, p. 398-399), the professional auditor might not have any comparative cognitive advantage over a lay person in terms of information processing.

The assessment of probable cause from an incomplete list of possible causes is relevant to the practice of auditing. That is, when an error in internal control is discovered, or when a significant fluctuation in an expected relationship is found in analytical review, the auditor will first consult with the audit client's management to determine possible causes (Smith, 1983, p. 30). Due to limited knowledge, management may be unable to provide a complete list of possible causes. To the extent that the audit is part of the performance evaluation of management, a manager may also have incentives for not disclosing all possible causes. The auditor is therefore in a position of having to assess probable cause from a partial list of possible causes. If the auditor can discern the missing causes, the audit may be conducted more efficiently, as those causes which are most probable may be investigated first. Thus it is relevant whether or not the auditor can augment that partial list by adding, from accumulated knowledge and experience, additional possible causes.

This study examined the issue of expertise in auditors' judgments of probable causes in an analytical review task. Experienced audit partners and managers were used as experts since they are held out to be experts by their firms and by the profession. Chartered accountancy students who had little auditing experience were used as novice subjects. These novice and expert auditors assessed the likelihood of either a short or long list of possible causes (and all other causes not listed). This study also provided a test of the availability explanation for the omission effect, in that the shorter lists of possible causes were categorized according to their relative availability (based on Libby's (1985) results in a free recall task).

The practical applications of causal analysis in auditing can be depicted in terms of a hierarchical listing of possible causes (i.e. a fault tree). Libby (1981, p. 103) suggested that fault trees may be applied to analyzing internal control systems (e.g. flowcharts) and in compliance testing.

When analytical review procedures indicate a deviation in an expected relationship, the auditor will first discuss the fluctuation with the client's management (Smith, 1983, p. 30). This discussion may indicate management's views as to the possible causes. Although auditors may not simply accept management's explanation for the fluctuation, they are in a position of having to assess both the completeness of the list of possible causes provided in the discussions with management and the probable cause among possible causes. Alternatively, even if the auditor did not consult with management regarding the probable cause of an error or fluctuation, he may be unable to generate a complete list of possible causes due to lack of complete information about the client's business environment and accounting system. The auditor's situation therefore appears similar to the experimental situation of Fischhoff, Slovic and Lichtenstein (1978) where subjects were asked to assess probabilities of partial lists of possible causes.

An audit consists of a sequential process of gathering sufficient evidence upon which to base an opinion regarding a set of financial statements. The fact that the process is sequential does not alter the auditor's situation concerning the completeness of a list of possible causes. That is, even if the auditor may eventually determine and investigate all possible causes in sequence, the audit may be more efficiently conducted if all possible causes may be discerned early in the investigation process.

The notion of auditors assessing probable cause from an incomplete list of possible causes provided by management may help to relate the incentive aspects of auditing into decision-theory models of the audit process. Fellingham and Newman (1985) note that the single person decision-theory model of the audit process (cf. Kinney, 1975a,b) does not allow for the auditor's actions to influence the behavior of his client. As noted by Fellingham and Newman (1985), there may be incentive reasons for the client not to divulge all information to the auditor. This information asymmetry and moral hazard could potentially be modeled analytically in a three-person (principal-manager-auditor) agency framework. However, as noted by Scott (1984, p. 186), it is difficult, even with the assumption of no coalitions, to analytically show what the implications of a three-person agency relationship would be since the manager (client) and the auditor may interact with each other and the principal in many different ways. There is therefore no unique analytical solution to the three-person agency problem. For the purposes of this study, it is sufficient that the client, if he knows the probable cause for an audit event, might not reveal this information to the auditor.

Duh and Sunder (1985) viewed accounting as the intersection of economic incentives (such as moral hazard in the principal-agent framework or maximization of utility) and cognitive limitations. Without both there would be no demand for accounting information, and hence no demand for audit services. The relationship between assessments of probable causes in auditing and incentives was referred to above. However, cognitive limitations (in the sense of the impossibility of knowing all information) would also be sufficient to give rise to an auditing problem in the estimation of probable cause. That is, even if the client manager were willing to provide all information to the auditor concerning possible causes for an auditing event, he may not be able to due to his own incomplete knowledge. Alternatively, since the client's perspective or causal field on the matter is different, the client manager may have a different assessment of which possible cause is most probable.

Thus either for incentive reasons or for cognitive reasons, the auditor may not obtain a complete listing of possible causes from which to make his own assessment of probable cause. It is therefore an important question as to whether, in such circumstances, the auditor's assessment of probable cause is affected in a manner similar to the estimation of probabilities in Fischhoff, Slovic and Lichtenstein (1978). If auditors' assessments also underestimate the relative frequency of any causes that are not directly considered, then it is important to understand the processes and conditions producing the omission effect so as to develop means to mitigate its effects on auditors' judgments of probable causes.

The Omission Effect and Availability

The omission effect in fault trees or similar listings of possible causes has been shown for both experts and novices in a number of studies. Fischhoff et al. (1978) suggested that this effect might be due to lack of availability (Tversky and Kahneman, 1973) of the omitted causes; in other words, "out of sight, out of mind." A number of studies have attempted to test this availability hypothesis.

Tversky and Kahneman (1973) proposed that when people are asked to make probability judgments about the likelihood of events, they may do so based on the recall of memory of similar events, or the ease with which they may construct or imagine the occurrence of the event. An event will be judged to have a high likelihood if many similar events are recalled or if it is easy to imagine similar events. Tversky and Kahneman

(1973, footnote 3) differentiated their meaning of availability from the earlier usage of the term availability in the verbal learning literature. In the verbal learning literature (e.g. Tulving and Pearlstone, 1966), availability referred to the ability of subjects to recall cued information. In verbal learning, availability was often contrasted with accessibility which refers to the unaided recall of information.

In a hypothesis generation task involving automobile malfunction (similar to Fischhoff et al.'s (1978) task), Mehle (1982) examined verbal protocols of subjects. He found that both expert and novice subjects had difficulty generating complete sets of hypotheses. In fact, the experts apologized for not being able to generate many hypotheses. He also found that subjects were overconfident in their subjective judgments of the probabilities associated with the hypotheses that they did generate. Mehle's results may indicate that automobile trouble shooting is not a task where it is necessary to generate a complete set of hypotheses prior to investigation. That is, the assessment of possible malfunctions may be dominated by cost rather than probability considerations. This could explain why Fischhoff, Slovic and Lichtenstein (1978) found that expert mechanics displayed similar omission effects as did novices.

Another study (Mehle, Gettys, Manning, Baca and Fisher, 1981) examined an availability explanation for subjects' overconfidence in the probability of generated hypotheses. The results were mixed. In one experiment their study found that overconfidence was reduced by having subjects complete the unspecified set. Presumably this lowering of overconfidence resulted from an increase in availability. However, in a second experiment, when the unspecified set was completed by the experimenter (by computer), overconfidence persisted. Although having subjects view the completed set should also have increased availability of the unspecified hypotheses, the results of the second experiment differed from the first.

Gettys, Mehle and Fisher (1986) concluded that subjects may be overconfident in hypotheses they generate and underestimate unspecified hypotheses because they have difficulty recalling hypotheses which are not specified. Consistent with Fischhoff et al. (1978), Gettys et al. (1986, p. 32) suggested that this overconfidence may be partially explained by the operation of an availability heuristic.

Hoch (1984) examined availability and output interference in a hypothesis-generation task. He suggested that the omission effect in fault trees (as in Fischhoff et al., 1978) could be due to subjects using availability to judge probability but being subject to output interference via the part-list cuing procedure (Slamecka, 1968, Moser, 1989). The part-list cuing procedure refers to experimental subjects recalling fewer of the remainder of a list of words than control subjects when the experimental subjects are provided (or cued to) part of the list by the experimenter. Output interference is an explanation for the part-list cuing phenomenon which suggests that previously retrieved or experimenter-supplied items interfere with subjects' ability to recall the remaining list items.

Hoch's (1984) study examined the ability of subjects to construct events for which they had no direct past experience. Thus he used availability in the sense of construction or simulation of future events, rather than that of direct recall, which involves one's memory of past events (although both were referred to in the context of availability by Tversky and Kahneman, 1973). Hoch found evidence to support his joint hypothesis that availability (i.e. simulation) and output interference may result in subjects being unable to generate all possible hypotheses. However, he stated (p. 660) that interference is more likely to influence novel judgments. Thus the availability explanation for the omission effect in terms of recall of past events was not examined in Hoch's study.

The omission effect was replicated using student subjects by Hirt and Castellan (1988) in an adapted version of Fischhoff et al.'s (1978) "reasons for a car not starting" problem. They used a within-subjects design, with subjects assessing the probabilities of both a complete (five categories) and a partial list (four categories) of possible causes. This design allowed them to examine category redefinition of the omitted cause by subjects when the list was shortened. They provided evidence that subjects idiosyncratically redefined category membership in the shorter list. That is, there were no discernible patterns in the change of the assessed frequencies of each category between the five category and four category lists. This study may have had some internal validity problems because subjects were told that the complete and partial lists of possible causes were from separate automobile manufacturers. As a result, it is not clear that the partial list was in fact a partial list. Rather, it could have been treated as an independent shorter list for a separate automobile. Thus it may not be too surprising that subjects idiosyncratically recategorized the omitted item in the shortened list. Hirt and Castellan also attempted to test whether availability was the mediating factor in the omission effect. They defined availability in the sense of subjects being cued in the experiment itself concerning the omitted causes by performing the assessment task on the complete list of possible causes. This conception of availability is consistent with that from the verbal learning literature.

Dube-Rioux and Russo (1988) similarly attempted to examine the availability explanation for the omission effect in fault trees. Again availability was defined in a verbal learning sense as retrieval failure, or the inability of subjects to bring known events to mind. Dube-Rioux and Russo differentiated causes into short-term and long-term which they presumed were respectively easier or harder to recall. Although their results were mixed (p. 231), they claimed that their results were consistent with an availability bias.

Another recent study by Weber, Eisenfuhr and Von Winterfeldt (1988), again stated (but did not test) that availability may be the mediating factor for what is referred to in this paper as the omission effect.

In an auditing context, Rennie and Johnson (1988) demonstrated the omission effect using chartered accountancy students (finalists and novices) in a task involving estimating probable causes for discrepancies in accounts receivable confirmations. There was no expertise effect shown in that the omission effect was not significantly reduced for the finalists over that of the novices. The subjects were junior employees of accounting firms and the task may not have been related to an actual audit judgment task which would require use of probable cause. There was an expertise effect in terms of the assessment of high and low availability causes. The finalists were better able to distinguish the high and low availability events and made more exteme judgments of the frequencies of these events. This may be due to better knowledge of actual frequencies. However, this more extreme assessment of the frequencies by the finalists could also be due to overconfidence on the part of the finalists, or lack of confidence on behalf of the novices. As noted above (e.g. Mehle, 1982), subjects tend to be overconfident in the hypotheses they generate. However, it is not known how experience affects confidence judgments in these tasks.

Rennie and Johnson's (1988) results indicated an omission effect for chartered accountancy students in a particular task. It is uncertain how robust Rennie and Johnson's results are for expert auditors and for more general auditing tasks.

The Rennie and Johnson (1988) study found that availability (in the sense of recall) was not a meditating factor in the omission effect. Subjects in the short list conditions assessed list items which (based on a prior test and a concurrent free recall task) were assessed as being either of high or of low recallability. Subjects in the long list condition assessed the probabilities of both the high and the low recall items. In the comparison of the probability assessments from the high and low availability short lists to the corresponding assessments from the long list, however, Rennie and Johnson obtained both high and low availability long list assessments from the same subjects. That is, subjects in the long list condition provided responses for both the high and low availability conditions, while the high and low availability conditions for the short lists were from separate groups of subjects. There was therefore possible statistical error due to the lack of independence in responses.

It is important to determine if the omission effect extends to judgments of probable cause more like those actually made by auditors and to different levels of auditor expertise, because auditors may commonly assess the probable cause of incomplete listings of possible causes. Different levels of auditor expertise have different opportunity costs as reflected in different billing rates. Certain tasks relating to the determination of probable cause are performed at different levels. For example, Libby (1985) noted that analytical review procedures are usually conducted only by managers and partners. Because the market for auditing services is competitive, the level of the organization at which certain tasks are performed has efficiency implications for the auditing profession.

By examining the judgments of expert auditors, this research may provide insights regarding what is meant by expertise in different fields of endeavor. That is, the relative probability of each possible cause may have efficiency implications for the audit investigation. On the other hand, consistent with Mehle's (1982) and Fischhoff et al.'s results, automotive trouble shooting may be dominated by cost rather than probability considerations. Thus one might expect that expert auditors would have a lower omission effect than would novices, while expert mechanics might display a similar omission effect as novices. Thus in order to access auditor expertise, it may be necessary, in an experimental study, to use a task which has both a high degree of experimental realism but, as well, has sufficient mundane realism that expert auditors may be able to access their knowledge of frequency information regarding the auditing event used in the task.

This study extended the results of prior studies by:

(1) Using a group of subjects (audit partners and managers) who are held out to the public, by virtue of their positions in the profession, to be experts;

(2) Using an analytical review task which has face validity as a task upon which expert auditors would be familiar with assessments of probable cause. This should have enhanced experimental realism and may have allowed experts to access the memory structures associated with their expertise;

(3) Having independent groups of subjects assess each of the high and low availability long list and short list conditions.

(4) Having subjects provide a subjective assessment of confidence in their probability assessments.

Research Goals and Predicted Effects

The first research goal was to test the "credibility" hypothesis that experts would be better able than novices to ascertain causes that were missing from a list of possible causes. From prior research using automobile mechanics (Fischhoff et al., 1978), it was predicted that there would be an omission effect for both experts and novices. In the case of mechanics investigating reasons for a car not starting, however, it may be that they follow diagnostic checklists which may be dominated by cost rather than probability considerations. Thus expert mechanics may not need to consider probabilities. In an auditing context consideration of relative probabilities of possible causes (as discussed previously) directly affects the efficiency of the audit. It was therefore predicted that the omission effect would be less for audit partners and managers than it was for novices.

The second research goal was to test the availability explanation for the omission effect. The availability explanation suggests that there should be a smaller omission effect when subjects in the short list conditions assess low recall causes than when subjects assess high recall causes. This is because high recall causes should already be available to subjects. Thus cuing subjects to the low recall causes by specifically listing them should mean that they are better able to fully populate the sample space of possible causes than subjects who assess only high recall causes and are not cued.

Note that availability and ecological frequency tend to be positively correlated (Tversky and Kahneman, 1973). Thus a comparison of the omission effects from lists of differing availability may be confounded with the underlying frequency of those high and low availability list items. However, the expert subjects (audit partners and managers) should have a better knowledge of the underlying frequencies of auditing events. Thus the second research goal is likely (at least for the novices) to be a joint test of whether frequency and availability are mediating factors in the omission effect. For the experts, there is a purer test of the availability explanation since the experts should be more knowledgable about the underlying frequencies.

The third goal of the research related to expertise in knowledge of frequencies of audit events. In Rennie and Johnson's (1988) study the more experienced senior auditors were better able to discriminate high and low recall events than were the novices. This result may be due to better knowledge of the underlying frequencies of the events or it may reflect a greater confidence in the assessment process (cf. Gettys et al, 1986). This study examined the relative confidence the partner/manager and novice subject groups had in their probability assessments. An assessment of confidence can help to distinguish whether the expertise effects found were due to differing memory structures of the frequency of actual events (as would be suggested by an expertise explanation) or due to overconfidence. Overconfidence by experts in such judgments might be caused by incomplete outcome feedback. That is, although experts may perform a task frequently, they may make judgment errors which are only rarely brought to their attention (such as where there is a law suit following a business failure).

The experiment in the paper was a 2^3 factorial of EXPERTISE (novices vs. audit partner/managers), OMISSION (list length of three or six possible causes) and AVAILABILITY (high or low recall of the possible causes). The dependent variable (discussed later) was based on subjects' probability assessments of "all other causes" not listed. The results were analyzed by analysis of variance. From the predictions above, it was hypothesized (in alternative form) that:

(1) Both expert groups would display significant OMISSION effects. This was tested by the significance of the main effect for the OMISSION variable. As well, separate follow-up tests by both expert groups were done.

(2) The partner/manager expert group would display a lower omission effect than would the novices. This was assessed by examining the group means and the significance of the interaction of OMISSION and EXPERTISE.

(3) There would be a lower omission effect for the low recall possible causes than for the high recall causes. This hypothesis follows from the availability explanation for the omission effect. This hypothesis was examined by testing the significance of the OMISSION by AVAILABILITY interaction.

(4) Due to better knowledge of the underlying frequencies of the pssible causes, the partner/manager expert group would be better able to discriminate the frequencies of high and low availability events than would the novices. This was assessed by examining the significance of the EXPERTISE by AVAILABILITY interaction. Since the expert group was predicted to

have a better knowledge of the underlying frequencies of the possible causes, a significant EXPERTISE by AVAILABILITY by OMISSION interaction was expected (following from prediction (3)) if availability is a mediating factor in the omission effect.

(5) The partner/manager group would have greater overall CONFIDENCE in their judgments (due to greater familiarity). This was assessed by examining the main effect for EXPERTISE in an ANOVA using the confidence judgments as the dependent variable. However, in this separate ANOVA using confidence as the dependent variable, it was hypothesized that confidence would not interact with either availability or omission. Thus any expertise effects found were hypothesized to be consistent with knowledge of underlying frequencies (and not overconfidence).

METHOD

Subjects and Administration

The 130 experimental subjects were grouped in terms of audit experience. The expert subjects were 74 audit partners and managers from Chartered Accountancy firms in Regina and Saskatoon, Saskatchewan. The novice subjects were 56 students registered with the Institute of Chartered Accountants of Saskatchewan and who had commenced employment with Chartered Accountancy firms within the past six months. Some of these novices may have received a few months of additional auditing experience in work terms in the co-operative work-study program at the University of Regina.

All subjects were volunteers. The sessions for the novices were held as the last agenda item at two new students' days in Regina and Saskatoon. One student at each of the two sessions chose not to participate in the study.

The expert subjects were solicited by contacting the managing partners (or equivalent title) of 14 offices of Chartered Accountants. The 14 offices consisted of 12 international firms, 1 regional firm and 1 local firm. Each office contacted agreed to participate in the study and solicited volunteers. Where possible, the experiment was conducted in a group setting by office. In the group administrations of the experiment, no clarifying questions of substance were asked. Therefore, for offices which were not able to arrange a common time for the volunteers to meet, the experiment was self-administered by the subjects. These self-administered subjects were asked to spend the same time at the task as had the group subjects and were given written instructions based on the instruction script read to group subjects. The self-administered subjects were debriefed by using a sealed envelope to be opened after completing the experiment. Because there were approximately equal numbers of group-administered (n=33) and self-administered (n=41) experts in each experimental condition, a separate ANOVA was performed on responses from the expert group with each expert subject coded as to type of administration. Neither the main effect for administration condition nor any interactions of administration condition and the other treatment effects were significant.

One result of using a self-administration procedure was that an exact prior count on expected numbers of subjects could not be achieved. As a result, the cell sizes, although very similar (refer to the Appendix), were not equal.

The expert auditor group consisted of audit partners and managers. Promotion to audit manager varies by firm with some firms having an intermediate rank of supervisor or an advanced rank of senior manager. When subjects were solicited, care was taken to ensure that a common definition of manager was used which included junior managers or supervisors as well as senior managers. As well, only those subjects who spent substantially all of their time in auditing were requested to volunteer.

The novices had a mean age of 23.8 years and mean auditing experience of 0.5 years. The expert group had a mean age of 33.4 years and mean audit experience of 11.9 years.

The responses for five subjects were eliminated due to minor addition errors by subjects in their completion of the task. Similar addition errors have been reported in other studies (e.g. Dube-Rioux and Russo, 1988). Dube-Rioux and Russo (1988) chose to retain such responses and normalize them to 100%, thereby allocating the error to each listed cause. Since addition errors can occur in both short and long lists such an allocation may arbitrarily increase the reported omission effect; thus normalizing the responses may not deal with the error in an unbiased manner. That is, these small addition errors could also represent frequencies which should be part of "all other causes" (which would not affect the omission effect).

As well, the responses for four subjects classified as experts were deleted because the subjects indicated that they were audit seniors, rather than managers. The final sample size was 121 consisting of

67 experts and 54 novices.

Materials

The experimental case was adapted from Libby (1985). Subjects were asked to read a narrative which provided information about a manufacturing company's accounting policies and listed the comparative results of three financial ratios (gross margin, current ratio and quick ratio) from the prior year's audited financial statements and the current year's unaudited accounts. Subjects were then asked to assess the likelihood (i.e. probability) of a list of possible causes for the fluctuation in the ratios. The narrative indicated that the possible causes had been provided by the client's staff in order that subjects would not assume that the list was exhaustive. Subjects assessed the relative frequency of each listed cause and all other possible causes not listed, on a 0 - 100 point scale. Subjects were reminded in the written instructions that their totals should equal 100. Following the probability assessment task, subjects were also asked to assess, on a 0 - 10 point scale, their confidence in the numerical assessments that they had made.

The six possible causes for the fluctuation in the ratios in the case were as follows:

1. Bad debt expense is underrecorded or not recorded.

2. Current period purchases on account have not been recorded or are recorded next period.

3. Current portion of long-term debt improperly classified as noncurrent.

4. Goods returned by customer in the current period but not recorded or recorded next period.

5. Purchase returns recorded but goods not returned or returned next period.

6. Payments on account recorded but not made or overrecording of such payments.

The narrative and possible causes were taken from Libby's (1985) results (in a free recall task) so that the first three causes (above) were high in recall and the last three causes were low in recall.

The text of the experimental narrative and instructions was as follows:

<u>*EAZ MANUFACTURING CO.*</u>

As part of the analytical review in connection with the audit of EAZ manufacturing Co., Linda Smith, C.A. compared the gross margin percentage, the current ratio and the quick ratio from the current year's unaudited financial statements with those of last year's audited statements. These ratios are shown below. Assume that any fluctuation in these ratios is due to a single mistake or multiple occurrences of the same mistake. The accounting practices followed by EAZ indicate:

-sales:

1. all sales are made on credit and the credit policy has not changed since last year.
2. prices have not changed since last year.
3. sales discounts are not material.
4. the allowance method is used for recognizing losses from uncollectible accounts; bad debt expense is an operating expense.

-inventories:

1. a perpetual inventory system is used and inventory costing methods have not changed since last year.
2. vendor's prices have not changed since last year.
3. all purchases are made on credit.
4. purchase discounts are not material.

Financial Ratios

Ratio	Prior Year Audited	Current Year Unaudited
GROSS MARGIN: (Net Sales - Cost of Goods Sold) / Net Sales	26.1%	26.2%
CURRENT RATIO: Current Assets / Current Liabilities	2.43	2.72
QUICK RATIO: Cash, Securities, Net Receivables / Current Liabilities	1.04	1.25

Linda Smith has discussed the fluctuation in these ratios with various members of EAZ's accounting staff. While these individuals could not state a precise cause for the fluctuation, they speculated that the cause could be due to any of the following possible causes listed below.

Required: Your task is to determine what the probable cause for the fluctuation in the ratios might be. Recall that the fluctuation has been caused by a single mistake or multiple occurrences of that single mistake. The possible causes provided by the client's staff are listed below. Following the listed causes are blank *lines for you to list possible causes not provided by the client's staff. From your experience and knowledge, please rate each of these possible causes by the likelihood (from 0 - 100) that you think it may have caused the fluctuation in EAZ's financial ratios. Think of this 100 point scale as being 100 separate occurrences of this same financial ratio fluctuation. Your total should add to 100 for the listed causes and any additional causes that you list, since you have listed and estimated the likelihood of any causes not provided by the client's staff.*

The narrative and the listed causes were on facing pages of the booklet so that subjects could review the narrative material in making their assessments. The instruction heading beside the listed causes read: "Estimated # of times out of 100 that ratio fluctuations would result from this possible cause."

The design of the experiment was a 2^3 factorial of expertise (novices or partner/managers), omission (3 or 6 possible causes listed) and availability (high or low recall). Subjects were randomly assigned to the experimental conditions. For a summary of the design and the cell sizes, please refer to the appendix.

There were three versions of the instrument corresponding to the long list (three high and three low recall possible causes), the high recall list (three causes), and the low recall list (three causes). On the instruments, the order of the items was randomized. Note that subjects in both high and low recall long list conditions used the same experimental materials.

Because subjects in the short list conditions assessed the likelihoods of different causes, direct comparisons of probability assessments on individual causes between these groups is not possible. The dependent measure for the probability assessment task (similar to Fischhoff et al., 1978) was a measure of the relative frequency (out of 100) for "other causes" for the condition of interest as follows:

Condition (high or low recall/ short or long list)	**Dependent Measure** "Adjusted All Other Causes"
high / short	100 - frequencies for high causes
low / short	100 - frequencies for low causes
high / long	100 - frequencies for high causes
low / long	100 - frequencies for low causes

Note that in all conditions the dependent measure is one hundred minus the assessed frequencies for the possible causes related to the applicable condition. If a subject in the high availability / short list condition assessed the listed causes to have a total frequency of 60 and "all other" causes to have a frequency of 40, then the subject would have a dependent measure of 40. Thus for the short list conditions, the dependent measure is equal to the assessed frequency for "all other causes." If a subject in the high availability / long list condition assessed the high recall causes to have a total frequency of 70, the low recall causes to have a total frequency of 25 and "all other causes" to have a frequency of 5, then the subject would have a dependent measure of 30. For the long list high (low) availability condition, the dependent measure treats low (high) availability causes as part of "all other causes." This dependent variable is referred to in this paper as "adjusted all other causes."

Although subjects in the long list condition responded to both the high and low availability lists, only one dependent measure was calculated for each subject. That is, subjects in the long list condition were randomly assigned to either the high or low availability conditions.

Note that the dependent measure could be viewed as partially confounding the effect of relative availability, in that subjects in the long list conditions assessed both high and low availability possible causes. The intention of the measure is to examine the probability assessments of high and low availability causes in both long and short lists of possible causes. There is no confounding, however, in that the difference between the long and short lists is that subjects in the long list also assessed the frequencies of additional causes, those being the ones of interest. This is what the "adjusted all other causes" measure reflects.

This experiment was administered together with four related auditor judgment experiments. The order of this experiment and another one adapted from Rennie and Johnson (1988), which also involved fault trees, was counterbalanced with this experiment being run either first or second (variable COUNTERBALANCE). No significant order effects or interactions were found (see Table 2-1). The total time taken by subjects for the complete set of experiments was about 45 to 50 minutes.

The instrument was pilot tested on 55 C.A. finalists who were attending a computer auditing course and expected to write their uniform final examinations in about two months. Although the instrument (apart from a minor variation in the instructions) did not change for the administration to the novices and experts, the finalist groups' responses were not included as a separate level of expertise. This is because approximately 80% of the finalists had participated in the Rennie and Johnson (1988) study as novices about one year prior to this study. Thus these subjects had been debriefed concerning the omission effect and the availability explanation for the effect and may have been cued to the omission effect and one of the hypotheses of interest.

The responses for this pilot group were consistent with prior knowledge of the hypotheses. When the overall results were analyzed including the finalists as a separate level of expertise, there was a significant Counterbalance by Expertise interaction ($p = .036$). A follow-up analysis of the expertise groups indicated that this interaction was due to the finalist subjects' responses. The finalist group results indicated a significant main effect for the order of the counterbalanced experiments ($p = .012$). These pilot subjects (i.e. the finalists) who assessed this experiment first, displayed a significant omission effect ($p < .001$) and a significant omission by availability interaction ($p = .005$). The pilot subjects who assessed this experiment following the experiment based on Rennie and Johnson (1988) (which most had participated in) displayed no significant omission effect ($p = .056$) and no significant omission by availability interaction ($p = .671$). The debriefing in the Rennie and Johnson (1988) study had discussed both omission effects and the availability explanation.

The nonsignificant omission effect for pilot subjects who assessed this task after the similar task from the previous year may indicate a learning effect. That is, when cued to the nature of the task, these subjects displayed no omission effect. This result may indicate that through direct experience with the omission effect and with proper cuing, the omission effect can be reduced. These results could also indicate willingness of the subjects to co-operate with the experimenter. Since these pilot test finalist subjects had participated in the prior experiment and this participation appeared to have affected their responses, the responses of the finalists in the pilot test were not included in the analysis.

Thus, the pilot test served its purpose of providing a review of the experimental materials. The comments of the subjects, who may have had some knowledge of the hypotheses of interest, provided some assurance regarding their understanding of, and the relevance of, the experimental task.

RESULTS

The results were analyzed by analysis of variance (see Table 2-1). To test the assumption of homogeneity of group variances, Bartlett's test was performed and the homogeneity assumption could not be rejected at conventional levels ($p = .054$). To test the robustness of the effects the ANOVA was also conducted on ranked data (Conover, 1980, p. 337). The significance of the effects was the same as those of the parametric ANOVA. Normality of the model's residuals was tested using Lilliefor's test. Normality could not be rejected ($p = .643$).

The main effect for length of list (omission) is shown in Figure 2-1. This figure indicates that, overall, subjects underestimated the probabilities of "all other causes" in the short list (relative to the full list). That is, in the short list, the mean frequency assessment of other causes (out of 100) was 33.5 while the corresponding amount in the full list condition was 56.0. This effect was significant ($F(1,105) = 24.1$; $p < .001$). This result was consistent with the first hypothesis.

The "credibility" hypothesis stated that this omission effect should be less for experts than for novices. Figure 2-2 shows that the experts in this sample did indeed have a lower omission effect than did the novices ($F(1,105) = 4.2$; $p < .05$). Note however, that the line for the expert group is still sloped indicating that there was still an omission effect for the experts. This slope was significantly different from zero using a simple effects follow-up test ($p <$ .05). This result was consistent with the second hypothesis.

Figure 2-3 illustrates the omission effect broken down by levels of availability. Note that the omission effect is nearly the same for both the high and low recall lists. The interaction was not significant ($F <$ 1). As well, the three-way interaction of availability, omission and expertise was not significant ($F < 1$) indicating that differential recall of the lists did not impact the omission effect for either group of subjects. It would appear that availability, in the sense of ease of recall, is not the mediating factor in the omission effect as had been speculated in the literature (e.g. Fischhoff et al., 1978) and had been predicted in the third hypothesis.

As reported above, experts in this study had a lower omission effect than did novices. Another aspect of expertise relates to the ability of experts to distinguish the high and low frequencies of events. As shown in Figure 2-4, experts were better able to distinguish the probabilities of the high and low availability possible causes. While the novices displayed little distinction between the high and low availability lists, the experts assessed the high (low) availability causes to also be high (low) in frequency (i.e. "all other causes" were assessed as high for the low recall causes and vice versa). This interaction was significant ($F(1,105) = 4.8$; $p < .05$). This result was consistent with the fourth hypothesis. Since, in general, availability and ecological frequency are positively correlated (Tversky and Kahneman, 1973), this result suggests that the partner/managers had better knowledge of the actual frequencies of the underlying events than did the novices. This result provides limited evidence that the expert subjects perceived the relative availability of the lists to be different. The novices did not make such an assessment and yet displayed an omission effect. Thus this result is not consistent with availability being the mediating factor in the omission effect. This result could mean that experts have better knowlege than do the novices of actual frequencies of events and therefore these events are more available to them. However, this difference could also be due to differential confidence in the assessments by the two groups.

Figure 2-5 illustrates that experts indicated more confidence in their probability judgments than did the novices ($t = 4.4$, $p < .001$; Bartlett's test for homogeneity of variances chi-square = .000, $p > .99$). However, as shown in Table 2-2 this confidence was not related to the availability (i.e. differential recall) of the items assessed ($F < 1$), contrary to Getty's et al.'s (1986) suggestion. This differential confidence was also not related to the length of the list assessed (i.e. the omission effect) ($F < 1$). Because the expertise by availability and expertise by omission interactions were not significant for the confidence assessments, this suggests that the expertise results for omission and availability in the probability assessments were not due to greater confidence on the part of the partner/manager group. These results were consistent with the fifth hypothesis.

Table 2-3 presents the means of the individual causes by the long and short list conditions. The probability values were taken from univariate ANOVAs for each of the two recall conditions. These results are inconsistent with an availability explanation for the omission effect, in that for both the high and low recall possible causes, subjects displayed a significant omission effect for some of the causes but not for others. That is, the availability explanation would predict omission effects for each of the high recall causes and lower omission effects (or none) for the low recall possible causes, since the availability explanation states that availability is the mediating factor in the omission effect.

DISCUSSION AND IMPLICATIONS

This study demonstrated that expert auditors were less susceptible than novice auditors to an omission effect in the assessment of probabilities of partial lists of possible causes. This expertise effect in fault tree judgments had not previously been demonstrated among other classes of experts. However, the audit partner and manager expert group still had a significant omission effect suggesting that they may have been unable to completely ascertain the causes which were missing from the partial list of possible causes.

It is also possible that the expertise effects found relate to some other systematic difference between the expert and novice groups such as a difference in experimental treatments. For example, for the experts, the experiments were administered at their own offices or at their own convenience. For the novices, the experiments were conducted in two large group settings at downtown hotel meeting rooms. There were also other obvious systematic differences between the novice and expert groups such as age. Thus there were potential confounding factors which could have caused the expertise effects found. However, these confounds in experimental design were necessary when studying expertise due to the limited time that experts, serving as volunteers, were willing to commit to the experimental task. Thus tradeoffs were made between internal and external validity. The most logical explanation for the difference between the two groups is expertise on the part of the partner/manager group. However, as with all research, these results should be replicated with a new task and a new sample of subjects.

The results also indicate that these auditing experts were better able to discriminate the relative frequencies of the high and low availability causes. This may be due to their knowledge of the underlying actual frequencies of such causes gained through experience. The experts also had greater confidence in their probability estimates.

The availability explanation (i.e. ease of recall) for the omission effect in fault tree judgments was examined. Contrary to conventional wisdom, availability was not the mediating factor in the omission effect.

Since availability and ecological frequency tend to be correlated (Tversky and Kahneman, 1973), the finding that relative availability was not a mediating factor in the omission effect may be confounded with relative frequency. However, availability did not mediate the omission effect for either the novices or the expert group (the EXPERTISE by AVAILABILITY by OMISSION interaction was not significant ($F < 1$)). Thus even for the experts, who should have had knowledge of the underlying frequencies of the possible causes, availability was not a mediating factor in the omission effect. This suggests that confounding of frequency and availability was not a problem in this study.

A potential weakness in this study is that there was no measure of the relative availability of each of the possible causes on a hold out sample of experts and novices. At the time the experiment was conducted, it was not possible to obtain sufficient additional subjects so that a separate test of availability could be made. As well, no independent assessment of prior availability was performed as a manipulation check because to do so using the experimental subjects would introduce a confounding factor into the experiments (i.e. subjects would have been focused). In prior studies (Fischhoff, Slovic and Lichtenstein, 1978, Rennie and Johnson, 1988) focusing was shown to reduce the omission effect. Thus not having an independent assessment of prior availability may make the finding of the expertise effect that much stronger. As a consequence, the results concerning availability in this study rely on the availability of the possible causes determined in Libby's (1985) study. Using Libby's results, however, has the benefit of providing an independent, objective classification of relative availability of the possible causes. That is, when interpreting subjects' responses in a free recall task, the experimenter must exercise some judgment when coding the responses. The availability results in this study replicate those found in Rennie and Johnson's (1988) study using chartered accountancy students. In that study a direct test was done of the relative availability of the possible causes used in that study.

Another potential weakness in the availability results is that although the relative availability of the individual possible causes was obtained from Libby's (1985) results, the possible causes were presented in groups of three (short list conditions) or six (long list conditions). Thus it is assumed that the availability of the groups of possible causes was not different from the sum of the availabilities of the individual possible causes. That is, it is assumed that no intralist cuing or interference took place.

An alternative explanation for the omission effect was suggested by Hoch (1984). This explanation comes from the part-list cuing phenomenon which suggests that providing a partial list of causes may interfere with subjects' ability to recall other possible causes. This output interference explanation suggests that viewing the incomplete list may strengthen the associations in memory between the listed causes and the causal event. This may decrease the probability that subjects will be able to retrieve possible causes which were omitted from the list. As well, Einhorn and Hogarth (1986) in their study on assessment of probable cause, present a model based on an anchoring-and-adjustment process. These explanations should be studied further.

The omission effects shown in this study are consistent with an anchoring-and-adjustment process for the assignment of probabilities. That is, subjects may have anchored on a probability equal to the reciprocal of the list length (e.g. 1/3 in the short lists and 1/6 in the long list) and then adjusted upwards or downwards from that anchor based on their beliefs of the underlying frequency. For example, for the BAD DEBTS and PURCHASES causes (see Table 2-3), subjects in the short list may have adjusted downwards from an anchor of 33 1/3 while subjects in the long list adjusted up from 16 2/3. There was therefore no omission effect. For the LONGTERM DEBT and RETURNS variables, the direction of adjustment was downwards from both anchors, but subjects in the short list did not adjust sufficiently, thereby producing an omission effect. When an anchoring-and-adjustment process is used, subjects typically do not adjust sufficiently from the anchor (Tversky and Kahneman, 1974). For the PURCHASE RETURNS and PAYMENTS variables the direction of adjustment is similarly downwards from both anchors. For these two possible causes there was a sufficient adjustment downwards from the anchor in the short list to reduce the significance of the omission effect. It is in these last two variables where experts made a greater adjustment than did the novices, which resulted in the significant expertise x omission interaction. Thus the different omission effect for the PURCHASE RETURNS and PAYMENTS causes may be due to the experts' greater knowledge of the underlying frequencies of these causes.

In an anchoring-and-adjustment explanation for the omission effect, as discussed above, it is assumed that subjects have sufficient knowledge of the underlying frequencies that they can make ordinal comparisons of the likelihoods of the possible causes. For example, for the BAD DEBTS possible cause, it is assumed that subjects in the long list believed from their backgrounds and experience that the underlying probability was greater than the anchor (i.e. the average probability of the listed causes) of 1/6 (the reciprocal of the list length). In the short list it is assumed that subjects knew that the underlying probability was less than the anchor of 1/3. Thus it is assumed that subjects have some knowledge of frequency information. With no knowledge of frequency information, an anchoring-and-adjustment strategy would predict that subjects would assess the probabilities of the individual causes as being equal to the anchor.

Figure 2-6 shows the mean probabilities of the individual causes for each of the list length (omission) conditions by level of expertise. As expected, the novices's responses are closer to the anchors of 16.67 (long list) and 33.33 (short list) for each possible cause. This is therefore further evidence that the expert group had a better knowledge of the underlying frequencies, since they were better able to discriminate the individual causes. This is also support for the anchoring-and-adjustment explanation for the omission effect because the novices, with less knowledge of the underlying frequencies, made probability estimates closer to the assumed anchor (i.e. the reciprocal of the list length).

The anchoring-and-adjustment explanation for the omission effect is consistent with the results of this study. This explanation requires that subjects have an ordinal knowledge of the underlying frequencies of the possible causes and use this information in the assignment of probabilities. Future research should try to assess to what extent auditors and other subjects actually have knowledge of frequencies of events. As well, future research should try to assess what anchor, if any, subjects actually use in making their probability assessments.

It could be suggested that the underlying cause of the omission effect is an experimental demand characteristic. That is, subjects in the short list condition might have assumed that the experimenter had provided them with a complete list of the most probable causes. While such an explanation cannot be ruled out, there are four factors that argue against it. First of all, the narrative told the subjects that the list of possible causes had been provided by the client's staff and subjects should have inferred that the list might be incomplete. Second, the expert group had a lower omission effect, indicating that if the effect is due to experimental demands, then it affected the subject groups differently. This nonconstant omission effect was also demonstrated by performing separate ANOVAs on the individual causes for each of the availability conditions. When this was done (see Table 2-3), the omission effects varied from highly significant to nonsignificant across the individual causes. Third, this omission effect has been demonstrated in a number of previous studies. Finally, even if the effect is a demand characteristic, there may be comparable professional judgment situations in audit practice. That is, auditors may often have to assess the probable cause of an event from an incomplete list of possible causes. Thus even if the omission effect results from experimental demands placed upon the subjects, there may be a counterpart in real judgment situations which is being captured in this phenomenon. If this is the case, then the omission effect cannot be said to be an experimental artifact.

Are there actual audit situations in which auditors have to make assessments of probable cause, and to which these results may apply? Responses to some of the post-experimental questions indicate that 80% of subjects use ratio analysis in analytical review and that 99% of subjects attempt to determine the possible cause for any fluctuations through discussions with the client. Similarly, when compliance errors are found in internal control evaluation, 97% of subjects attempt to determine the cause of such errors and 97% discuss these errors with the client's staff in order to determine their probable cause.

This study has demonstrated that although experts had a lower omission effect than did novices in this experimental task, auditors in general were not able to assess the probabilities of what was missing from a list of possible causes. If this effect exists in actual audit judgment situations, what can be done to minimize its impact on the judgment process? These results suggest that, since there was no difference in frequency assessments between the experts and novices in the full list condition, that use of complete checklists as a decision aid may be appropriate where inexperienced auditors must exercise judgment. However, such checklists must be complete and applicable to the judgment task, or a similar omission effect may result. As well, the impact of the omission effect may be reduced by ensuring that professional judgment decisions are congruent with the level of experience necessary to make such judgments. This last statement begs the question of how inexperienced auditors are to gain the necessary experience to exercise professional judgment. Further research into the underlying cognitive reason for the omission effect is necessary before one can speculate how the effect may be reduced.

Table 2-1

Analysis of Variance on Means of "Adjusted All Other Causes"

Effect		Mean-Square	F-Ratio	p
Expertise	(E)	1862.9	2.690	.104
Omission	(O)	16663.2	24.057	.000
Availability	(A)	5838.2	8.429	.005
Counterbalance	(C)	2.3	0.003	.954
E x O		2889.1	4.171	.044
E x A		3322.8	4.797	.031
E x C		1642.5	2.371	.127
O x A		177.4	0.256	.614
O x C		652.4	0.942	.334
A x C		28.3	0.041	.840
E x O x A		0.2	0.000	.987
E x O x C		920.0	1.328	.252
E x A x C		1916.0	2.766	.099
O x A x C		486.7	0.703	.404
E x O x A x C		232.4	0.336	.564
Error (105 df)		692.7		

Table 2-2

Analysis of Variance on Means of Confidence Judgments

Effect	Mean-Square	F-Ratio	p
Expertise (E)	77.7	18.209	.000
Omission (O)	0.2	0.044	.835
Availability (A)	0.5	0.113	.737
E x O	2.6	0.620	.433
E x A	9.2	2.166	.144
O x A	0.6	0.147	.702
E x O x A	5.6	1.315	.254
Error (109 df *)	4.3		

* N.B. 4 cases were omitted due to missing data

Table 2-3

Means of Individual Causes

Possible Cause	Long List (Six Causes)	Short List (Three Causes)	*p* *
High Recall	n = 60		
BAD DEBTS	20.5	19.8	.862
PURCHASES	21.5	28.9	.167
LONGTERM DEBT	11.2	23.9	.002
Low Recall	n = 61		
RETURNS	13.5	26.6	.011
PURCHASE RETURNS	11.0	18.5	.061
PAYMENTS	10.4	16.1	.189

* from separate univariate ANOVAs by availability condition

Figure 2-1
Omission Main Effect

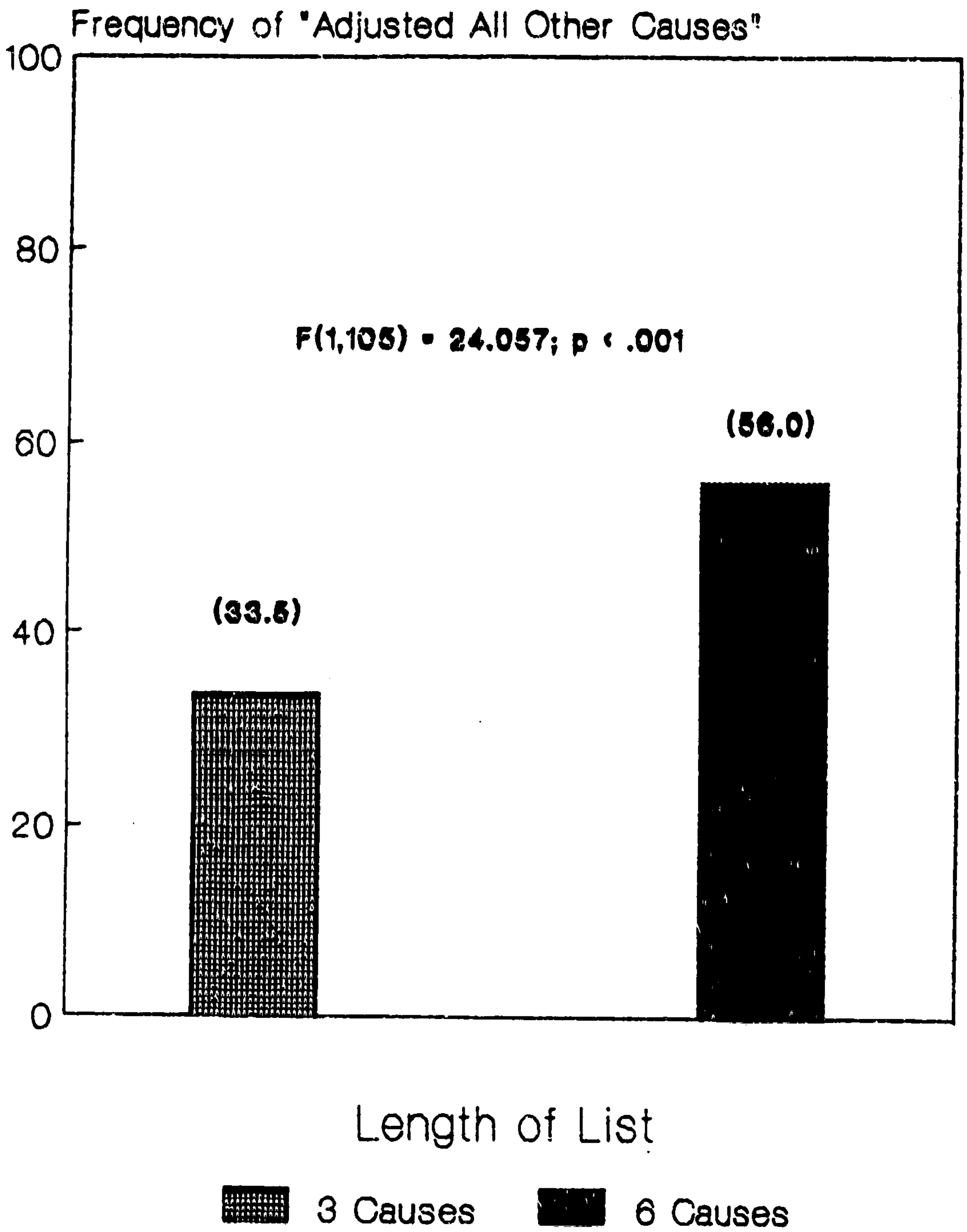


Figure 2-2
Expertise x Omission
Interaction

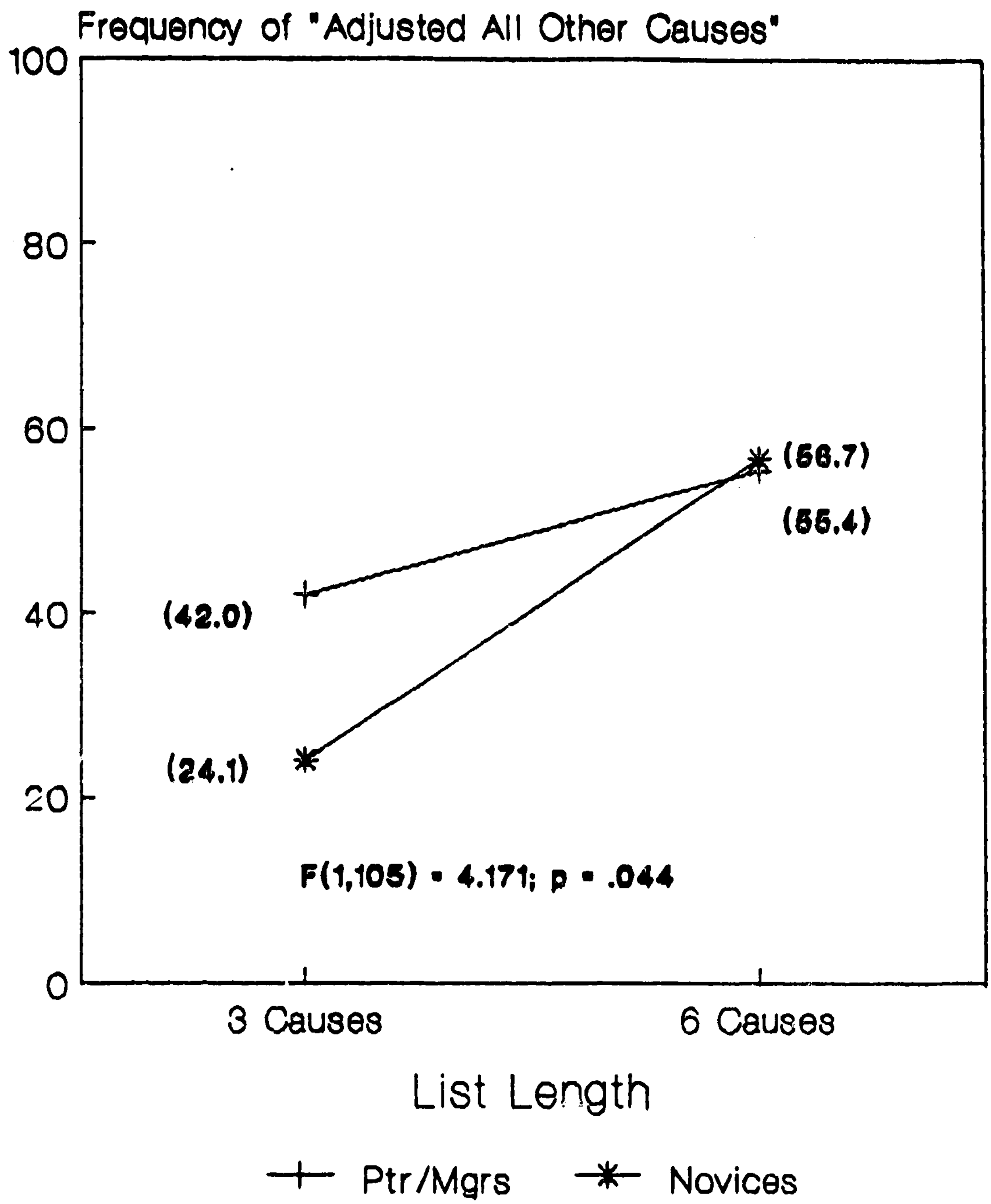


Figure 2-3
Availability x Omission Interaction

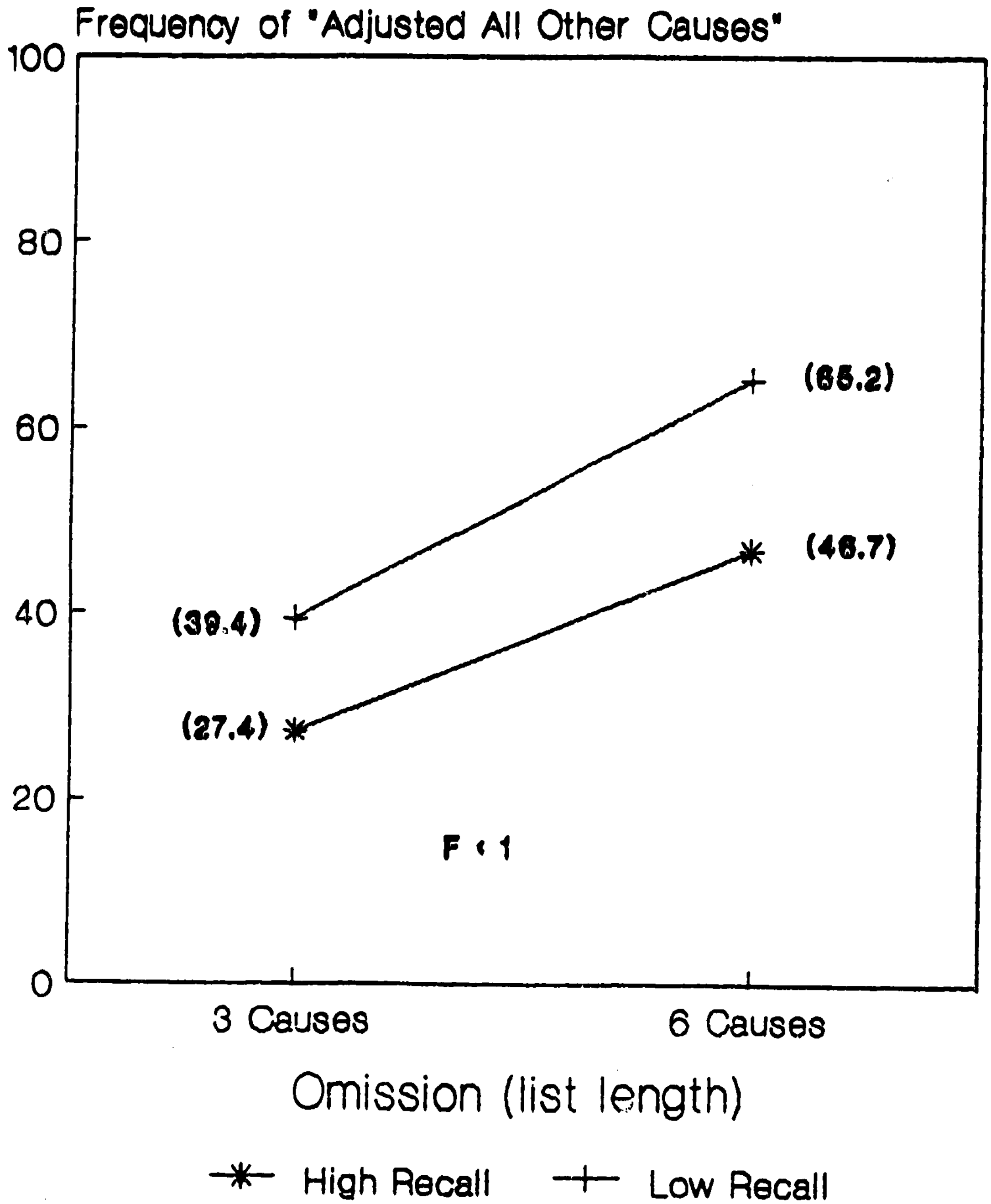


Figure 2-4
Expertise x Availability
Interaction

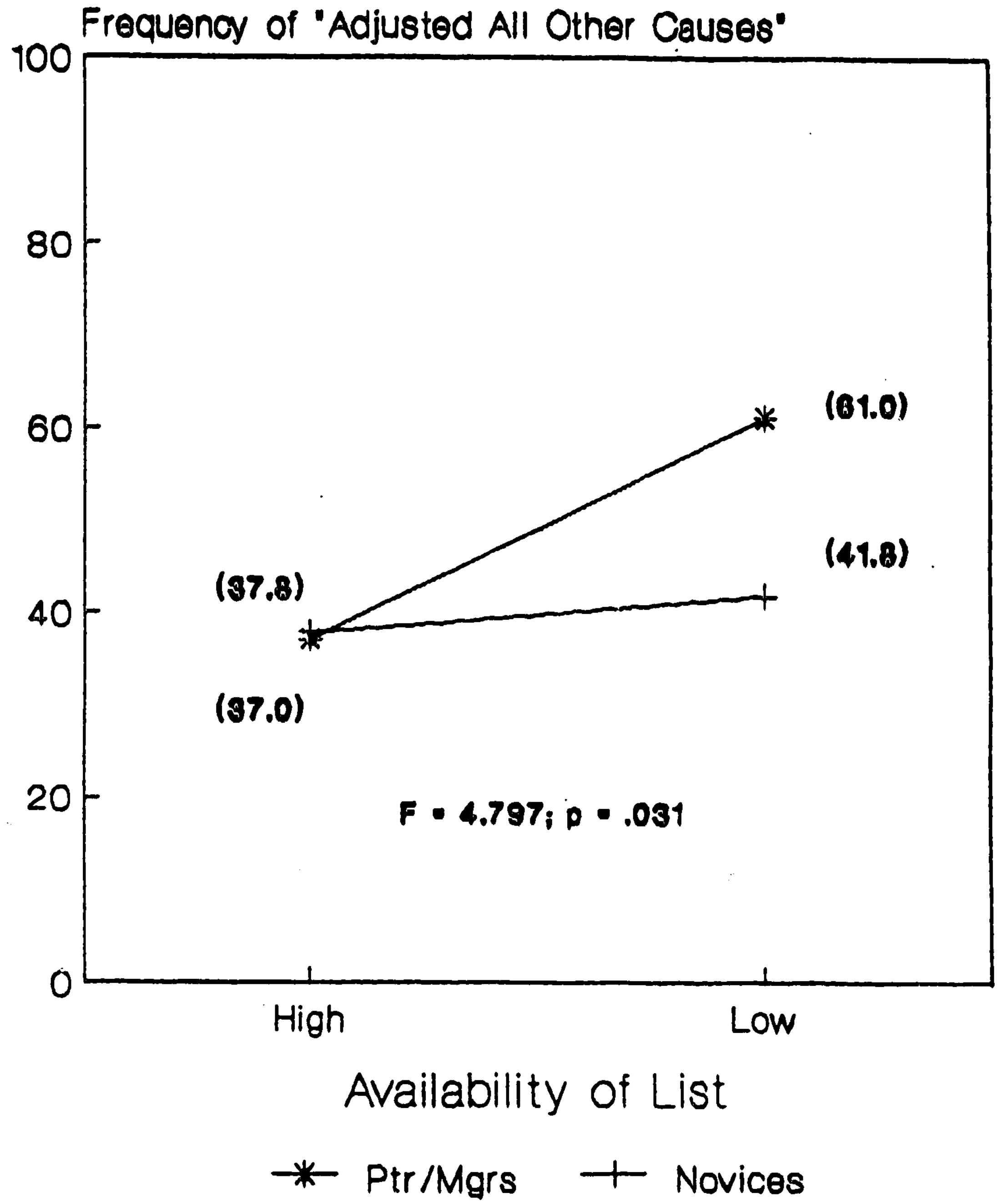


Figure 2-5
Confidence Judgments
by Expertise

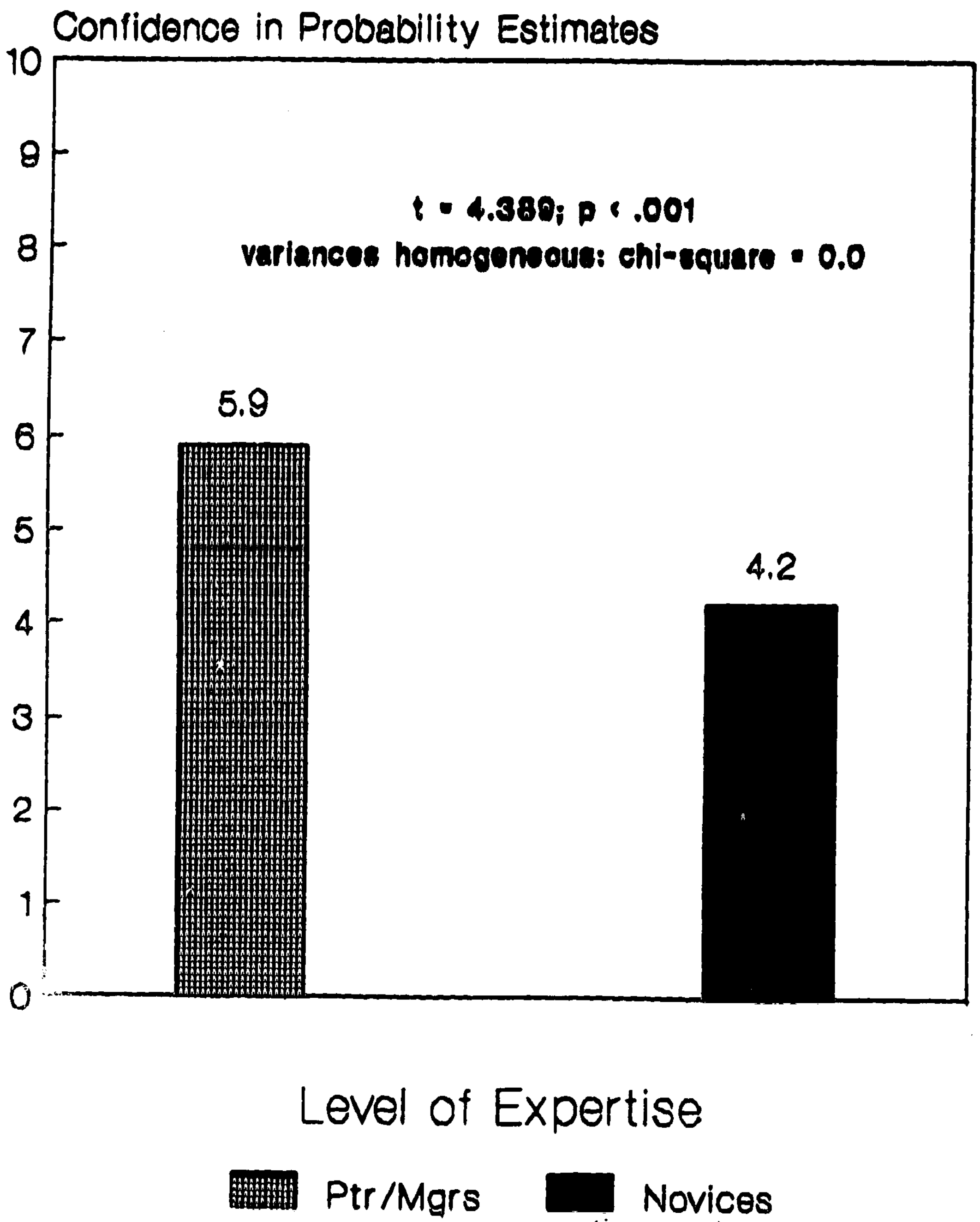
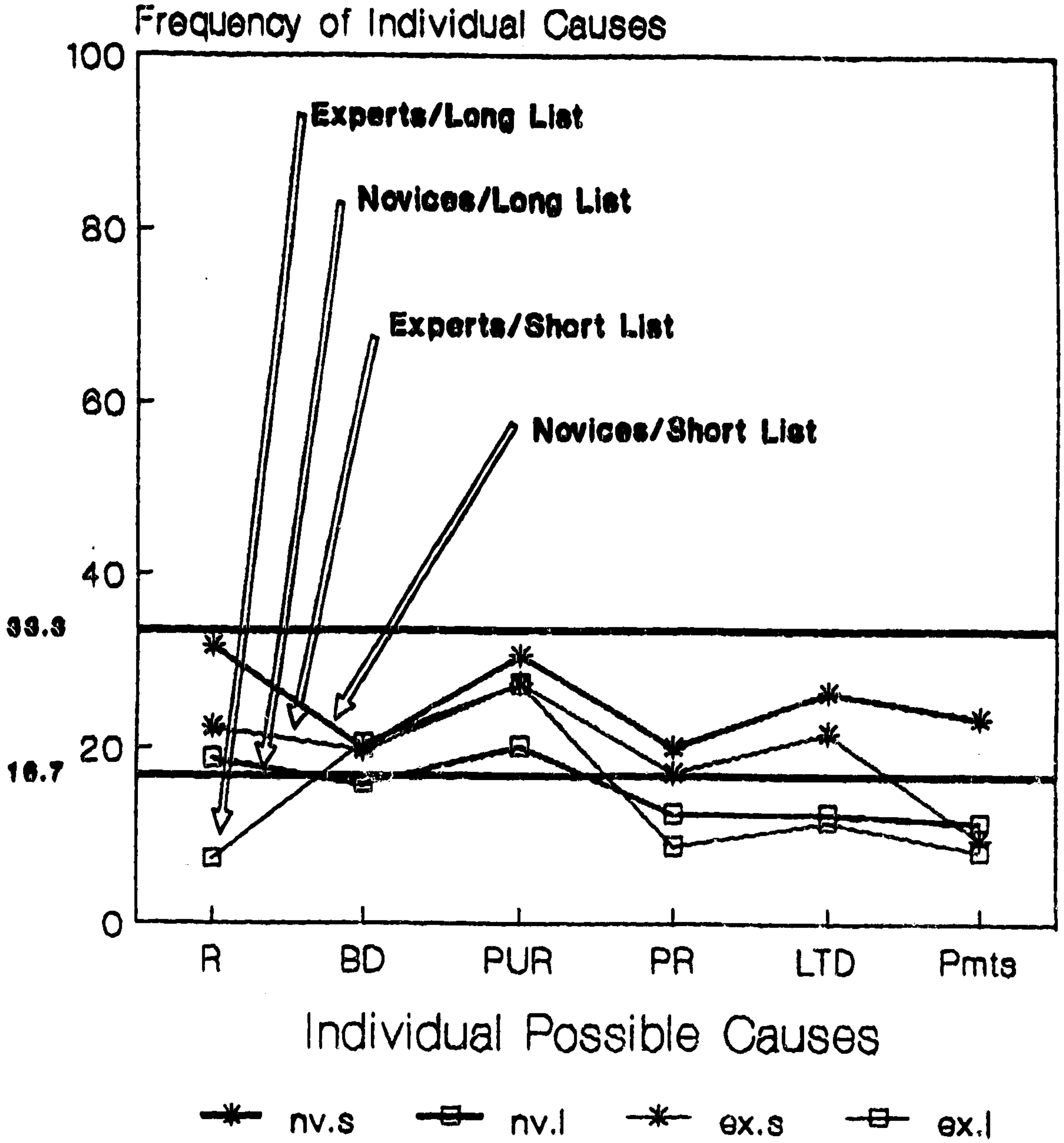


Figure 2-6
Probability Estimates of Individual Causes by Expertise



nv.s = novices/short list
ex.l = experts/long list

REFERENCES

Anderson, R. J. (1977). The external audit I: Concepts and techniques. Toronto: Pitman.

Burgstahler, D. & Jiambalvo, J. (1986). Sample error characteristics and projection of error to audit populations. The Accounting Review, 61, 233-248.

Conover, W. J. (1980). Practical nonparametric statistics, 2nd ed. New York: John Wiley.

Dube-Rioux, L. & Russo, J. E. (1988). An availability bias in professional judgment. Journal of Behavioral Decision Making, 1, 223-237.

Duh, R. R. & Sunder, S. (1985, September). Information in a market environment: An examination of the base-rate fallacy. Unpublished manuscript, University of Minnesota.

Einhorn, H. J. & Hogarth, R. M. (1986). Judging probable cause. Psychological Bulletin, 99, 3-19.

Fellingham, J. C. & Newman, D. P. (1985). Strategic considerations in auditing. The Accounting Review, 60, 634-650.

Fischhoff, B., Slovic, P. & Lichtenstein, S. (1978). Fault trees: Sensitivity of estimated failure probabilities to problem representation. Journal of Experimental Psychology: Human Perception and Performance, 4, 330-344.

Gettys, C. F., Mehle, T. & Fisher, S. (1986). Plausibility assessments in hypothesis generation. Organizational Behavior and Human Decision Processes, 37, 14-33.

Gibbins, M. (1984). Propositions about the psychology of professional judgment in public accounting. Journal of Accounting Research, 22, 103-125.

Hirt, E. R. & Castellan, N. J., Jr. (1988). Probability and category redefinition in the fault tree paradigm. Journal of Experimental Psychology: Human Perception and Performance, 14, 122-131.

Hoch, S. J. (1984). Availability and interference in predictive judgment. Journal of Experimental Psychology: Learning, Memory, and Cognition, 10, 649-662.

Kida, T. (1984). The effect of causality and specificity on data use. Journal of Accounting Research, 22, 145-152.

Kinney, W. R., Jr. (1975a). A decision theory approach to the sampling problem in auditing. Journal of Accounting Research, 13, 117-132.

Kinney, W. R., Jr. (1975b). Decision theory aspects of internal control system design / compliance and substantive tests. Journal of Accounting Research, 13, supplement, 14-29.

Libby, R. (1981). Accounting and human information processing: Theory and applications. Englewood Cliffs NJ: Prentice-Hall.

Libby, R. (1985). Availability and the generation of hypotheses in analytical review. Journal of Accounting Research, 23, 648-667.

Mehle, T. (1982). Hypothesis generation in an automobile malfunction interference task. Acta Psychologica, 52, 87-106.

Mehle, T., Gettys, C. F., Manning, C., Baca, S. & Fisher, S. (1981). The availability explanation of excessive plausibility assessments. Acta Psychologica, 49, 127-140.

Moser, D. V. (1989). The effects of output interference, availability, and accounting information on investors' predictive judgments. The Accounting Review, 64, 433-448.

Rennie, R. D. & Johnson, R.D. (1988, October). Auditors' judgments of probable causes: Effects of availability, experience, focusing and omission. Presentation at ORSA/TIMS, Denver.

Scott, W. R. (1984). The state of the art of academic research in auditing. Journal of Accounting Literature, 3, 153-200.

Slamecka, N. J. (1968). An examination of trace storage in free recall. Journal of Experimental Psychology, 76, 504-513.

Smith, D. G. (1983). Analytical review: A research study. Toronto, Canada: The Canadian Institute of Chartered Accountants.

Tulving, E. & Pearlstone, Z. (1966). Availability versus accessibility of information in memory for words. Journal of Verbal Learning and Verbal Behavior, 5, 381-391.

Tversky, A. & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. Cognitive Psychology, 4, 207-232.

Tversky, A. & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. Science, 185, 1124-1131.

Waller, W. S. & Felix, W. L., Jr. (1984). The auditor and learning from experience: Some conjectures. Accounting, Organizations and Society, 9, 383-406.

Waller, W. S. & Felix, W. L., Jr. (1989). Auditors' causal judgments: Effects of forward vs. backward inference on information processing. Accounting, Organizations and Society, 14, 179-200.

Weber, M., Eisenfuhr, F. & Von Winterfeldt, D. (1988). The effects of splitting attributes on weights in multiattribute utility measurement. Management Science, 34, 431-445.

III. OMISSION EFFECTS IN FAULT TREES: AN ANCHORING-AND-ADJUSTMENT PHENOMENON

Research has shown that when subjects are asked to assess the probabilities of possible causes for an event (such as in a "fault tree" when the causes are organized hierarchically) they tend to underestimate the probability of "all other causes." This failure to determine what causes are missing from a list is referred to in this paper as the "omission effect." This omission effect has been demonstrated even for expert mechanics (Fischhoff, Slovic and Lichtenstein, 1978, Mehle, 1982) and professional hospitality industry managers (Dube-Rioux and Russo, 1988). As noted by Slovic, Fischhoff and Lichtenstein (1982), the use of listings of possible causes together with their estimated probabilities are relevant to environmental protection, management of nuclear hazards, use of automobile seatbelts, etc. Thus the potential effects of underestimation of the probabilities of all other possible causes not provided to a decision maker has broad implications for risk and probability assessment.

None of the long lists nor the short lists in this paper or in previous studies can possibly contain complete listings of all possible causes. Therefore this paper refrains from using the established (and possibly misleading) terminology of "full tree" and "pruned tree." As well, the listings of possible causes are not called fault trees in the current study because they contain only items from a single level of the tree, and not hierarchical listings of nested causes. Because Fischhoff, Slovic and Lichtenstein (1978) found similar omission effects in both single and hierarchical levels of fault trees, this study does not address the issue of whether the omission effect is different for different levels within a fault tree.

The purpose of this research is to try to discover the cognitive process that underlies the omission effect. Unless and until a satisfactory explanation for this omission phenomenon is established, measures to improve fault tree judgments will be ad hoc. This paper contributes to the literature by examining an anchoring-and-adjustment explanation for the omission effect. As well, the availability explanation (Fischhoff, Slovic and Lichtenstein, 1978) and the output interference (Hoch, 1984) explanations for the omission effect are examined.

An anchoring-and-adjustment explanation states that when subjects are asked to assess the probabilities of a list of possible causes, they may anchor on a probability estimate which is the reciprocal of the number of listed causes. They then increase or decrease their estimate from the anchor based on their perception of the underlying frequency. When large adjustments are needed (e.g. for low frequency causes), the adjustment may not be sufficient, thereby producing an omission effect. As is common where an anchoring-and-adjustment heuristic is employed, there is an insufficient adjustment from the anchor (Tversky and Kahneman, 1974) thereby producing an omission effect.

The Availability Explanation

Fischhoff, Slovic and Lichtenstein (1978) speculated that the omission effect was due to the lack of availability of the omitted causes, that is "out of sight, out of mind." They did not, however, directly manipulate availability in their study. The availability explanation suggests that there should be a lower omission effect when subjects in the short list conditions assess low recall causes than when subjects assess high recall causes. This is because high recall causes should already be available to subjects. Thus cuing subjects to the low recall causes by specifically listing them should mean that they will be better able to fully populate the sample space of possible causes than will subjects who assess only high recall causes and who are therefore not cued. Weber, Eisenfuhr and Von Winterfeldt (1988), in a study involving value trees, similarly state that the most straightforward explanation (for what is called the omission effect in the current study), is availability.

A descriptive theory of how individuals make assessments of probable cause comes from norm theory (Kahneman and Miller, 1986). Normality of a possible cause is assessed by comparing it with the norms it evokes after the causal event. A normal cause is one that has few alternatives; an abnormal cause is one which has highly available alternatives. A normality measure is obtained by assigning a normality of 1.0 to the most available causal event.

In comparing normality and statistical probability, the normality of a possible cause can therefore increase without reducing the normality of another possible cause (because the value of 1.0 is obtained merely by scaling). This is not the case with statistical probability because the total sample space must sum to one. That is, there is a total probability of 1.0 so that if the probability of one cause were to increase then the probability of at least one other cause must decrease. If subjects confuse statistical probability and normality, then this use of availability may account, at least in part, for the omission effect. The number of available alternatives (or the list length) may determine the normality of a particular cause. Thus relative to a short list (few alternatives) a particular cause may be assessed as high in normality while the same possible cause relative to a long list (more alternatives) may be assessed as lower in normality. These differences in normality assessments between the long and short lists would result in an omission effect.

Norm theory could, therefore, make similar predictions to those from an anchoring-and-adjustment explanation for the omission effect, since both explanations depend upon list length. However, an anchoring-and-adjustment mechanism merely starts the probability estimate at the anchor (based on list length) and then adjusts insufficiently for the perception of the underlying frequency. Under a norm theory explanation, the adjustment from list length would be solely due to the number of other causes that were available. Thus simultaneously manipulating the list length (by informing subjects of the length in the introduction to the task) and the number of causes that are available (by sequentially presenting the causes to be assessed) provides a means of contrasting the predictions of norm theory with those of an anchoring-and-adjustment explanation for the omission effect.

Rennie and Johnson (1988) examined the role of availability in auditors' (novices and seniors) assessments of probable cause. Availability was defined as ease of recall (Tversky and Kahneman, 1973, 1974). In Rennie and Johnson's study subjects were asked to assess the probabilities of either six or three possible causes for an auditing discrepency. Two short lists (three possible causes) were constructed by selecting either the three most or the three least easily recalled items from the long list (six possible causes). Ease of recall was determined from a pilot study. The overall results indicated a strong omission effect ($F(1,125) = 44.49$, $p < .001$). That is, subjects overestimated the probabilities associated with the reduced list of three possible causes relative to the full list. However, there was no difference (between-subjects) in the omission effects between the high and low recall partial lists ($F < 1$). Rennie and Johnson's results depended upon there being a valid manipulation of availability between the high and low recall lists of possible causes. To provide evidence of such a distinction, subjects were asked to list possible causes either prior to, or after, the probability assessment task. For the high recall causes, 77% of the subjects in the prior-elicitation group and 73% of subjects in the post-elicitation group listed at least one of the high recall possible causes (chi-square not significant). For the low recall causes, however, only 9.4 % of prior-elicitation subjects listed at least one of the low recall causes, compared with 39.6% of subjects in the post-elicitation group (chi-square significant, $p < .001$). This difference between the high and low recall lists suggests that the high recall possible causes were more available to subjects in the probability assessment task than were the low recall causes, which were only made available through being cued in the probability assessment task. Yet, even with this cuing, there was no difference in the omission effect between the high and low recall lists.

Fischhoff, Slovic and Lichtenstein (1978) speculated that the omission effect in partial lists of possible causes was due to availability (citing Tversky and Kahneman, 1973 for the meaning of availability). The results of Rennie and Johnson (1988) cast some doubt on Fischhoff et al.'s interpretation, if availability is defined as recall. However the concept of availability is somewhat ambiguous. When Tversky and Kahneman (1973) introduced the availability heuristic they referred to either the recall from memory of stored knowledge of an event, or the ease of construction of scenarios, where the subject had no direct prior experience with a particular event. They specifically stated that their concept of availability was different from availability in a verbal learning context (cf. Tulving and Pearlstone, 1966). The verbal learning conception of availability is closely associated with aided recall (as contrasted with unaided recall or accessibility). Kahneman and Tversky (1982) later called the process of construction of scenarios the simulation heuristic. It would appear, then, that the meaning of availability in Fischhoff et al. (1978) is uncertain, since Slovic, Fischhoff and Lichtenstein (1982), in referring to the results of Fischhoff et al. (1978), discuss both the recall and simulation concepts of availability.

As a result of the diverse interpretation of availability, it is not clear that many of the prior studies that have attempted to test availability as a mediating factor in the omission effect have in fact done so.

Hirt and Castellan (1988) attempted to examine whether availability was a mediating factor in the omission effect. Their experiments were based on Fischhoff et al.'s (1978) "reasons for a car not starting" experiment and used the same list of causes for the long list. For the short list they omitted one cause. Hirt and Castellan used students in a within-subjects design to examine the omission effect, category redefinition, and the availability explanation for the omission effect. In their first experiment, they had one-half of the subjects generate possible causes for a car not starting prior to the probability assessment tasks. They suggested that this generation task should increase the availability of the omitted cause. This manipulation would appear to be more related to focusing (i.e. thinking about or attending to possible causes) in the sense of Fischhoff et al. (1978) rather than availability. The result of this manipulation was that there was no significant difference in the omission effect between the generation and no generation groups.

The within-subjects design of Hirt and Castellan's study allowed them to examine a second test of availability. The subjects assessed the probabilities of both the long and short list of possible causes. This assessment order was counterbalanced so that an order effect could be assessed. Hirt and Castellan reasoned that if availability were the mediating factor in the omission effect, then subjects assessing the long list first should have a lower omission effect than subjects assessing the short list first. This conception of availability, or the immediate recall of a cued item, is more consistent with Tulving and Pearlstone's (1966) conception in a verbal learning context (which was specifically differentiated from "availability" by Tversky and Kahneman, 1973). The results of this manipulation were that subjects who assessed the long list first had a significantly lower omission effect than did subjects assessing the short list first. Although their result was consistent with this limited definition of availability, Hirt and Castellan did not accept the availability hypothesis because the results of their availability manipulations were mixed.

Hirt and Castellan used a within-subjects design so that they could examine category redefinition between the long (five category) and short (four category) lists of possible causes. They found evidence to support idiosyncratic redefinition of categories by subjects between the short and long lists. However, the within-subjects design may have created an internal validity problem. That is, all subjects assessed both the long and short lists. In order to make the assessment of the second list of possible causes appear plausible to subjects, subjects were told that (p. 124) "The chart you just saw was only one automobile manufacturer's guide in thinking about the problem if starting failure.... Other manufacturers dealing with the same problems may decide to construct their charts differently...." Thus subjects were told that the population of automobiles in the assessment task was different for the long and short lists. It is not clear, therefore, what an omission effect would mean in such a situation, since the long and short lists may not be comparable.

Dube-Rioux and Russo (1988) also claimed to examine the availability explanation for the omission effect. For their meaning of availability they referred to retrieval failure. Retrieval failure is a verbal learning concept that contrasts availability with accessibility. In their study they had professional managers assess the probabilities of reasons why a restaurant may fail. Subjects assessed lists of possible causes which were categorized as either long-term or short-term causes. Since they expected short-term causes to be more easily retrieved, they attempted to test the role of availability in the omission effect. In both the short lists (long-term and short-term), however, they included combinations of the short-term and long-term causes (p. 227). This has the effect of confounding high and low recall so that the study cannot adequately comment on the availability explanation. Even though the short-term causes as a group may have been more easily generated by subjects than were the long-term causes as a group, it is unknown what the relative recall of the two combinations of long-term and short-term causes were. That is, consider the following: (a) the sum of a list of two high numbers (40 and 50) and one low number (10) (sum equals 100); (b) the sum of a list of two low numbers (5 and 10) and one high number (90) (sum equals 105). Thus even though a list may contain more low items, the presence of a sufficiently high item may result in its total value exceeding the value of a list containing a greater frequency of high items. As well, the results of the availability manipulations in Dube-Rioux and Russo (1988, p. 231) were mixed. In fact, the major test of availability, that predicted there would be a smaller estimation bias for the short-term cause condition, was not significant.

It would appear that neither the Hirt and Castellan (1988) study nor the Dube-Rioux and Russo (1988) study examined availability either as recall of past events or as ease of construction of scenarios (Tversky and Kahneman, 1973).

Rennie and Johnson's (1988) results are inconsistent with the availability explanation for the omission effect, where availability is defined as ease of recall. However, since availability (Kahneman and Tversky, 1973) may also be defined as ease of construction (or simulation) of alternative scenarios, these results do not rule out an alternative conceptualization of availability as the mediating factor in the omission effect.

The Output Interference Explanation

A second possible explanation (Hoch, 1984) for the omission effect in lists of possible causes is output interference via the "part-list cuing" phenomenon (Slamecka, 1968). In the part-list cuing procedure, subjects are asked to study a list for later recall. A "cued" group of subjects is provided a subset of the list and is asked to recall the remaining items while a control group is asked to recall the entire list. The result is that the control group consistently recalls more of the remaining items than does the cued group. It appears that the experimenter-provided items interfere with a subject's ability to recall (or output) other items.

Rundus' (1973) model which deals with probabilistic recall from associative memory can account for output interference in fault trees. Providing subjects with a partial list of possible causes may increase the associative strength between these listed causes and the causal event (and hence the probability that they will be recalled). Since the memory search process is assumed to be with replacement, these listed causes will have a higher probability of subsequent retrieval; thus providing subjects with the short list of causes will interfere with their retrieval of other possible causes.

In Hoch's (1984) study three experiments were conducted to investigate the joint influence of output interference and the availability heuristic on predictive judgments in hypothesis generation tasks. Hoch assumed that likelihood judgments of future events would be based upon the operation of an availability heuristic. The omission effect could result from output interference within the retrieval process upon which availability-based frequency judgment was made. Unlike the studies that tested the availability explanation for the omission effect, Hoch's study involved generating reasons (and their likelihood) for a future event (e.g. a consumer purchase) that had not yet taken place. Thus the conception of availability in Hoch's study was that of simulation or construction of scenarios, rather than recall of past events. In the first experiment subjects generated pro and con reasons for the event (the order was counterbalanced). The order of generation significantly affected the number of reasons generated, which supported an output interference explanation within a process based on an availability heuristic. In the second experiment the time taken to generate the second set of reasons was recorded. The third experiment replicated the second experiment except that a distraction task (delay) condition was introduced between the generation of the pro or con reasons. In the third experiment in the no-delay condition, the probability judgments were higher in the pro-con order than in the con-pro order. However, when the intervening distraction task was introduced, this difference reversed and the probabilities were lower in the pro-con order than in the con-pro order. Hoch concluded (p. 658) that this interaction raises questions about the relation between hypothesis-generation and an availability-based inference rule. Hoch stated that his results could be due to output interference. The primacy and recency effects could be due to attention decrement. Hoch suggested that this attention decrement is also consistent with an anchoring-and-adjustment process with insufficient adjustments to the anchor: in the no-delay condition subjects could anchor their probabilitity estimates on the first set of reasons and adjust insufficiently for the second set; in the delay condition subjects could anchor on the second set. Hoch further suggested that interference is more likely to result in events requiring construction of scenarios than those requiring recall of past events.

Some of the results of the studies that examined the availability explanation for the omission effect can be interpreted in terms of output interference. In Fischhoff et al.'s (1978) study, the omission effect was reduced but not eliminated by having subjects focus on all other causes (subjects were told to specifically consider what was missing). Focusing (or thinking about all possible causes) immediately prior to assessing a partial list of possible causes may reduce output interference because the associations in memory among all list items have been primed just prior to the interference task. Thus there is a small relative difference in level of activation between the listed and unlisted causes in a fault tree when a subject has been focused. Without focusing, only the associations among the presented list items are activated leaving a large relative difference in level of activation between the presented and omitted causes. With focusing, therefore, there is a higher probability that the primed causes (whether or not presented in the list) will be recalled (and less output interference) than would be the case without the focusing.

In one of the conditions of Rennie and Johnson's (1988) study, subjects generated possible causes prior to the probability generation task. This focusing significantly reduced the omission effect ($F(1,125 = 4.79$, $p < .05$). Furthermore, the focusing effect was greater for the low availability causes than it was for the high availability causes ($F(1,125 = 4.42$, $p < .05$). In fact, contrary to the availability explanation, the omission effect was less for focused subjects when they assessed the high availability list than when they assessed the low availability list. These positive results for focusing suggest that having subjects try to populate the sample space prior to the probability assessment task may reduce the output interference.

A similar result was found in Dube-Rioux and Russo (1988) where extending the fault tree led to lower omission effects. On the other hand, using a similar focusing procedure, Hirt and Castellan (1988) found no effect for having subjects generate possible causes prior to the probability assessment task. Hirt and Castellan's failure to find a positive effect for focusing may have been due to recategorization by subjects. That is, since subjects were told the long and short lists were from different automobile manufacturers, the subjects may have defined the overlapping categories differently in each of the long and short list assessment tasks.

Hoch's (1984) results, together with consistent findings in other studies, support an output interference explanation for the omission effect. Hoch (1984) examined the joint operation of availability and output interference. His conception of availability was that of ease of construction of scenarios, rather than recall of past events. His results were consistent with the joint operation of output interference and an availability-based inference rule (based on ease of construction of scenarios). Because Hoch did not study availability as recall of past events, it is unknown whether output interference or some other process may mediate the omission effect in such circumstances. On the other hand, Hoch also stated that his results might be accounted for by anchoring-and-adjustment.

The Anchoring-and-adjustment Explanation

Einhorn and Hogarth (1986) presented a descriptive framework by which people may judge probable cause. The first three components of the framework are: (1) the context (causal field); (2) probabilistic indicators of causal relations (cues-to-causality) --covariation, temporal order, contiguity in time, similarity; and (3) strategies for combining the causal field and cues-to-causality. This framework determines the gross strength of each single possible cause. The fourth component of the framework deals with situations where there are alternate possible causes for an event. In such cases, the net strength of a causal explanation is determined by subtracting the weighted strength of the alternatives. This is a Bayesian-like updating model which involves an anchoring-and-adjustment process. The model can therefore explain why even a relatively weak explanation is given considerable support if there are no alternative explanations.

There have been no studies that have examined an anchoring-and-adjustment explanation for the omission effect. However, if subjects form an anchor based on the reciprocal of the list length and then adjust this frequency to reflect their beliefs in the likelihood of the particular cause, then insufficient adjustments from the anchor would be sufficient to produce an omission effect.

One of the characteristics of the anchoring-and-adjustment heuristic is that subjects do not adjust sufficiently from the anchor (Tversky and Kahneman, 1974, Slovic, Fischhoff and Lichtenstein, 1982). For example, Tversky and Kahneman (1974) discussed an experiment where subjects were asked to indicate the percentage of African countries belonging to the United Nations by moving a wheel of fortune to the correct position. The initial position of the wheel was determined by spinning the wheel in the subjects' presence. The initial random position of the wheel affected subjects' responses, with subjects who had a higher initial position of the wheel making higher estimates of the percentage of African countries in the United Nations. Slovic, Fischhoff and Lichtenstein (1982) discussed an experiment where subjects were asked to estimate the frequencies of death from 40 different causes. One group of subjects was told about the 50,000 annual fatalities from motor vehicle accidents while another group of subjects was told about the 1,000 annual deaths by electrocution. Subjects who were told about the deaths by electrocution made lower estimates of the frequencies for most of the 40 causes of death than did subjects who were anchored on the large number of motor vehicle deaths. Thus, with incomplete knowledge of frequency information for all items in a list, subjects may anchor on a number (even a random one) and adjust insufficiently from that anchor when making a probability or frequency estimate.

Overview of Experiments

This study reports the results of three experiments. The first experiment examined the anchoring-and-adjustment explanation for the omission effect. As well, the experiment examined the availability explanation for the omission effect (using a simulation or construction concept of availabilty). Evidence from Rennie and Johnson (1988) was discussed above suggesting that the recall conception of availability is not a mediating factor in the omission effect. The second and third experiments dealt with focusing effects in order to examine the output interference explanation for the omission effect. For summaries of the variables, the experimental designs and cell sizes, and the statistical method, please refer to the appendix.

Experiment 1

The goal of this research was to attempt to determine the cognitive processes underlying auditors' judgments of probable cause which could give rise to the omission effect. This experiment tried to distinguish between one of the propositions of Einhorn and Hogarth's (1986) probable cause framework and a corresponding proposition from Kahneman and Miller's (1986) norm theory. Einhorn and Hogarth's (1986) probable cause framework stated that competing possible causes are judged as to their relative probability on the basis of anchoring-and-adjustment. This was assessed by telling subjects how many causes they are going to assess and then having subjects assess these causes sequentially. Norm theory proposes that normality of a possible cause is assessed by comparing it with the number of available alternatives. Norm theory's proposal was assessed by using low recall possible causes and counterbalancing the order of the sequentially presented alternative possible causes in order to assess the effect of making another cause available to subjects. Because the recall notion of availability was examined in the Rennie and Johnson (1988), the first experiment used the construction concept of availability, by attempting to hold degree of recallability constant at a low level. That is, list items which were low in recall were chosen so that when they were made available by being cued in the task, they would be unlikely to result in recall of specific past instances involving these items (i.e. by design they were low recall items). Similarly, the verbal learning conception of availability as cued recall (vs. accessibility) was avoided since the list items were low recallability and they had not been previously presented.

The design of the first experiment was a 2^3 factorial of expertise (novices vs. audit partner/managers), anchor (number of possible causes [three or six] that subjects were told they would be asked to sequentially assess) and order (the order of the first two causes, both low in recall, was counterbalanced so that their availability could be controlled).

It was predicted under an anchoring-and-adjustment explanation for the omission effect, that informing subjects of the different list lengths would affect their responses. That is, subjects would anchor on a probability of the inverse of the list length and then adjust this probability subjectively to reflect the perceived likelihood of occurrence. Thus a significant main effect for the ANCHOR variable was predicted from the anchoring-and-adjustment explanation.

It was predicted, from norm theory, using an availability explanation for the omission effect, that normality would be judged inversely to the number of available alternatives. Thus a significant ORDER main effect was predicted by availability, because these low recall causes would be more available as alternative possibilities when the second possible cause was assessed.

Method

Subjects and Administration

The 130 experimental subjects were grouped in terms of audit experience. The expert subjects were 74 audit partners and managers from Chartered Accountancy firms in Regina and Saskatoon, Saskatchewan. The novice subjects were 56 students registered with the Institute of Chartered Accountants of Saskatchewan and who had commenced employment with Chartered Accountancy firms within the past six months.

All subjects were volunteers. The sessions for the novices were held as the last agenda item at two new students' days in Regina and Saskatoon. One student at each of the two sessions chose not to participate in the study.

The expert subjects were solicited by contacting the managing partners (or equivalent title) of 14 offices of chartered accountants. The 14 offices consisted of 12 international firms, 1 regional firm and 1 local firm. The managing partners of all offices contacted agreed to participate in the study and solicited volunteers. Where possible, the experiment was conducted in a group setting by office. In the group administrations of the experiment, there were no clarifying questions of substance asked. For offices which were not able to arrange a common time for the volunteers to meet, the experiment was self-administered by the subjects. These self-administered subjects were asked to spend the same time at the task as had the group subjects and were given written instructions based on the instruction script read to group subjects. The self-administered subjects were debriefed by using a sealed envelope to be opened after completing the experiment. Because there were approximately equal numbers of group-administered (n=33) and self-administered (n=41) experts in each experimental condition, a separate ANOVA was performed on responses from the expert group with each expert subject coded as to type of administration. Neither the main effect for administration condition nor any interactions of administration condition and the other treatment effects were significant.

The expert auditor group consisted of audit partners and managers. Promotion to audit manager varies by firm with some firms having an intermediate rank of supervisor or an advanced rank of senior manager. When subjects were solicited, care was taken to ensure that a common definition of manager was used which included junior managers or supervisors as well as senior managers. As well, only those subjects who spent substantially all of their time in auditing were requested to volunteer.

The novices had a mean age of 23.8 years and mean auditing experience of 0.5 years. The expert group had a mean age of 33.4 years and mean audit experience of 11.9 years.

Responses for 9 subjects were deleted due to minor addition errors in totaling frequencies of responses. As well, since the issue of expertise was of interest in the experiment, the responses of 4 experts who were were audit seniors rather than managers were deleted. The final sample size was 117.

This experiment was run in conjunction with four other experiments involving auditor judgment. The total time taken by subjects to complete the set of experiments was about 45 - 50 minutes. This position of this experiment within the set was either first or second with the order counterbalanced. When separate ANOVAs were performed on the response variables with order of the experiments coded as an independent variable, none of the interactions of order with the other effects nor the main effect for order was significant.

Instrument and Experimental Design

This first experiment extends the research begun in Rennie and Johnson (1988) and uses a similar assessment task. The experimental design was a 2^3 factorial between-subjects design. The first independent variable was audit experience (novice auditors vs. expert audit managers and partners). The second independent variable was the number of causes (3 or 6) which subjects were asked to assess. The third independent variable was the order in which the first two causes was listed.

Subjects were asked to assume the role of an auditor investigating possible reasons for 100 disputed account balances from accounts receivable confirmations sent by the auditor to customers of a client firm, St. Albert Corp asking the customers to confirm the balance owing to St. Albert as of December 31, 1987. Subjects were then asked to assess the number of the 100 disputed accounts which would have been due to each of the three (or six) possible causes, plus the number due to "all other" causes not specifically considered. This reference to "all other" causes should have caused subjects to realize that the list was not necessarily complete. Subjects were reminded that their total should equal 100. Only the first three causes (common to both long and short list conditions) were analyzed. The case and the possible causes were adapted from an American Institute of Certified Public Accountants uniform final examination question which appears in a number of current auditing textbooks. Because these textbooks have not been used in the auditing courses at either Saskatchewan university, it is unlikely that subjects would have seen the question previously.

The text of the narrative and instructions was as follows:

ST. ALBERT CORP.

Tom Jones, C.A. has been assigned to the audit of a new client, St. Albert Corp. St. Albert Corp. is a merchandise wholesaler and has a December 31, 1987 year end for accounting purposes. Early in January 1988, as part of the audit procedures for accounts receivable, Tom Jones sent confirmation letters to a sample of St. Albert's customers asking them to confirm the balance owing to St. Albert Corp. as of December 31, 1987. On December 31, 1987 St. Albert's records indicated a balance of accounts receivable of $8,567,250, consisting of 2,000 accounts. Seven hundred and fifty (750) of these accounts with balances totaling $5,907,250 were selected for confirmation. The results are summarized below:

Description	*# of Confirmations*
Signed without comments	*300*
Minor differences (cleared satisfactorily)	*150*
No reply	*200*
Disputed items	*100*
Total Confirmation Requests	*750*

Instructions

On the pages which follow you will be asked to consider three (3) [six (6)] possible reasons for the one hundred disputed items. You will be asked to state the number (of the one hundred) confirmations which you estimate were in dispute for the stated possible reason. Following your assessment of the 3 [6] possible reasons you will be asked to estimate the number of confirmations which were in dispute due to all causes other than those you have specifically considered. The total of your answers should be one hundred, corresponding to the 100 disputed items.

Please consider the possible reasons in the order presented and record your answers in ink. You may refer back to your previous answers in this problem for comparison purposes if you wish, but please answer the questions in the order presented.

The possible causes for the disputed account balances, common across conditions, were:

1. The balance is disputed by the customer because an advance payment had been made by the

customer earlier in 1987. St. Albert had not offset the invoice charges against this advance at December 31, 1987.

2. The balance is disputed by the customer due to a clerical error by St. Albert who failed to post one of the invoices to the customer's account.

3. The balance is disputed because the final payment made early in December 1987 has not been evidenced on the statement accompanying the confirmation as a result of lapping of accounts receivables (i.e. fraud) by the accounts receivable clerk who also handles all deposits (which include cash sales).

The first presented possible cause (cause 1 or 2 above) was on a separate sheet in the experimental booklet. On the following page (backing the first presented cause) was the second cause. The third (and fourth through sixth causes for the long list conditions) plus "all other possible causes" were on the following page (facing the second presented cause).

The fourth through sixth causes in the long list condition were:

4. Balance is disputed by the customer because the balance was paid prior to December 31, 1987

and was incorrectly posted by St. Albert to another customer's account due to a clerical error.

5. Balance is disputed by the customer due to an error in the customer's accounting system which failed to record the invoice.

6. Balance is disputed by the customer because the balance was paid prior to December 31, 1987 and was in transit at year end.

Note that the responses to the fourth through sixth causes were not analyzed because they did not appear on in the short list condition.

On the top of each page adjacent to the possible causes was the heading: *Estimated # of the 100 Disputed Confirmations which were Caused by this Possible Reason.*

The instrument was pilot tested on 55 C.A. finalists who were attending a computer auditing course and expected to write their uniform final examinations in about two months. This pilot test provided some assurance that the tasks were familiar to the subjects and that the wording of the cases and instructions was clear. Although the instrument did not change for the administration to the novices and experts, the finalist groups' responses were not included as a separate level of expertise. This is because approximately 80% of them had participated in the Rennie and Johnson (1988) study as novices about one year prior to this study. Thus these subjects had seen the entire list of possible causes and had been debriefed concerning the omission effect and the availability explanation for the effect. They may therefore have been cued to the omission effect and one of the hypotheses of interest. The responses for this pilot group were consistent with prior knowledge of the hypotheses. That is, there was no significant effect of the list length for two of the three possible causes (i.e. no omission effect). This may mean that if subjects are appropriately cued to be concerned about the completeness of a list, the omission effect may be reduced. These learning effects should be studied further.

In order to assess whether the omission effect of assessments of probable cause in a partial list of possible causes is due to anchoring-and-adjustment, subjects, after reading the case but before the probability assessment task, were either instructed that they would be asked to sequentially assess the probabilities of six or three possible causes for an auditing event (second independent variable). This should have created an anchor of probability 1/6 or 1/3 for each possible cause which subjects assessed. In the assessment task that followed, subjects were instructed to assess each possible cause in the order presented and to record their answers in ink. To ensure lack of review of all presented causes, the first possible cause was listed on a single sheet in the experimental booklet. The second cause was presented on the back of the first sheet and faced the remaining cause(s). Under an anchoring-and-adjustment explanation for the omission effect in the assessment of probable cause, there should be a difference between these two conditions due to the differing initial anchors. Norm theory and availability would not predict any differences, since the specific causes are not yet available.

To assess the effect on probable cause assessments of availability (from norm theory) defined as construction or simulation, the order of the first and second possible causes was exchanged (third independent variable). The third possible cause in both conditions was the fraud related cause (lapping of accounts receivable). These first two possible causes were chosen from the results of Rennie and Johnson (1988) so that both of the causes were of low recall. Low recall items were employed in order to eliminate the effect of recallability by holding it constant at a low level. The fraud explanation was never mentioned as a possible cause by subjects in Rennie and Johnson (1988), so it was also of low recall. When a possible cause is presented first, norm theory would predict that it would be rated higher in normality than when it is presented second. This is because norm theory proposes that the normality of an item is judged as an inverse relation of the number of available alternatives. Where the possible cause is presented first, the other possible causes have not yet been cued and should not yet be available since they are low in recall (see discussion following). Thus the finding of an order effect would be consistent with a simulation notion of availability from norm theory. As well, since fraud was not mentioned by any of the subjects in Rennie and Johnson (1988), any positive frequency assignment to fraud may indicate simulation.

The manipulation of availability in this experiment was not that of recall (as in Rennie and Johnson, 1988) since the causes compared were both of low recall. Instead of testing the effect of differential recall, this manipulation attempted to test the simulation notion of availability. When a low recall cause was in the second position, the first low recall possible cause had already been cued and its relative frequency had been assessed. Thus we know that at least the first possible cause was available. (The remaining cause(s) may also have been available because they were on the page which faced the page which listed the second possible cause.) Where a low recall possible cause was assessed first, we have evidence from Rennie and Johnson (1988) that other low availability causes were not available to all subjects. That is, in Rennie and Johnson (1988), over 90% of subjects in the prior elicitation group did not recall any of the low availability causes. Therefore any order effect should not be attributable to recallability. The manipulation of availability in this experiment was also not that of aided recall (Tulving and Pearlstone, 1966) again since the causes chosen were of low recall.

Note that no order effect would be predicted under anchoring-and-adjustment. That is, even if the order affected the initial anchor, the anchor would be approximately the same (i.e. low) under both orders so that both list length conditions would be similarly affected.

Results

Bartlett's chi-square test was used to test the hypothesis that the correlation matrix of the dependent variables was an identity matrix. The test statistic indicated that none of the three dependent variables were significantly correlated ($p > .05$). Therefore the results were analyzed by separate univariate ANOVAs.

The main effects for the list length (ANCHOR variable) for the three dependent variable possible causes are shown in Figure 3-1. These results indicated a strong omission effect for the different list lengths (anchors) for all three possible causes (the univariate F statistics were all significant at $p < .01$ --see Table 3-2).

The results in Figure 3-1 are consistent with an anchoring-and-adjustment explanation for the omission effect. That is, the experiment provided subjects with anchors equal to the inverse of the list lengths. These anchors could therefore be 1/6 and 1/3. Since the possible causes listed were all low frequency causes, subjects adjusted downwards from their anchor. However, the magnitude of the required adjustment for the three cause list was greater than for the six cause list and there was an insufficient adjustment from the anchor. An insufficient adjustment to the anchor is characteristic of an anchoring-and-adjustment process (Tversky and Kahneman, 1974). The means and standard errors of the means for the possible causes by list length (anchor) condition are shown in Table 3-1.

The means in Figure 3-1 are collapsed over the order of the Advances and Clerical Error possible causes because the order in which these causes were presented was not significant (univariate F ratios all < 1 --see Table 3-2). This non-significant order effect is contrary to the prediction from norm theory regarding probable cause being judged by the availability of other alternatives.

Although the nonsignificant order effect suggested that availability was not a mediating factor in the omission effect, there was some limited evidence in support of availability (in the sense of simulation) as noted by the positive frequency associated with the fraud explanation. Recall that fraud was never mentioned as a possible cause for this case in the free recall task in Rennie and Johnson (1988). Thus any non-zero probability assignment to fraud may indicate simulation. The results indicated that the mean frequency of the FRAUD variable was in fact significantly different from zero (assuming normality). However, the distributions of the fraud variable were skewed right (and bounded at zero). Using a binomial test, at $p = .05$, for the long list group (anchor of six causes), the hypothesis that the median was zero could not be rejected. For the anchor on three causes (short list group), the hypothesis that the median was zero was rejected at $p < .05$. However, the hypothesis that the median of the short list group had a frequency of 5.0 could not be rejected at $p = .05$. These results imply that the median assessed frequency for fraud was very low, although it was positive for the short list group. Thus, in terms of an availability explanation, the ORDER variable was not significant, and the median assessments for the FRAUD variable were also low. There is, therefore, only limited support for availability (as simulation) as a mediating factor in the omission effect.

There was also a significant expertise main effect for the Advances possible cause (univariate $F(1,109) = 11.6$; $p < .001$) (see Figure 3-2). These results indicate that the expert group may be more aware, from their experience, that advances do not occur very often.

The results were analyzed using univariate analysis of variance (Table 3-2). In order to stabilize the variances, logarithmic transformations of the dependent variables were performed. The logarithmic transformation was used because each of the possible causes was low in frequency and was bounded at a value of zero. The logarithmic transformations stabilized the variances for all three of the possible cause dependent variables across groups (Bartlett's test). The arcsin transformation was also used with similar results. However, since the residuals from the univariate ANOVAs were not normal even under the logarithmic transformation for one of the variables (Lilliefor's test, p = .047), a rank transformation (Conover, 1980, p. 337) was used to test the robustness of the effects across different distributional assumptions. The rank transformation stabilized both the variances of the group means and resulted in residuals which were normally distributed (Lilliefor's test), and the significance of the effects was similar to those under the logarithmic transformation. Only one of the significant effects using the untransformed data became nonsignificant when these transformations were performed. This was the Expertise by Anchor interaction for the Clerical error cause which dropped in significance from p = .044 to p = .205 (ranks) and p = .371 (log). A review of the data indicated that there were a small number of experts in the long list (anchor on six causes) condition who assessed Clerical error as having a high frequency. This could be due to a misreading of this cause and an interpretation that it included all clerical errors instead of just clerical errors by the client company. When these outliers were omitted the interaction became non-significant. Because ranked and logarithmic data are more robust to the effects of outliers and because there were no Expertise by Anchor effects for the ADVANCES and FRAUD variables, the nonsignificant result is interpreted as being more reliable.

In the logarithmic and rank transformed data, the three-way interaction among EXPERTISE, ANCHOR and ORDER became significant ($p < .05$) for the FRAUD variable. That is, fraud was assessed at a low frequency in all the long list (anchor on six causes) groups, but the novices, in one of the orders of the short list (anchor on three causes) groups assessed fraud as being higher (median = 12.5). One interpretation for this result is that the novices in this group made an insufficient adjustment to their anchor (based on three possible causes). Experts due to better knowledge of the underlying frequencies of fraud may have made a better adjustment. This is additional evidence of an expertise effect.

Discussion

This experiment provides evidence consistent with an anchoring-and-adjustment explanation for the omission effect. In the manipulation, an anchor of either six or three total causes was provided to subjects. The result was that, although subjects did not preview any of the other possible causes (i.e. the causes were presented sequentially), this anchor affected the responses. In the six cause list (anchor of 1/6), the probability assessments were lower for each of the three possible causes than in the three possible cause condition (anchor of 1/3). Because an anchor was given to subjects, this experiment cannot assess what anchor, if any, subjects actually used in making their frequency assignments. However the results are consistent with subjects having used an anchor based on list length and then having made insufficient adjustments to that anchor for the low frequency of the possible causes presented.

There is little support in this experiment for an availability explanation for the omission effect. The results suggest that experts may assess frequencies differently than do novices. However, the effect of the anchor was no different for the experts in this experiment. The lack of a significant Expertise by Anchor interaction suggests that experts were no less affected by the anchor than were the novices. Perhaps the experimental task (assessing probabilities for discrepancies in accounts receivable confirmations) is not a situation where these experts would likely have a working knowledge of the underlying frequencies. That is, reconciling accounts receivable confirmations is not normally done by partners and managers but by more junior staff and staff are promoted quickly beyond the level at which they do such reconciliations. It is not unreasonable, then, that experts would use an anchoring-and-adjustment strategy for making probability assignments similar to that used by novices. In order to assess expertise, future research should use tasks and domains where knowledge of frequency information would be expected to be part of experts' knowledge representation.

Experiment 2

This experiment examined the effect of focusing on fault tree judgments. Focusing, or thinking about possible causes prior to the probability assessment task, has been shown to reduce, but not eliminate the omission effect (cf. Fischhoff, Slovic and Lichtenstein, 1978, Rennie and Johnson, 1988). Assuming an output interference explanation for the omission effect, focusing prior to making probability assessments should reduce the amount of interference in recall because the associations in memory have been primed just prior to the interference task. However, focusing during the probability assessment task should not reduce the omission effect since the interference will already have occurred by the subject reading the list of presented causes.

This experiment manipulated focusing at four different levels: no focusing, minimal focusing, prior focusing and concurrent focusing. The minimal focusing manipulation was similar to that of Fischhoff et al. (1978) where subjects were told in the experimental narrative that the list may be incomplete. Prior focusing (as in Rennie and Johnson, 1988) involved having subjects list possible causes before the probability assessment task. In concurrent focusing, subjects were asked to specifically list "other causes" during the assessment task. Assuming an output interference explanation for the omission effect it was predicted that the omission effect would be reduced under prior focusing and would not be reduced under concurrent focusing. Thus for the focusing by omission interaction, there should have been a significant contrast between the prior focusing and no focusing conditions. As mentioned above, prior focusing should have increased the associations of all possible causes which were listed, thereby reducing the amount of output interference when the partial list was presented.

For concurrent focusing, there was no priming of the associations for the omitted causes, so output interference should still have occurred. Thus, for the focusing by omission interaction, no significant contrast between concurrent focusing and no focusing conditions was predicted.

Minimal focusing may reduce the omission effect under an output interference explanation. This would mean that should have been a significant contrast between minimal focusing and no focusing for the focusing by omission interaction. However, minimal focusing, as a practical measure is of limited benefit, since all lists are necessarily incomplete. Thus if a reduction in the omission effect were found only for the minimal or concurrent focusing groups, this may suggest that the focusing result was due to experimental demands.

The design of the experiment was a 4 x 2 x 2 between subjects factorial of focusing (at four levels), omission (list length of two or four causes) and expertise (novice auditors and partner/managers). The dependent measures were the frequencies out of 100 for the two causes which were common to both lists.

Method

Subjects and Administration

The subjects for this experiment were the same 130 auditors as in experiment 1. Subjects were randomly assigned to the experimental conditions. After deleting the four seniors who had not yet been promoted to the manager level and six responses which were not complete or did not total to 100, the final sample size was 120.

This experiment was the last of five experiments which subjects were asked to complete in the experimental session. The total time taken by the subjects was about 45 - 50 minutes. This experiment was positioned following two auditor judgment experiments unrelated to fault trees, so that these prior tasks could serve as a distraction from the first two fault tree experiments. As in experiment 1, when a separate MANOVA was performed on the response vector including administration mode (group or self-administered) as a separate effect, neither response mode nor any of the interactions of response mode with the other experimental effects were significant.

Instrument

The experimental narrative was adapted from a current auditing text book. Since this text had not been used at either Saskatchewan university, it was unlikely that subjects would have been familiar with the question. The narrative was pilot tested on the same group of finalists as in experiment 1. The narrative asked subjects to assume the role of an auditor assessing the probability that weaknesses in an internal control system of a nonprofit art gallery would cause an error or omission to cash receipts. Following the description of the control features of the art gallery (end of second paragraph of the narrative -- see below), subjects in the prior focusing condition were asked to list potential weaknesses in internal control over cash receipts. Subjects in all conditions (on the following page of the experimental booklet) were asked to assess the "# of times out of 100 that this weakness would result in an error or omission to cash receipts." Subjects in the prior focusing condition were instructed that they could look back at the weaknesses they had listed in assessing the probabilities of the listed weaknesses. The listed weaknesses were described as having been identified by a previous auditor. In the minimal focusing group subjects were told to think about possible weaknesses which were not listed by the previous auditor: "In your consideration of all other potential weaknesses, think specifically about possible weaknesses which may not have been listed by Mr. Horwich [the previous auditor]." Subjects in all groups were asked to assess the frequency of all other weaknesses not listed, with subjects in the concurrent focusing group being asked to list these other weaknesses. Subjects in all conditions were reminded that their totals should equal 100.

The text of the experimental narrative for the no focusing conditions was as follows:

Friends of Seven Society

The Friends of Seven is an art appreciation society which operates a gallery to promote Canadian paintings. When the gallery is open to the public, two clerks who are positioned at the entrance collect a five dollar admission fee from each nonmember patron. Members of the Friends of Seven Society are permitted to enter free of charge upon presentation of their membership cards.

At the end of each day, one of the clerks delivers the day's cash receipts to the treasurer. The treasurer counts the cash in the presence of the clerk and places it in a safe. Each Friday afternoon the treasurer and one of the clerks deliver all cash held in the safe to the bank, and receive an authenticated deposit slip that provides the basis for the weekly entry in the cash receipts journal.

Paul Horwich, Chartered Accountant, had acted as auditor for the society on a donated fee basis since the society was founded. In response to a recent management letter sent to the society by Mr. Horwich, the board of directors of the Friends of Seven Society discussed the possibility of improving their system of internal control over cash admission fees. The board has decided that altering the physical layout of the gallery by installing turnstiles, etc. would not be cost effective. None of Mr. Horwich's recommendations had been acted upon prior to his amicable and voluntary retirement as the Society's auditor, due to poor health.

Assume that your firm has been appointed as the new auditors of the Friends of Seven Society this year and you are in charge of the audit. You have been asked to draft the management letter to the board of directors outlining possible weaknesses in internal control over cash receipts. Since the society has limited resources to deal with the weaknesses you identify, the board of directors has asked you to indicate which weaknesses are most important, in terms of potentially causing errors or omissions in cash receipts.

Required:

Listed below are a number of potential weaknesses in internal control over cash receipts for the Society which were identified by Mr. Horwich, the previous auditor. Please rate each of these potential weaknesses, together with all other potential weaknesses not listed, on a scale of 0 to 100 in terms of their likelihood of causing a possible error or omission in cash receipts. Since you will have considered all potential weaknesses, your total should equal 100.

The two weaknesses which were common to both the long and short lists were:

1. There is no basis for establishing the documentation of the number of paying patrons. (variable PATRONS)

2. There is no segregation of duties between persons responsible for collecting admission

fees and persons responsible for authorizing admission. (variable FEES)

The third and fourth weaknesses (not analyzed since they appeared in the long list condition only) were:

3. An independent count of paying patrons is not made.

4. There is no proof of accuracy of amounts collected by the clerks.

Results

Bartlett's chi-square test was used to examine the hypothesis that the correlation matrix of the dependent variables was an identity matrix. The test statistic indicated that the dependent variables were significantly correlated (correlation coefficient = -.356; chi-square(1) = 7.94; $p < .005$). Thus both multivariate and univariate analyses of variance results were performed.

The means of the FEES and PATRONS variables by length of list and across focusing conditions are shown in Figures 3-3 and 3-4. The overall omission effect was significant (multivariate $F(2,103) = 22.3$; $p < .001$) Note in Figure 3-4 that there was a relatively constant omission effect for the PATRONS variable across all focusing conditions. This omission effect was significant for the PATRONS variable (univariate $F(1,104) = 23.7$; $p < .001$). For the FEES variable (Figure 3-3), note that again there was an omission effect across all focusing conditions (univariate $F(1,104) = 4.3$; $p < .05$). However, the omission effect was quite small in the no focusing and concurrent focusing conditions.

The interaction of focusing and omission (see Figures 3-3 and 3-4 and Table 3-3) was not statisically significant ($F < 1$). As well, none of the three planned contrasts on the focusing by omission interaction between the no focusing condition and the three focused conditions were significant for either dependent variable (all p's > .25 -- see Tables 3-4, 3-5 and 3-6) Thus focusing did not significantly reduce the omission effect.

There was a significant main effect for focusing (multivariate $F(6,206) = 2.4$; $p < .05$). The focusing manipulation was not significant for the PATRONS variable (univariate $F < 1$). The main effect for focusing was significant for the FEES variable (univariate $F(3,104) = 3.4$; $p < .05$). Orthogonal contrasts on this effect indicated (Table 3-7) that the concurrent focusing variable was different from the other conditions and led to higher assessed mean frequencies than did the other focusing conditions.

Unlike previous studies (except Hirt and Castellan, 1988), there was no significant reduction in the omission effect through focusing (univariate and multivariate F statistics < 1). None of the contrasts on the interaction effect approached significance. There was also no effect of expertise in this experiment.

The results were analyzed by multivariate and univariate analyses of variance on the responses to the two common causes, or internal control weaknesses (Table 3-3). Using Bartlett's test, the assumption of homogeneity of group variances could not be rejected for either variable. However for the multivariate results, using Box's M test, the assumption of homogeneity of covariance matrices was rejected ($p > .05$). Because the cell sizes were similar in size, this heterogeneity should not have greatly affected the results.

When the univariate residuals were tested for normality (Lilliefor's test), the normality assumption was rejected for the FEES variable. Therefore, to test the robustness of the effects over different distributional assumptions, a rank transformation was done (Conover, 1980, p. 337). The results were unchanged from the parametric ANOVA. The assumption of bivariate normality of the residuals from the MANOVA was further assessed by a chi-square plot of the ordered squared residuals against a chi-square distribution (Johnson and Wichern, 1982, p. 158). This plot indicated a straight line (the correlation was .991) suggesting that the assumption of bivariate normality was reasonable.

Discussion

This study demonstrates that focusing does not always reduce the omission effect. Subjects may have found the task difficult in that they may not normally think of the relative importance of internal control weaknesses in probability terms. That is, the decision to remedy an internal control weakness may be dominated by cost factors, and not the probability of a material cash shortage or error. For example, the experimental narrative indicated that the art gallery had limited funds to deal with any internal control weaknesses. This was mentioned in order to provide subjects with a justification for assigning probabilities to the weaknesses.

Under an output interference explanation for the omission effect, prior focusing should have reduced the amount of interference. This reduction in interference would result from the priming of associations (within memory) of the items omitted from the list. However, to obtain a positive effect from focusing, there must be associations in memory able to be primed. That is, subjects must have access to possible weaknesses or be able to construct such weaknesses. If situations such as this experimental task are dominated by cost factors, then there may be no associative memory to activate. This could explain the failure to reduce the omission effect through focusing. That is, there was no output interference which could be reduced through focusing. This explanation may also mean that output interference is not a necessary condition for the existence of an omission effect, since a significant omission effect was found in this experiment. Thus the positive effects from focusing found in previous studies may have been in situations where output interference did result. Hirt and Castellan's (1988) results which failed to find any reduction in the omission effect through focusing were similar to this experiment. That is, omission effects were found, but they were not reduced through focusing. These results suggest that even if output interference is sufficient to cause an omission effect, it is not necessary. The role of output interference in the omission effect was explored further in the next experiment.

Some of these results could be interpreted in terms of Hirt and Castellan's redefinition of the categories hypothesis. In the common causes listed in the experiment, there was some degree of overlap. That is, although the listed causes were meant to be discrete, it might have been possible for subjects to consider the documentation of number of paying patrons (PATRONS) and the segregation of duties to be related, in that number of patrons times admission price (FEES) equals total cash receipts. Thus some amount of recategorization could have occurred, and some of the assessed probability for one of the dependent variables could have been assigned to the other dependent variable. The responses, however, were not consistent with this hypothesis. That is, when the sum (for each subject) of the frequencies for the listed weaknesses was run as the dependent variable in an ANOVA, there was still a highly significant omission effect ($p = .004$). Thus subjects in the long list condition assigned lower average frequencies to the listed weaknesses than did subjects in the short list condition. This means that subjects were not simply redefining the listed categories.

The failure to find a significant effect for prior focusing means that the output interference hypothesis does not explain the results of this experiment. In fact, the only significant effect for focusing was the main effect for focusing for the FEES variable and that was not due to prior focusing but to concurrent focusing. Alternatively, the lack of a significant interaction between focusing and the omission effect could mean that the subjects had no knowledge of what was missing from the list. This might be expected for the novices, but not for the expert group.

A possible alternative reason for nonsignificance of the focusing effect may have been fatigue. Recall that this experiment was the last of five experiments in a 45 - 50 minute session. As evidence against a fatigue hypothesis it may be noted that subjects in the prior elicitation condition listed an average of 3.1 possible internal control weaknesses each, and only one subject did not list any possible weaknesses.

The results of this experiment were not consistent with either category redefinition or with output interference. Although both of these phenomena may account for some of the variance in the omission effect in some fault tree judgments, they do not appear to be have been mediating factors here. On the other hand, anchoring-and-adjustment may explain some of these results. That is, the omission effect may have resulted from subjects anchoring on the number of causes in the long and short lists and making an insufficient adjustment for the actual frequency in the short list. In this experiment there was a greater proportionate omission effect for the higher frequency variable (PATRONS) (see Figure 3-4). For the FEES variable, as shown in Figure 3-3, the assessed frequencies were already quite low in the long list. Therefore less of an adjustment in the short list was necessary and the omission effect was correspondingly lower. Thus it appears that the omission effect was not uniform between the two dependent variables. When the hypothesis that the two univariate models were different was tested, the results were significant ($F(16,104) = 2.959$; $p < .001$).

Experiment 3 further examined the effects of focusing.

Experiment 3

This experiment examined further the output interference explanation for the omission effect. Because output interference should not be affected by concurrent focusing (viz. the partial list has already been presented and the interference has occurred) only prior focusing was manipulated in this experiment. Prior focusing may reduce output interference because the associations in memory for the omitted list items have been primed just prior to the interference task. In Rennie and Johnson's (1988) results, it was shown that the omission effect was no different for high and low recall possible causes. In the first experiment in this paper, an omission effect was shown for low recall possible causes. This study used high recall causes to examine the omission effect.

By using high recall items which had different frequencies, the results of this experiment may be able to comment on the availability explanation for the omission effect. That is, in many situations availability and ecological frequency are confounded due to their high correlation (Tversky and Kahneman, 1973). Under an availability explanation for the omission effect, we would expect to find similar omission effects for each high recall cause. If the omission effects are different between the causes (as was found in this experiment), this suggests that some other factor is mediating the omission effect.

The design of the experiment was a 2^2 factorial of focusing (present or absent) and omission (long or short list). The focusing variable was manipulated by having subjects in the focusing-present condition perform a free recall of a list of auditing assertions (which formed the basis for the long and short lists in the assessment task). Subjects in the focusing-absent condition performed a free recall of auditing assertions after the assessment task so that the prior recall of the list items and the effects of interference could be assessed. The assessment task involved subjects allocating 100 auditing hours among two or four audit assertion areas (second independent variable) plus all other areas. Subjects in all conditions were also asked to list and assess the hours for any areas not listed (concurrent focusing) with the number of blank lines plus listed items being equal across list length (omission) conditions. The purpose of this concurrent focusing was to minimize any demand characteristics associated with the different list lengths. The dependent variables were the number of hours allocated to the two audit assertion areas which were common to the short and long lists.

By using number of hours (rather than probability or frequency) as the dependent variable, this experiment was able to show that the omission effect does not exit simply because an unnatural response scale (i.e. probability) had been used in the other experiments and in past studies. Using hours as the response scale also meant that the response scale value and the recallability of the listed items would not necessarily be highly correlated.

Assuming an output interference explanation for the omission effect, the focused subjects should have had a lower omission effect than did the unfocused subjects. Thus output interference predicted a significant FOCUSING x OMISSION interaction for both dependent variables.

Method

Subjects and Administration

The subjects were 64 auditors, registered as students with the Institute of Chartered Accountants of Saskatchewan, who were attending an auditing procedures course. The subjects were volunteers and one subject (not included in the 64) declined to participate. The responses for three subjects were deleted due to addition errors in their responses, leaving a final sample size of 61. The subjects had an average of 6.9 months of auditing experience.

The experiment was conducted in a classroom setting by the Director cf Education of the Institute of Chartered Accountants of Saskatchewan. Subjects were randomly assigned to each condition. The experimental session, including debriefing, took less than 15 minutes.

Instrument

The instrument was reviewed with a number of practicing chartered accountants and academics prior to the experiment for clarity of wording and instructions. The experiment consisted of two parts: 1) a free recall of audit assertions and 2) an allocation of 100 audit hours among two or four listed audit assertion areas plus all other areas. Subjects in the focusing present condition performed the free recall task prior to the assessment of the audit hours.

Subjects were expected to be very familiar with the audit assertions in the task, since this was one of the syllabus areas of the audit procedures course which they were attending. Section 5300.17 of the Handbook of the Canadian Institute of Chartered Accountants (C.I.C.A.) lists the following assertions regarding financial statement items: existence, occurrence, completeness, ownership, valuation, measurement, statement presentation. The experiment was administered towards the end of the course to ensure that audit assertions had been covered in class.

The experimental narrative asked subjects to assume the role of an audit senior who was planning the allocation of the 100 audit hours assigned to the audit of inventory among the various audit assertions which the audit is trying to satisfy. The client company was described so that a large number of audit assertions would be relevant to the audit. The text of the narrative and the instructions was as follows:

Audit Judgment Situation

In this situation you are to assume the role of an audit senior making planning decisions for the inventory section of the yearend audit of Consumer Products Ltd., which is a new audit client of your firm. Some background information concerning Consumer Products Ltd. is as follows:

Consumer Products Ltd. is a successful merchandiser of small kitchen gadgets and appliances. Many of its sales are made through television commercials where interested purchasers may immediately telephone a local number to place an order. As well, the company distributes its products through a number of retailers on a consignment basis. Most of the gadgets are impulse or fad items. Accordingly, they have to be sold at a reduced price once the initial promotion campaign is over.

Required:

Your audit manager has advised you that he roughly expects the inventory section of the audit to take about 100 hours. Based on the description of the company above, he has asked you to estimate how many hours you expect should be allocated to each of the following areas:

The two audit assertion areas which were common to both the short and long list conditions were: valuation of the inventory (VALUATION) and ownership of the inventory (OWNERSHIP). The third and fourth assertions (for which the responses were not analyzed) which appeared in the long list condition were: disclosure in accordance with generally accepted accounting principles; and completeness/cutoff of transactions. The instrument reminded subjects that the total hours allocated should equal 100.

Results

Bartlett's chi-square test was used to examine the hypothesis that the matrix of dependent variables was an identity matrix. Because the test statistic indicated a significant correlation ($r = -.356$; chi-square(1) = 7.943; $p < .005$) between the dependent variables, the results were analyzed by multivariate and univariate analyses of variance.

The means of the ownership and valuation variables are shown in Figure 3-5. There was an overall omission effect (multivariate $F(2,56) = 4.0$; $p < .05$). There was a significant omission effect for the OWNERSHIP variable which had mean assessments of 15.1 hours in the long list condition and 24.3 in the short list condition (univariate $F(1,57) = 7.2$; $p = .009$). For the VALUATION variable, however, the means for both list length conditions were about 42 hours and there was consequently no significant omission effect (univariate $F < 1$). Thus the two dependent variables appeared to be different. A contrast of the two dependent variable vectors in the MANOVA model indicated that the two models for the dependent variables were different ($F(4,57) = 12.885$; $p < .001$)

The auditing assertions were highly available to subjects. Subjects in the focusing-present condition listed an average of 7.5 auditing assertions (some subjects listed additional assertions beyond those listed in the C.I.C.A. Handbook such as validity and reliability). In the focusing-absent condition, where subjects recalled assertions after the assessment task, the mean number of assertions listed was 7.0. An analysis of variance on the means across the four experimental conditions indicated no significant differences in mean numbers of listed assertions.

All subjects in both focusing conditions listed the valuation assertion and only two of the 61 subjects (both in the post-elicitation group) failed to list the ownership assertion. This result implies that there was no output interference in the assessment task. Yet, there was a significant omission effect for the OWNERSHIP variable. These results also suggest that the assertions were highly available to subjects prior to the assessment task and were not cued by the task.

There was no difference between the two focusing conditions for either dependent variable, nor was the omission effect for the OWNERSHIP variable reduced through focusing (see Table 3-8). That is, contrary to the output interference prediction, the focusing by omission interaction was not significant (univariate and multivariate F's < 1). Thus if output interference is a mediating factor in the omission effect, it is not a necessary condition. That is, there was an omission effect, yet there was no output interference.

The results were analyzed by univariate and multivariate analyses of variance (Table 3-8). To stabilize the variances of the OWNERSHIP variable, a logarithmic transformation was performed. After this transformation the variance homogeneity assumption (Bartlett's test) and the assumption of normality of the residuals (Lilliefor's test) could not be rejected. The significance of the effects was the same as that of the untransformed data. (Similar results were obtained using an arcsin transformation). The assumptions of homogeneity of variances and normality of the residuals could not be rejected on the untransformed data for the VALUATION variable.

To assess the equality of the covariance matrices for the MANOVA, Box's M test was conducted. The hypothesis of equality could not be rejected ($p < .05$) on the untransformed data. A chi-square plot (Johnson and Wichern, 1982, p. 158) was used to provide further evidence concerning the reasonableness of the assumption that the residual matrix was bivariate normally distributed. When the ordered squared residuals were plotted against their chi-square values, the result was a straight line (the correlation was .972). This result indicates that the assumption of bivariate normality was a reasonable one.

Discussion

This experiment replicated the focusing results from the second experiment and demonstrates that prior focusing does not consistently reduce the omission effect. Prior focusing should have reduced output interference of the omitted list items, if interference is the mediating factor in the omission effect. This is because focusing should have activated the associative links in memory between the task and the omitted items in the list. However, in this experiment, the free recall task results suggest that there was no output interference. Because no omission effect was shown for one of the variables while a significant omission effect was shown for the other variable, this result suggests that it was not output interference that mediated the omission effect in this experiment. That is, an output interference explanation could not account for the significant omission effect for one variable and lack of an omission effect for the other variable.

Because the audit assertions do not overlap, the significant omission effect which was found can similarly not be accounted for by recategorization by subjects.

In this experiment the lists of items were highly recallable by subjects. Yet there was an omission effect only for one of the variables. Such a finding cannot be accounted for by the availability explanation for the omission effect because, under an availability explanation, we would expect similar omission effects for items which had a similar degree of recall.

These results suggest that focusing may be effective where the list items are ones that must be constructed or simulated by the subjects, rather than recalled from memory as was the case in this experiment. This is consistent with Hoch's (1984) observation that output interference is more likely to occur in novel judgments than in situations involving direct recall from memory.

The finding of a significant omission effect for one variable and no omission effect for the other variable can be accounted for by an anchoring-and-adjustment process for the assignment of probabilities (or in this case allocation of hours). In the case of the nonsignificant omission effect for the VALUATION variable, subjects may have adjusted upwards or downwards from their anchor to the mean assessment of about 42 hours. If we assume that subjects in the long list (four audit assertions) had an anchor of about 25, then they adjusted upwards for the additional importance by 17 hours. In the short list (two audit assertions) subjects might have anchored on 50 hours and adjusted downward by eight hours to 42 hours. Because the adjustments required were both small, similar in size, and in opposite directions, the means could have turned out to be similar. For the OWNERSHIP variable, however, the adjustments were not symmetric. For both the long and short lists there was a need to adjust downwards from the anchor. In the short list there was a very large downward adjustment and the adjustment may not have been sufficient to prevent an omission effect.

An anchoring-and-adjustment strategy for the assignment of numerical values to the list items can explain the lack of an omission effect in this experiment. Note however that the assumed anchors of 25 and 50 (based on list lengths of 4 and 2 items respectively) are crucial to the explanation. If one alternatively assumes anchors which include "all other" (i.e. list lengths of 5 and 3), then the anchors become 20 and 33. With such anchors, there would be no possibility that the offsetting adjustments from the anchors could result in similar means with a value of 42. However, it is not likely that subjects would have included "all other" as part of their anchors. This is because the experimental booklets presented blank lines and asked subjects to write in "all other" with the number of listed areas plus blank lines being equal across conditions. Therefore the only reasonable anchors, based on list length, which could be assumed, are 50 and 25 (as above). Future research should attempt to determine what anchor, if any, subjects actually used in making their numerical estimates.

The lack of an omission effect for one variable could also suggest an expertise effect on the part of the auditor subjects. However, it is quite likely that ownership is not the most important audit assertion and that determining the physical existence through observing the inventory count may take more audit hours than does determining ownership.

The results of this experiment are consistent with an anchoring-and-adjustment process. This experiment may also provide insight into why availability has often been interpreted as a mediating factor in the omission effect. That is, the availability of an item will usually be highly correlated with its ecological frequency (Tversky and Kahneman, 1973). Thus differing availability may be seen as resulting in different omission effects when what is actually happening is that the different omission effects are caused by different frequencies in relation to the list length or anchor. This experiment used list items which were highly available but which had different perceived frequencies. Thus, this experiment was able to differentiate frequency and availability. This experiment was also able to show that the omission effect applies to lists in general and not just those where the response scale is probabilities.

GENERAL DISCUSSION

This study examined the phenomenon that when people, (including experts) are asked to assess the probabilities of a partial list of possible causes (as in a fault tree), they tend to underestimate the probability of what is missing from the list. This phenomenon was called the omission effect. The results supported an anchoring-and-adjustment explanation for the omission effect.

Because this anchoring-and-adjustment explanation has not been previously studied, it is difficult to fully interpret the results of prior studies in terms of this explanation. However, in Fischhoff et al.'s (1978) study (Tables 1 and 5 pp. 335 and 342), the mean assessed frequency for each listed possible cause was higher for the short list than the corresponding frequency for the long list. This is consistent with subjects anchoring on a higher probability value in the short list. In Dube-Rioux and Russo (1988) only aggregated data were presented and it is not possible to examine the effect of list length on the individual listed causes. Because Hirt and Castellan (1988) indicated to subjects that the long and short lists were from different automobile manufacturers, the resulting differential categorization would not make such a comparison of frequencies meaningful.

This study also examined the availability explanation and the output interference explanations for the omission effect. The results of these experiments were not consistent with either of these alternative explanations.

The availability explanation is rejected from the results of Rennie and Johnson (1988) and from the results of experiments one and three. That is, there was no difference in the omission effect for high and low recall lists. As well, there was no mediation of the omission effect when low availability causes were cued in the first experiment. Finally, in the third experiment, there was an omission effect for list item and not the other, yet both items were highly available.

The output interference explanation for the omission effect is rejected from the results of the second and third experiments. In the third experiment there was no output interference yet there was an omission effect. As well, there was no reduction in the omission effect for prior focusing in the second experiment.

It is quite possible that, in other domains, these and other explanations may account for a significant portion of variation in subjects' responses. For example, output interference may be a sufficient condition for the occurrence of the omission effect, but, as shown in this study, it is not necessary for an omission effect to result. Thus, as shown in previous studies, focusing may sometimes reduce the omission effect.

The results of this study suggest that when subjects are asked to assess the probabilities of a list of possible causes, they may anchor on a probability estimate which is the reciprocal of the number of listed causes. They then increase or decrease their estimate from the anchor based on their perception of the underlying frequency. When large adjustments are needed (e.g. for low frequency causes), the adjustment may not be sufficient. This insufficient adjustment through the anchoring-and-adjustment process is similar to that found by Tversky and Kahneman (1974) and by Einhorn and Hogarth (1987) in belief revision.

Future research should continue to explore the anchoring-and-adjustment and other explanations for the omission effect. Since assessment of the probabilities of possible causes is vital to so many areas, prescriptive measures to improve the quality of such judgments may be needed. However, an anchoring-and-adjustment explanation for the omission effect does not immediately lead to any easy means of alleviating this potential problem. While focusing, or thinking about other causes, has been shown to sometimes reduce (but not eliminate) the omission effect, this study demonstrates that the effects of focusing are inconsistent. Anchoring-and-adjustment suggests that decision-makers should attempt to use lists that are as complete as possible. Some limited expertise effects were also found in this study; however, experts' responses were also consistent with anchoring-and-adjustment. Thus if individual experts can only be relied on to a limited extent to provide complete lists, perhaps future research should examine whether groups of experts working together are able to overcome the omission effect. Future studies should also examine domains where normative frequency data exist so that the magnitude of the omission effect can be assessed.

Table 3-1

Experiment 1

Means of Possible Causes by List Length

Possible Cause	List Length: Three Causes Mean	S.E.	Six Causes Mean	S.E.
	(n = 58)		(n = 59)	
Advances	18.6	2.4	8.6	1.3
Clerical Error	17.9	2.1	8.1	1.1
Fraud	10.1	2.0	3.8	.9

Table 3-2

Experiment 1

Analysis of Variance on Means of Possible Causes

	Univariate F-Ratios (1,109 df) Possible Causes:		
Effect	Advance	Clerical	Fraud
Expertise (E)	11.6***	< 1	2.8
Anchor (A)	14.2***	14.8***	8.7**
Order (O)	< 1	< 1	< 1
E x A	2.1	4.2*	1.2
E x O	< 1	< 1	< 1
A x O	1.1	< 1	< 1
E x A x O	< 1	< 1	2.1

* $p < .05$
** $p < .01$
*** $p < .001$

Table 3-3

Experiment 2

Analysis of Variance on Means of Possible Causes

Effect	Univariate F-Ratios (df)	Possible Causes: Patrons	Fees	Multivariate F (df)	(Wilk's Lamba)
Omission(O)	(1,104)	23.7***	4.3*	(2,103)	22.3***
Focusing(F)	(3,104)	< 1	3.4*	(6,206)	2.4*
Expertise(E)	(1,104)	< 1	< 1	(2,103)	< 1
O x F	(3,104)	< 1	< 1	(6,206)	< 1
O x E	(1,104)	< 1	< 1	(2,103)	< 1
F x E	(3,104)	1.1	< 1	(6,206)	< 1
O x F x E	(3,104)	< 1	< 1	(6,206)	< 1

* $p < .05$
*** $p < .001$

Table 3-4

Experiment 2

Planned Contrast on Focusing by Omission Interaction -- Prior Focusing vs. No Focusing

	Dummy Variables for Effects							
		Omission	Focus			Interaction		
	C	O	f_1	f_2	f_3	Of_1	Of_2	Of_3
Treatment								
No Focusing								
--long list	1	1	1	0	0	1	0	0
--short list	1	-1	1	0	0	-1	0	0
Difference (omission effect)	0	2	0	0	0	2	0	0
Prior Focusing								
--long list	1	1	-1	-1	-1	-1	-1	-1
--short list	1	-1	-1	-1	-1	1	1	1
Difference (omission effect)	0	2	0	0	0	-2	-2	-2
Contrast	0	0	0	0	0	4	-2	-2

FEES variable: $F(1,112) = .182$; $p = .670$
PATRONS variable: $F(1,112) = .007$; $p = .933$

Table 3-5

Experiment 2

Planned Contrast on Focusing by Omission Interaction -- Concurrent Focusing vs. No Focusing

	Dummy Variables for Effects							
		Omission	Focus			Interaction		
	C	O	f_1	f_2	f_3	Of_1	Of_2	Of_3
Treatment								
No Focusing								
--long list	1	1	1	0	0	1	0	0
--short list	1	-1	1	0	0	-1	0	0
Difference (omission effect)	0	2	0	0	0	2	0	0
Concurrent Focusing								
--long list	1	1	0	0	1	0	0	1
--short list	1	-1	0	0	1	0	0	-1
Difference (omission effect)	0	2	0	0	0	0	0	2
Contrast	0	0	0	0	0	2	0	-2

FEES variable: $F(1,112) = .089$; $p = .766$

PATRONS variable: $F(1,112) = .020$; $p = .866$

Table 3-6

Experiment 2

Planned Contrast on Focusing by Omission Interaction -- Minimal Focusing vs. No Focusing

	Dummy Variables for Effects							
		Omission	Focus			Interaction		
Treatment	C	O	f_1	f_2	f_3	Of_1	Of_2	Of_3
No Focusing								
--long list	1	1	1	0	0	1	0	0
--short list	1	-1	1	0	0	-1	0	0
Difference (omission effect)	0	2	0	0	0	2	0	0
Minimal Focusing								
--long list	1	1	0	1	0	0	1	0
--short list	1	-1	0	1	0	0	-1	0
Difference (omission effect)	0	2	0	0	0	0	2	0
Contrast	0	0	0	0	0	2	-2	0

FEES variable: $F(1,112) = 1.181$; $p = .280$
PATRONS variable: $F(1,112) = .093$; $p = .761$

Table 3-7

Experiment 2

Orthogonal Contrasts on Focusing Main Effect:

Means of FEES variable

Focusing Condition	Mean	(S.E.)		Orthogonal Contrasts 1st	2nd	3rd
No Focusing	21.3	(2.9)		+1	+1	+1
Minimal Focusing	23.5	(3.1)		+1	+1	-1
Concurrent Focusing	32.7	(3.9)		-3	0	0
Prior Focusing	18.4	(2.5)		+1	-2	0
			p	.003	.247	.685

Table 3-8

Experiment 3

Analysis of Variance on Mean Budgeted Hours

Effect	Univariate F-Ratios (1,57 df) Listed Areas: Valuation	Ownership	Multivariate F (2,56 df)
Omission	< 1	7.2**	4.0*
Focusing	2.8	< 1	2.9
Omission x Focusing	< 1	< 1	< 1

* $p < .05$
** $p < .01$

Figure 3-1
Experiment 1
Main Effect for Anchor Variable

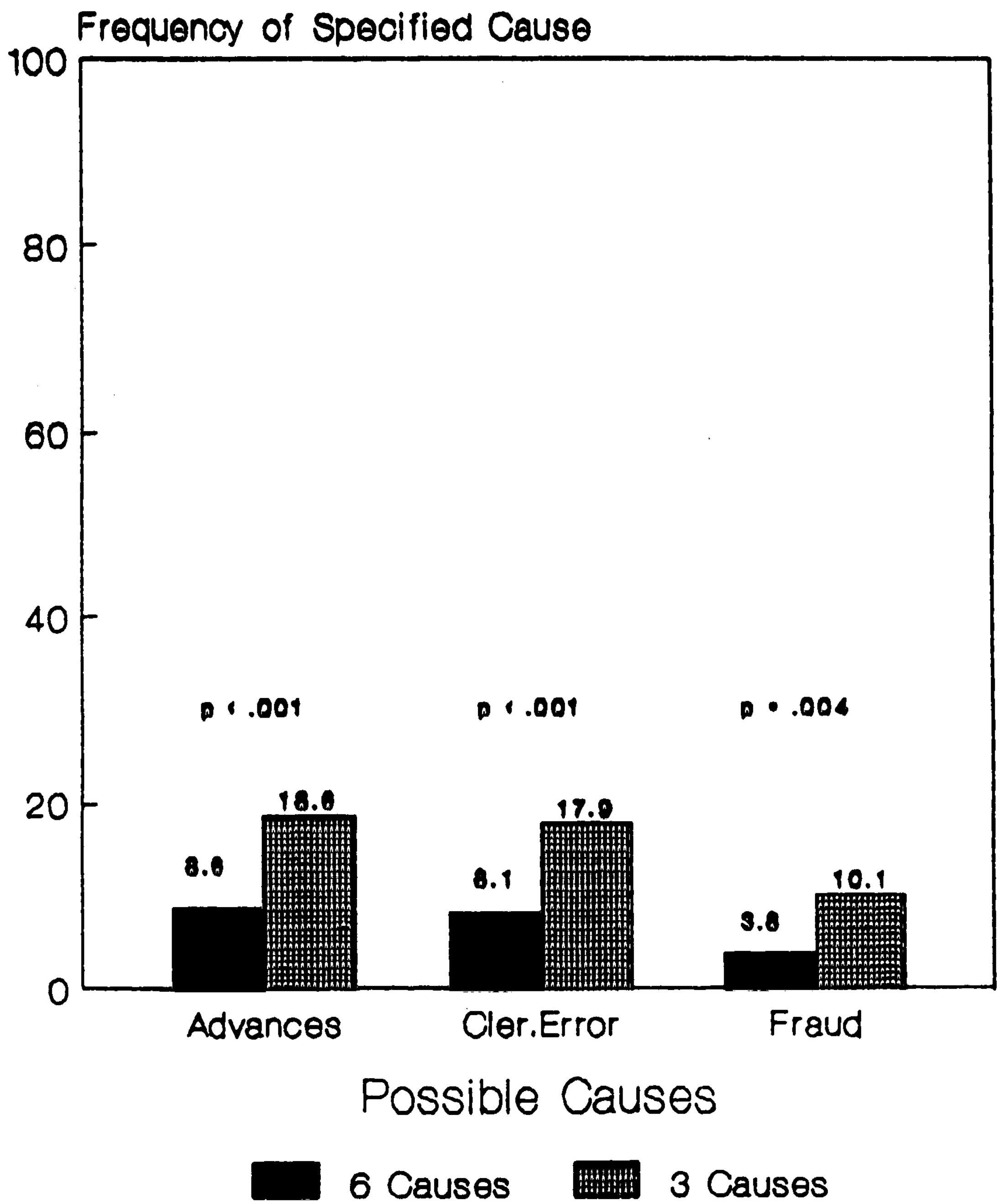


Figure 3-2
Experiment 1
Expertise Main Effect

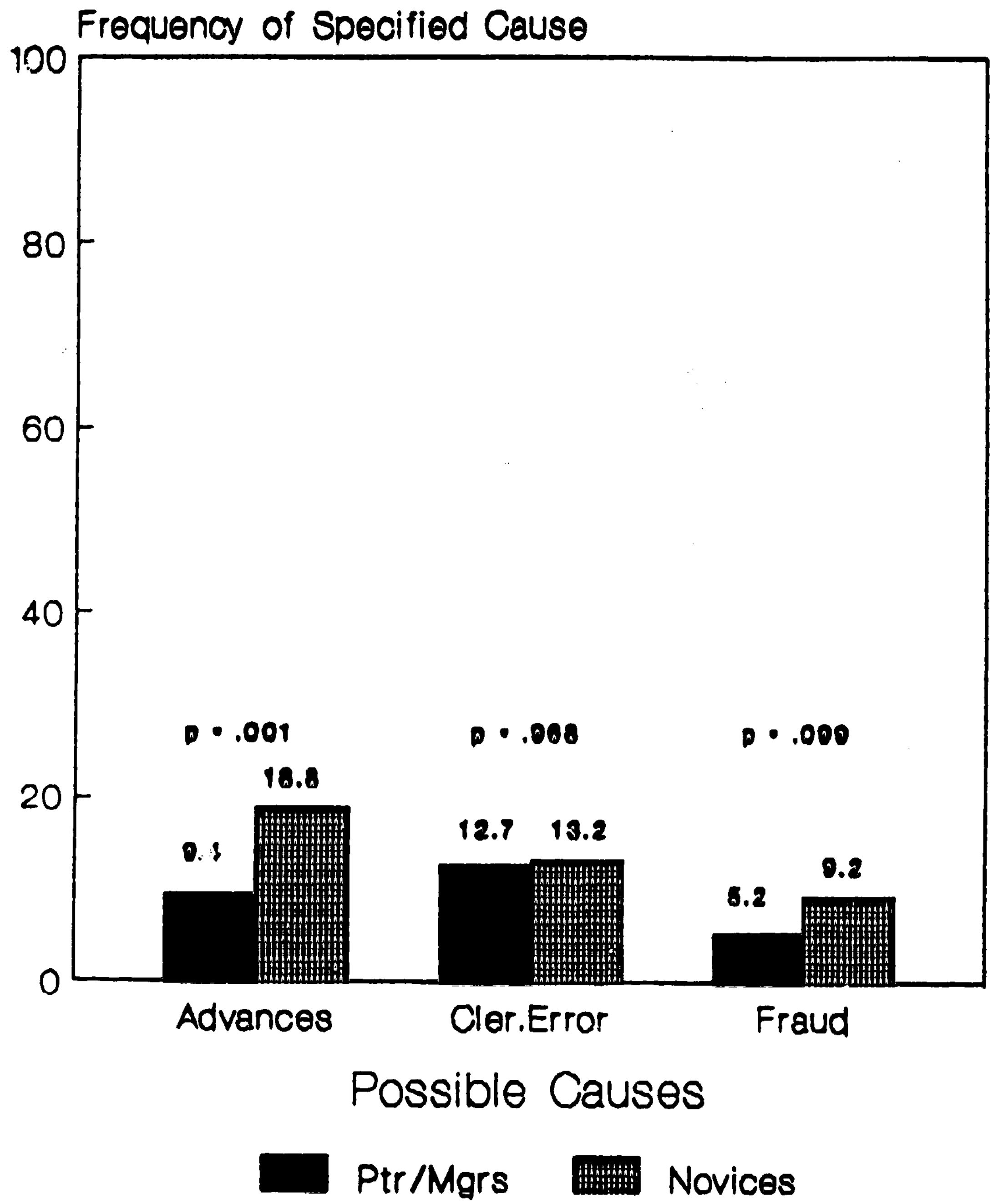


Figure 3-3
Experiment 2
Omission Effect by Focusing

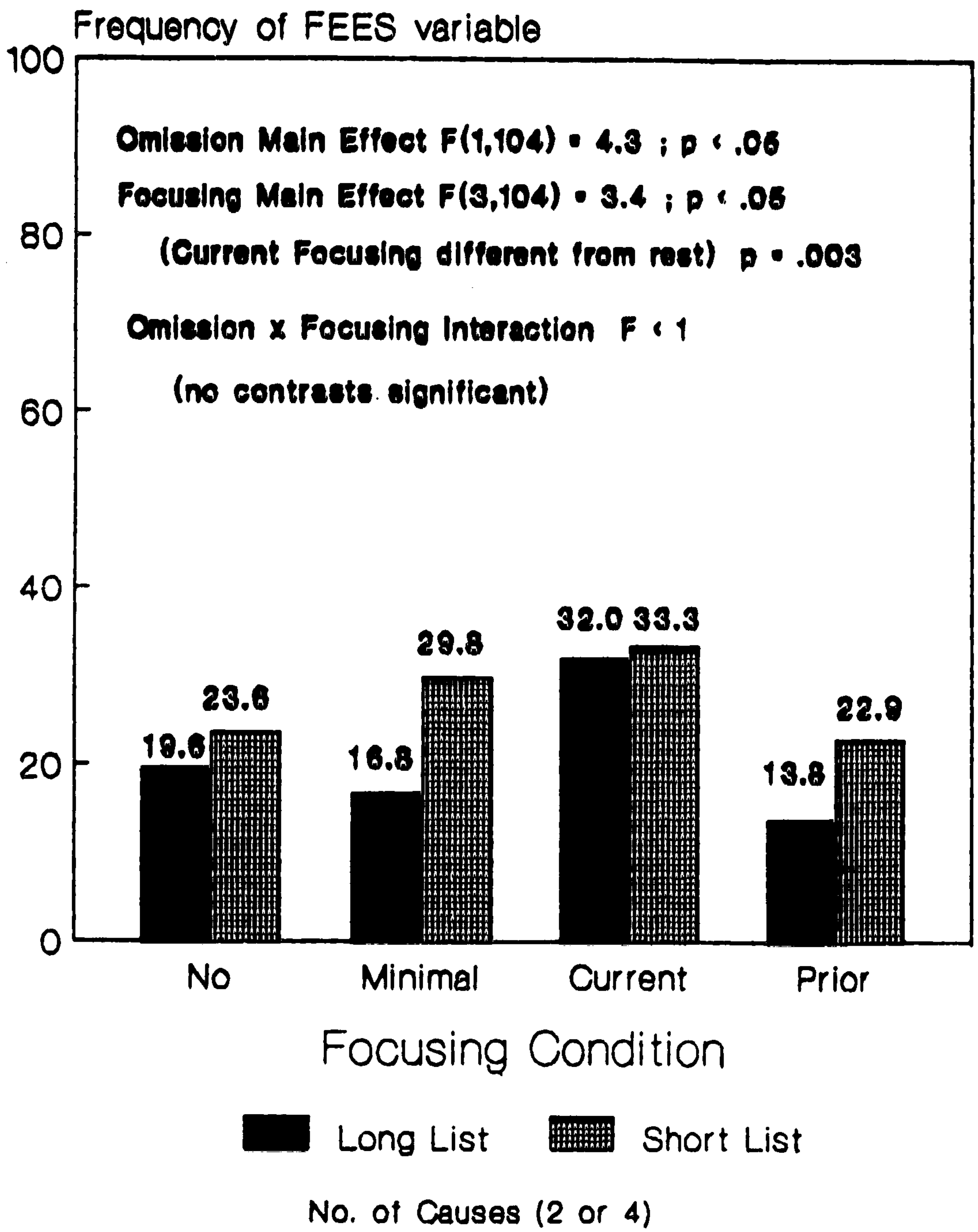
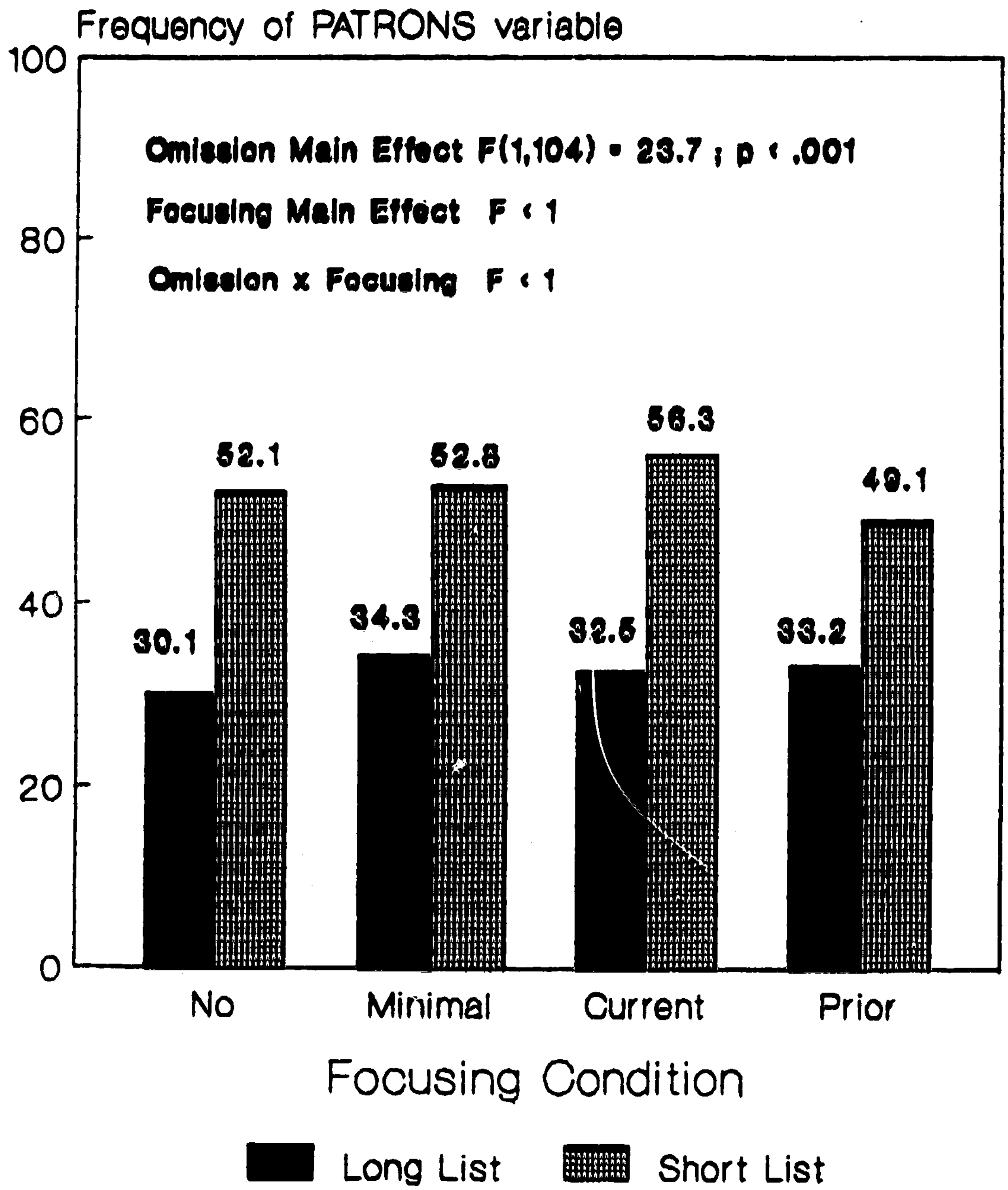
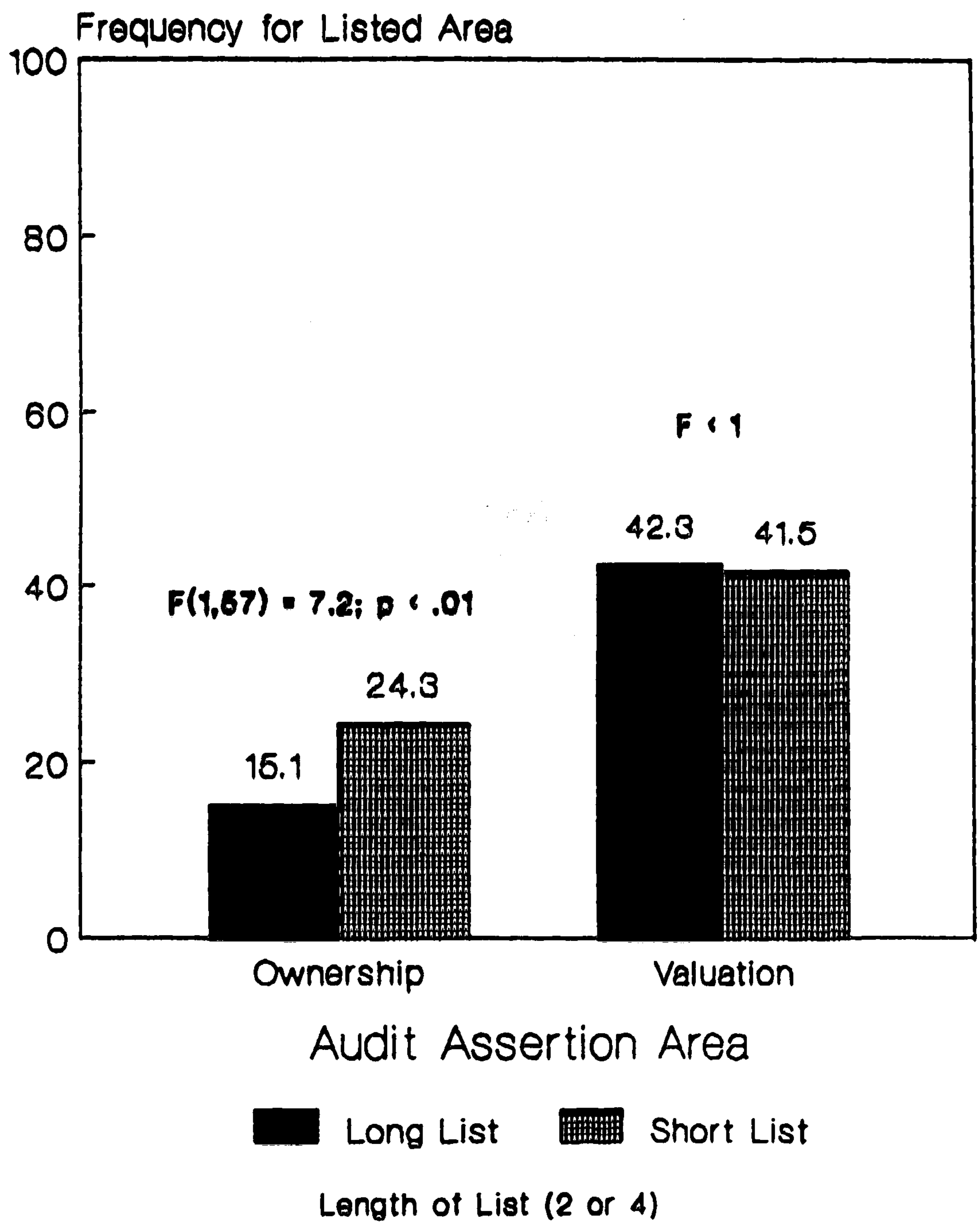


Figure 3-4
Experiment 2
Omission Effect by Focusing



No. of Causes (2 or 4)

Figure 3-5
Experiment 3
Omission Main Effects



REFERENCES

C.I.C.A. Handbook. Toronto, Canada: The Canadian Institute of Chartered Accountants.

Conover, W. J. (1980). Practical nonparametric statistics, 2nd ed. New York: John Wiley.

Dube-Rioux, L. & Russo, J. E. (1988). An availability bias in professional judgment. Journal of Behavioral Decision Making, 1, 223-237.

Einhorn, H. J. & Hogarth, R. M. (1986). Judging probable cause. Psychological Bulletin, 99, 3-19.

Einhorn, H. J. & Hogarth, R. M. (1987, October). Adaptation and inertia in belief updating: The contrast-inertia model. Unpublished working paper, University of Chicago.

Fischhoff, B., Slovic, P. & Lichtenstein, S. (1978). Fault trees: Sensitivity of estimated failure probabilities to problem representation. Journal of Experimental Psychology: Human Perception and Performance, 4, 330-344.

Hirt, E. R. & Castellan, N. J., Jr. (1988). Probability and category redefinition in the fault tree paradigm. Journal of Experimental Psychology: Human Perception and Performance, 14, 122-131.

Hoch, S. J. (1984). Availability and interference in predictive judgment. Journal of Experimental Psychology: Learning, Memory, and Cognition, 10, 649-662.

Johnson, R. A. & Wichern, D. W. (1982). Applied Multivariate Statistical Analysis. Englewood Cliffs NJ: Prentice-Hall.

Kahneman, D. & Miller, D. T. (1986). Norm theory: Comparing reality to its alternatives. Psychological Review, 93, 136-153.

Kahneman, D. & Tversky, A. (1982). The simulation heuristic. In D. Kahneman, P. Slovic and A. Tversky (eds.), Judgment under uncertainty: Heuristics and biases (pp. 201-208). New York:

Cambridge University Press.

Mehle, T. (1982). Hypothesis generation in an automobile malfunction interference task. Acta Psychologica, 52, 87-106.

Rennie, R. D. & Johnson, R.D. (1988, October). Auditors' judgments of probable causes: Effects of availability, experience, focusing and omission. Presentation at ORSA/TIMS, Denver.

Rundus, D. (1973). Negative effects of using list items as recall cues. Journal of Verbal Learning and Verbal Behavior, 12, 43-50.

Slamecka, N. J. (1968). An examination of trace storage in free recall. Journal of Experimental Psychology, 76, 504-513.

Slovic, P., Fischhoff, B. & Lichtenstein, S. (1982). Facts versus fears: Understanding perceived risk. In D. Kahneman, P. Slovic and A. Tversky, (eds.), Judgment under uncertainty: Heuristics and biases, (pp. 463-489). New York: Cambridge University Press.

Tulving, E. & Pearlstone, Z. (1966). Availability versus accessibility of information in memory for words. Journal of Verbal Learning and Verbal Behavior, 5, 381-391.

Tversky, A. & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. Cognitive Psychology, 4, 207-232.

Tversky, A. & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. Science, 185, 1124-1131.

Weber, M., Eisenfuhr, F. & Von Winterfeldt, D. (1988). The effects of splitting attributes on weights in multiattribute utility measurement. Management Science, 34, 431-445.

IV. GENERAL DISCUSSION

The results of these two papers demonstrate that both experienced and novice auditors underestimated (relative to a longer, more complete list) the probabilities of list items which were omitted from a list of possible causes for an auditing event. This phenomenon was called an omission effect.

Unlike previous studies, however, an expertise effect on the part of experienced auditors was found. That is, the omission effect was significantly lower for the experienced auditors than it was for the novice auditors. This finding of an expertise effect supports a view that expertise is domain specific and that expertise includes knowledge of relative frequencies of events. The failure to find an expertise effect in most prior studies may be due to the experimental tasks used. That is, unless an experimental task is derived from a domain in which subjects can be expected to use knowledge of frequency information, it is not surprising that expertise effects are not found. For example, in Fischhoff et al.'s (1978) study of reasons for a car not starting, expert mechanics were asked to assess the probability of each listed possible cause. Mechanics may not need to use relative frequency information. Decisions about which possible cause to investigate first may be dominated by cost considerations. In fact, many automobiles (e.g. Toyota) include a trouble shooting guide that appears to be based on amount of time required to investigate each possible cause. Alternatively, mechanics may follow a frequency-based guide without knowing the frequencies by which it was generated.

Dube-Rioux and Russo (1988) also examined professional hospitality industry managers' probability judgments of why a restaurant might fail. However, they provided no evidence to support their assumption that these experts should use and have access to the underlying frequencies.

The initial chapter of this thesis demonstrated how knowledge of the underlying frequencies of possible causes could aid in efficiently conducting an audit. The analytical review task in the first study was chosen as an area where expert auditors might reasonably be expected to have and use such expertise. The result was that the expert auditors had a lower omission effect than did the novices (at $p < .05$). This expertise is domain specific, however, as demonstrated by the failure to find any strong expertise effects in the first experiment of the second study. These findings together suggest that in order to reduce any omission effect, auditors should attempt to accumulate actual frequency information concerning important audit decisions.

The results of these studies are consistent with an anchoring-and-adjustment process for the assignment of probabilities to the possible causes. The use of this process should not be confused with the biases associated with the anchoring-and-adjustment heuristic as discussed by Tversky and Kahneman (1974). In situations where auditors have no access to frequency information from their experience, an anchoring-and-adjustment process is a reasonable strategy to follow even though the use of this strategy can lead to an omission effect. However, this omission effect cannot be viewed as a bias since it is simply the comparison of probabilities from long and short lists of possible causes and there is no available standard of comparison. Unless veridical or normative probabilities are available as a standard, one cannot say whether the probabilities in the long or in the short lists are less biased.

The conclusion that the omission effect is mediated by an anchoring-and-adjustment strategy is consistent with Einhorn and Hogarth's (1986) model for determination of probable cause. A similar strategy for the combination of audit evidence has been found in other studies involving auditors (e.g. Ashton and Ashton, 1988).

This dissertation used laboratory experiments to examine auditors' cognitive processes. While an experimental approach may be necessary to examine these processes (cf. Nisbett and Wilson, 1977), it does present problems of external validity. That is, while novice and experienced auditors displayed omission effects in the laboratory tasks, one cannot say from these results that similar judgments would prevail in actual audit settings. A laboratory setting can never attempt to simulate a real world setting. For example, auditors face a different incentive structure when making audit decisions and most of these decisions are subject to at least two levels of review. However, the finding of omission effects in the judgments of auditors in this study suggests that further experimental and field research is needed. For, it is not hard to imagine audit situations which may be similar to those presented in these experiments such as the second review of an audit file, the use of audit checklists which may be incomplete, and the actual determination of the probable cause for compliance deviations in internal control and for fluctuations found in analytical review. If such situations do exist and if there are omission effects in auditors' judgments, then the accumulation of actual frequency data for important auditing decisions might help to reduce any negative consequences of anchoring-and-adjustment by providing auditors with realistic anchors.

REFERENCES

Ashton, A. H. & Ashton, R. H. (1988). Sequential belief revision in auditing. The Accounting Review, 63, 623-641.

Dube-Rioux, L. & Russo, J. E. (1988). An availability bias in professional judgment. Journal of Behavioral Decision Making, 1, 223-237.

Einhorn, H. J. & Hogarth, R. M. (1986). Judging probable cause. Psychological Bulletin, 99, 3-19.

Fischhoff, B., Slovic, P. & Lichtenstein, S. (1978). Fault trees: Sensitivity of estimated failure probabilities to problem representation. Journal of Experimental Psychology: Human Perception and Performance, 4, 330-344.

Nisbett, R. E. & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. Psychological Review, 84, 231-259.

Tversky, A. & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. Science, 185, 1124-1131.

APPENDIX

Overview of Experiments, Statistical Method and Experimental Design

The experiments in this dissertation examined the phenomenon that when subjects are asked to assess the probabilities of a partial list of possible causes for an event (plus the probability of all other causes), they underestimate the probabilities of all other causes relative to other subjects who assess a more complete list. This phenomenon is referred to as the omission effect and it is measured between subjects (to avoid experimental demands of a within-subjects design). Another way of looking at the omission effect is that subjects in the partial list overestimate the probabilities of the individual causes, relative to subjects assessing those same causes in the more complete list. Thus the omission effect was assessed by comparing the mean assessed probabilities for the listed causes for each of the list length conditions.

Beyond demonstrating that the omission effect exists using auditors as subjects, this dissertation was primarily concerned with factors which may mediate the omission effect. Thus factors such as subject expertise and prior recallability of the list items were varied factorially by list length condition. It is therefore the interaction of these possible mediating variables and the list length condition which was of interest in assessing their impact on the omission effect.

Since the concern of the study was whether the means of the groups by preassigned condition differ, analysis of variance (ANOVA) was used to test the significance of the differences in mean responses for the individual causes. A fixed effects model was assumed since no attempt is being made to generalize these results beyond the levels of the variables tested in the experiments. In those situations (in the second paper) where the probabilities of more than one individual cause were examined, multivariate analysis of variance (MANOVA) was also used (on the matrix of responses for the common causes by subject by condition) to test the hypothesis of equality of mean vectors. MANOVA was used since the responses to the individual causes by subject may be negatively correlated (since the total of all probabilities listed for each subject must sum to one). The significance of the correlation among the dependent variables was assessed using Bartlett's chi-square statistic which tests whether or not the correlation matrix is an identity matrix. Where Bartlett's statistic indicated that there was a significant correlation between at least two dependent variables, a MANOVA was performed on the vector of probability responses by subject by condition. (Where there was no significant correlation only univariate ANOVAs were reported). The MANOVA results were examined first and where the hypothesis of equality of mean vectors was rejected, a hypothesis contrasting the dependent variables was done. Then univariate ANOVAs on the independent variables were used to determine which means were different.

Since multiple responses were generated by each subject (i.e. probabilities for each possible cause), the analysis could be viewed as a repeated measures design. However, such an analysis may not be meaningful since it is the differences on the individual presented causes (and not an "average" cause) between list length conditions (i.e. the omission effect) which is of primary concern in the study.

ANOVA (and MANOVA) assumes that the variances (and the covariance matrices) of the experimental conditions are homogeneous. Where cell sizes are equal or proportional (or approximately so) this assumption is not a major concern. In this dissertation these assumptions were examined by Bartlett's test for homogeneity of variance (for the univariate ANOVAs) and by Box's M test for the MANOVAs. Where Bartlett's test rejected the homogeneity of variance assumption, the variables were transformed to achieve homogeneity. The significance of the effects was not altered in the transformed data. Thus, since the cell sizes are similar in size, any lack of homogeneity of variances (and of covariance matrices) is not likely to be a problem.

ANOVA and MANOVA assume that the underlying distribution of the dependent variables of each experimental group is normally distributed. Because the cell size for each group is quite small it would be difficult to detect departures from normality. Thus the normality assumption was assessed by looking at the residuals of the ANOVA and MANOVA models over the entire sample rather than on a group basis. Normality for the univariate models was tested by Lilliefor's test. The hypothesis of normality of the residuals was rejected for one of the models even for the transformed data. Therefore to test the robustness of the effects found over different distributional assumptions, this ANOVA was rerun using a rank transformation of the dependent variable. The significance of the effects was unchanged. For the bivariate models the reasonableness of the bivariate-normal distribution was assessed by relying on the Lilliefor's test results for the univariate residuals and by looking at the chi-square plots of the ordered squared distances of the residuals. The results of these plots indicated that the bivariate normal distribution was a reasonable assumption in each case.

The next part of this appendix reviews the models tested by each of the experiments.

Paper 1

The experiment in the first paper (chapter 2), examined whether the availability (defined in terms of prior recall) was a mediating factor in the omission effect. Thus subjects in the short list conditions examined either low or high recall lists of possible causes while subjects in the long list conditions assessed the probabilities of both high and low recall possible causes. Since subjects in the short list conditions did not (necessarily) respond to the same causes, their responses on individual possible causes could not be directly compared. Thus a dependent measure ("ADJUSTED All OTHER CAUSES" or ADJOTHER) was computed which measures the relative frequency for other causes for the condition of interest (see discussion in text). The dependent measure was the frequency for "all other causes" for the short lists. For the long lists, the corresponding dependent measure was the frequency for all other causes plus the frequency for either the low or high availability causes. Thus in comparing the the dependent measures between long and short list conditions, the dependent measure reflected the assessed frequency associated with those list items which were different between the long and short lists.

The independent (i.e. explanatory) variables in the 2 x 2 x 2 design were expertise (novices or partner/managers), omission (long or short list length of 3 or 6 possible causes listed) and availability (high or low recall). A summary of the variables and cell sizes is shown below:

N=121

EXPERTISE:	NOVICES n=54				PARTNER/MANAGERS n=67			
OMISSION (List Length):	LONG n=28		SHORT n=26		LONG n=31		SHORT n=36	
AVAILABILITY (Prior Recall):	HIGH	LOW	HIGH	LOW	HIGH	LOW	HIGH	LOW
Cell Sizes:	14	14	13	13	15	16	18	18

The ANOVA model which was of experimental interest was:

```
ADJOTHER
 = Bo + B1(EXPERTISE) + B2(OMISSION) + B3(AVAILABILITY)
 + B4(EXPERTISE*OMISSION)
 + B6(EXPERTISE*AVAILABILITY)
 + B7(OMISSION*AVAILABILITY)
 + B9(EXPERTISE*OMISSION*AVAILABILITY)
```

The effects which were of interest are the OMISSION main effect and any of the interactions involving OMISSION. The interaction term for EXPERTISE and OMISSION measures the tendency for the OMISSION effect to vary over the two levels of EXPERTISE. Therefore examining the significance of this effect and the related means provides a test of the hypothesis that experts should have a lower omission effect than novices. Similarly the interaction term for OMISSION and AVAILABILITY allows the OMISSION effect to vary over the two levels of AVAILABILITY. The significance of this interaction and an examination of the related marginal means provides a test of the availability hypothesis that the omission effect should be lower for the low availability list of possible causes.

In the administration of the experiments, the order of this experiment and one other were counterbalanced, with one-half of the subjects doing this experiment first and one-half doing it second. The variable COUNTERBALANCE (Table 2-1) was used to reflect this difference in order. Since all other effects were factorially crossed with the COUNTERBALANCE variable, this order effect was incorporated into the model by adding another independent variable (COUNTERBALANCE) main effect and all two-way, three-way and four-way interactions of COUNTERBALANCE with the other independent variables. None of these order effects or interactions were significant.

Paper 2 -- Experiment 1

The first experiment of the second paper examined whether subjects' probability assessments were affected by EXPERTISE (novices or partner/managers), list length or ANCHOR (subjects were told they would sequentially assess a long or a short list of length three or six), and the ORDER in which two of the possible causes were presented (first or second). The design was therefore a 2 x 2 x 2 factorial. The cell sizes were as follows:

N=117								
EXPERTISE:	NOVICES n=52				PARTNER/MANAGERS n=65			
ANCHOR (List Length):	LONG n=27		SHORT n=25		LONG n=31		SHORT n=34	
ORDER	1st	2nd	1st	2nd	1st	2nd	1st	2nd
Cell Sizes:	14	13	12	13	16	15	17	17

The dependent variables were the three causes which were common across both list length conditions: ADVANCES, CLERICAL ERROR and FRAUD. Since Bartlett's chi-square statistic indicated no significant correlations among the dependent variables, separate univariate ANOVAs were reported for each dependent variable, testing the following models:

Dependent variable-- each of:
ADVANCES, CLERICAL ERROR, FRAUD =

Bo + B1(EXPERTISE) + B2(ANCHOR) + B3(ORDER) +
B4(EXPERTISE*ANCHOR) + B5(EXPERTISE*ORDER) +
B6(ANCHOR*ORDER) + B7(EXPERTISE*ANCHOR*ORDER)

The main effect for ANCHOR was of experimental interest because an anchoring-and-adjustment explanation for the omission effect would predict that when subjects are given an anchor based on a longer list length, their probability assignments should be lower. The main effect for the ORDER variable was also of interest because having a possible cause presented second rather than first would mean that at least one other cause had already been cued and was therefore available. Thus if the order of presentation of the possible causes affected the magnitude of subjects' estimated probabilities, this would support the availability explanation for the omission effect. Expertise effects were also of interest in this experiment. Since this experiment was counterbalanced with the experiment in the first paper, a similar test of the order effects was done. None of the order effects or interactions of order with the other experimental effects were significant.

Paper 2 -- Experiment 2

The second experiment of the second paper examined various focusing techniques as a means of reducing the omission effect. The design of the experiment was a 2 x 2 x 4 between-subjects factorial of EXPERTISE (novices or partner/managers), OMISSION (list length -- long or short --4 or 2 causes), and FOCUSING at four levels (no focusing, minimal focusing, concurrent focusing and prior focusing). The cell sizes (N = 120) were as follows:

Novices: (n=50)

Long List (n=25)

Focusing:	None	Minimal	Current	Prior
cell sizes:	6	7	5	7

Short List (n=25)

Focusing:	None	Minimal	Current	Prior
cell sizes:	5	6	6	8

Partner/Managers: (n=70)

Long List (n = 35)

Focusing:	None	Minimal	Current	Prior
cell sizes:	12	8	5	10

Short List (n = 35)

Focusing:	None	Minimal	Current	Prior
cell sizes:	9	10	7	9

The multivariate model which was tested was:

[FEES , PATRONS] =
Bo + B1(EXPERTISE) + B2(OMISSION) + B3(FOCUSING) +
B4(EXPERTISE*OMISSION) + B5(EXPERTISE*FOCUSING) +
B6(OMISSION*FOCUSING) + B7(EXPERTISE*OMISSION*FOCUSING)

The effects which were of experimental interest were the OMISSION main effect and the FOCUSING by OMISSION interaction. Any expertise effects were also of interest. Since the FOCUSING variable was at 4 levels, three orthogonal constrasts were performed on the main effect. Planned comparisons of interest were done on the interactions to assess the effect of the various levels of focusing on the omission effect. Since the MANOVA rejected equality of the mean vectors, univariate ANOVAs were done on the individual dependent variables to assess the differences.

This experiment was positioned as the last of five experiments which were run in an experimental session. The first two experiments were discussed above. The other two experiments concerned auditors judgments but were unrelated to the dissertation and so they served as a distraction for this experiment, from the first two experiments.

Paper 2 -- Experiment 3

The third experiment of the second paper further explored the effects of focusing on the omission effect. It was run at a different time with a different group of subjects. The independent variables in this 2 x 2 factorial design were: OMISSION (list length of 2 or 4) and FOCUSING (prior or post). Expertise was not examined in this experiment. The cell sizes were as follows:

N = 61

OMISSION (list length):	Four n=32		Two n=29	
FOCUSING:	Prior	Post	Prior	Post
cell sizes:	17	15	14	15

The multivariate model which was examined in this experiment was:

[VALUATION , OWNERSHIP] =
Bo + B1(OMISSION) + B2(FOCUSING) +
B3(OMISSION*FOCUSING)

The effects which were of experimental interest were the OMISSION main effect and the OMISSION by FOCUSING interaction. Since the MANOVA rejected equality of the mean vectors, univariate ANOVAs were done on the individual dependent variables to assess the differences.