

1-1-1997

An examination of methods of improving audit judgment

Janne Chung
Edith Cowan University

Follow this and additional works at: <https://ro.ecu.edu.au/theses>



Part of the [Accounting Commons](#)

Recommended Citation

Chung, J. (1997). *An examination of methods of improving audit judgment*. <https://ro.ecu.edu.au/theses/887>

This Thesis is posted at Research Online.
<https://ro.ecu.edu.au/theses/887>

Edith Cowan University

Copyright Warning

You may print or download ONE copy of this document for the purpose of your own research or study.

The University does not authorize you to copy, communicate or otherwise make available electronically to any other person any copyright material contained on this site.

You are reminded of the following:

- Copyright owners are entitled to take legal action against persons who infringe their copyright.
- A reproduction of material that is protected by copyright may be a copyright infringement. Where the reproduction of such material is done without attribution of authorship, with false attribution of authorship or the authorship is treated in a derogatory manner, this may be a breach of the author's moral rights contained in Part IX of the Copyright Act 1968 (Cth).
- Courts have the power to impose a wide range of civil and criminal sanctions for infringement of copyright, infringement of moral rights and other offences under the Copyright Act 1968 (Cth). Higher penalties may apply, and higher damages may be awarded, for offences and infringements involving the conversion of material into digital or electronic form.

AN EXAMINATION OF METHODS
OF
IMPROVING AUDIT JUDGMENT

J. O. Y. Chung
Ph. D.
1998

USE OF THESIS

The Use of Thesis statement is not included in this version of the thesis.

**AN EXAMINATION OF METHODS OF IMPROVING AUDIT
JUDGMENT**

by

Janne Oi Yin Chung

A Thesis Submitted in Fulfilment of the Requirements for the Award of

Doctor of Philosophy

At the Faculty of Business,
Edith Cowan University

Date of submission: December 8, 1997

AN EXAMINATION OF METHODS OF IMPROVING AUDIT JUDGMENT

Abstract

This study examined three methods of improving audit judgment, namely inoculation, the group process, and counterexplanation. Prior research found that auditors' judgments were not always optimum. An important question this study asked was whether the use of inoculation, the group process and counterexplanation leads to more effective judgments.

In Experiment One, I proposed that inoculation will reduce the effect of framing so that participants exposed to the inoculation treatment will not display a framing effect, while participants not exposed to the treatment will. In addition, I hypothesised that the relation between audit experience and responsiveness to training has an inverted-U shaped value function. The results showed this to be the case. No interaction effect between framing and inoculation was observed, but the order of writing supporting and opposing arguments as required by the inoculation treatment led to a primacy effect.

The second experiment posited that the use of audit groups will improve individual auditors' judgments in a going-concern evaluation task. The results showed that not only were audit group judgments more conservative, but consensus was also higher among audit group judgments. To explain the hypothesised differences between audit group and individual auditor judgments, I examined groups' and individuals' self-reported cue usage, cue selection and cue weighting. The results showed that both groups and individuals were equally proficient in cue weighting, but the groups selected more

going-concern problem indicators as well as mitigating factors compared to the individuals.

Experiment Three contained two main parts. In part one, groups and individuals performed an audit of an inventory balance. The results showed that the group process results in more accurate judgments among auditors. The results also showed that groups were more confident in addition to being better calibrated. In the second part of Experiment Three, I conjectured that there are positive as well as negative aspects of counterexplanation, and that the group process may be an effective way of enhancing the positive aspects and reducing the negative aspects of counterexplanation. To test this, I required participants to counterexplain their initial judgment prior to making a final judgment. The results showed that counterexplaining inaccurate judgments led to more accurate judgments, whereas counterexplaining accurate judgments led to less accurate judgments. I also found inconclusive evidence that the group process may enhance the positive aspects and mitigate the negative aspects of counterexplanation.

DECLARATION

"I certify that this thesis does not incorporate, without acknowledgement, any material previously submitted for a degree or diploma in any institution of higher education and that, to the best of my knowledge and belief, it does not contain any material previously published or written by another person except where due reference is made in the text".

Janne Oi Yin Chung

Date: December 8, 1997

ACKNOWLEDGEMENTS

*"Unless the Lord builds the house, those who build it labour in vain
Psalms 127:1a"*

I will forever be indebted to all of my family for the loving support and encouragement they have provided in helping me attain this goal. My sister, Anne, provided me with spiritual and emotional support and I dedicate this thesis to her. My parents, Teck Min and Syn Soon, were a constant source of inspiration and motivation to me. I also thank my other siblings, nieces and nephews whose unfailing love sustained me during this project.

I wish to thank Gary Monroe, my principal supervisor, for his continuous help and encouragement throughout this project. In addition, I express my sincere appreciation to Geoff Soutar, my associate supervisor, who was never too busy to deal with my queries, and to my colleagues, Vincent Chong and Theo Christopher, for their friendship. Thank you also to Roger Simnett and Peter Schelluch for their assistance with data collection, Karen Pincus for the use of her case materials, and the School of Accounting, Edith Cowan University and Small and Medium Enterprise Research Centre, Edith Cowan University for financial assistance.

Table of Contents

Chapter 1 - Introduction.....	1
Chapter 2 - Methods for Improving Audit Judgment	7
Introduction	7
Inoculation Theory.....	7
Group Process.....	12
Counterexplanation	35
Chapter 3 - Experiment One: The Effects of Inoculation, Framing and Audit Experience on Internal Control Evaluation.....	44
Introduction	44
Internal Control Evaluation	47
Framing Effect.....	48
Order effects	52
Theory Development	53
The Experiment	59
Results and Discussion	64
Conclusions	74
Limitations.....	76
Chapter 4 - Experiment Two: An Investigation of Information Processing by Audit Groups	77
Introduction	77
Going-Concern Assessment	80
Theory Development	81
The Experiment	89
Results	94
Discussion and Conclusions	102
Limitations.....	104
Chapter 5 - Experiment Three: Counterexplanation and Audit Groups	105
Introduction	105
Explanation Effect.....	107
Hypotheses Development	127
The Experiment	137
Results and Discussion	146
Conclusions	171
Limitations.....	175
Chapter 6 - Conclusions.....	176
Overview	176
Limitations.....	179
Future Research	180
References	183

List of Tables

Table 3.1 - The Effects of Inoculation, Framing and Audit Experience on Task Difficulty.....	66
Table 3.2 - The Effects of Inoculation and Framing on the Evaluation of Internal Controls - Mean (sd)	68
Table 3.3 - ANOVA Results: The Effects of Inoculation, Framing and Audit Experience on the Evaluation of Internal Controls	69
Table 3.4 Effects of Inoculation and Audit Experience - Mean (SD) Evaluation of Internal Controls	71
Table 3.5 - Order Effects	74
Table 4.1 - Descriptive Statistics.....	89
Table 4.2 - Going-Concern Judgments of Groups and Individuals	96
Table 4.3 - The Effects of Groups/Individuals and Audit Experience on Going-Concern Judgments	96
Table 4.4 - The Effect of Groups/Individuals on the Mean Number of Information Items Used (sd).....	98
Table 4.5 - The Effects of Groups/Individuals and Audit Experience on the Mean Number of Cues Used	98
Table 4.6 - Cue Selection.....	99
Table 4.7 - The Effects of Group/Individuals and Audit Experience on Cue Selection.....	100
Table 4.8 - The Effect of Groups/Individuals on Cue Weighting.....	101
Table 4.9 - The Effects of Groups/Individuals and Audit Experience on Cue Weighting .	102
Table 5.1 - Experimental Design	140
Table 5.2 The Effect of Groups/Individuals on Accuracy	149
Table 5.3 Groups/Individuals Dichotomous Accuracy Judgment	149
Table 5.4 The Effect of Groups/Individuals on Confidence.....	151
Table 5.5 - The Correlation Between Accuracy and Confidence Using the Continuous Accuracy Scale	152

Table 5.6 - The Relation Between Accuracy and Confidence Using the Dichotomous Scale	152
Table 5.7 - The Effects of Explanation Mode and Groups/Individuals on Belief Revision.....	154
Table 5.8 - ANCOVA Analyses of Belief Revision	155
Table 5.9 - The Effects of Explanation Mode and Groups/Individuals on Inaccurate Judgments.....	156
Table 5.10 - The Effects of Explanation Mode, Groups/Individuals on Inaccurate Judgments.....	157
Table 5.11 - The Effects of Explanation Mode and Groups/Individuals on Accurate Judgments.....	158
Table 5.12 - The Effects of Explanation Mode, Groups/Individuals and Audit Experience on Accurate Judgments.....	161
Table 5.13 - Belief Revision Using Dichotomous Scale	162
Table 5.14 - The Effects of Explanation Mode and Groups/Individuals on Confidence Revision.....	164
Table 5.15 - The Effects of Explanation Mode, Groups/Individuals and Audit Experience on Confidence Revision	167
Table 5.16 - The Effects of Explanation Mode, and Groups/Individuals on Difficulty of Explaining/Counterexplaining.....	168
Table 5.17 - The Effects of Explanation Mode and Groups/Individuals on the Number of Explanations and Counterexplanations Generated.....	169
Table 5.18 - Difficulty of Explaining/Counterexplaining and Number of Explanations or Counterexplanations Generated	171

List of Figures

Figure 3.1 - Experimental Procedures.....	61
Figure 3.2 - The Effects of Audit Experience and Inoculation on the Evaluation of Internal Controls.....	70
Figure 4.1 - Experimental Procedures.....	92
Figure 5.1 - Experimental Procedures.....	142
Figure 5.2 - Interaction Effect Between Explanation Mode and Groups/Individuals on Accuracy.....	160
Figure 5.3 - Interaction Effect Between Explanation Mode and Groups/Individuals on Confidence Revision	166

List of Appendices

Experiment One.....	194
Experiment Two.....	202
Experiment Three.....	209

CHAPTER 1 INTRODUCTION

Auditing studies in the 1970's and early 1980's typically examined audit judgments in an uncertain environment. These studies reveal various fallacies in audit judgments (see Solomon and Shields, 1995). According to Solomon and Shields, later studies attempt to explain observed auditor behaviour using the heuristics and biases approach of Tversky and Kahneman (1974) and Kahneman et al. (1982). Consequently, numerous studies report that audit judgments are subject to various heuristics and biases (refer Smith and Kida, 1993). For example, auditors have a confirmatory bias (Church, 1991), their judgments are affected by the curse of knowledge¹ (Kennedy, 1995), the most recently obtained information cues are weighted more heavily than the earlier-obtained information cues² (Asare, 1992), and their judgments are also affected by the explanation effect³ (Koonce, 1992). While identifying and explaining auditor behaviour is vital and necessary, the next phase in the development of audit judgment research should be aimed at discovering means of correcting sub-optimal audit judgments. There is evidence in the literature of the beginnings of such a trend, for example, Heiman (1990), Koonce (1992) and Kennedy (1993; 1995) examine how audit judgments may be improved.

Judgment improving techniques include accountability⁴, inoculation⁵, and counterexplanation⁶, and has been found to lead to higher judgment consistency (Libby

¹ The curse of knowledge is the inability of people to ignore their own knowledge when making predictions about other people's knowledge (Kennedy, 1995).

² Prior psychology literature also finds a primacy effect (refer Hogarth and Einhorn 1992).

³ When the event explained is judged more likely to occur than the unexplained event, an explanation effect occurs.

⁴ Accountability refers to the need to justify one's judgment to others.

and Libby, 1989). These techniques work by requiring decision-makers to consider all appropriate information cues, combine these cues more effectively, and thereby, reduce cognitive strain (Jiambalvo and Waller, 1984). It is often assumed that the effect of judgment improving techniques is always positive. However, these techniques sometimes lead to ineffective judgments (Jiambalvo and Waller, 1984). This study examines three judgment improving techniques - namely, inoculation, group process and counterexplanation - and their effect(s) on audit judgments.

One technique suggested in the psychology literature is based on inoculation theory (Anderson, 1982). The inoculation treatment requires decision-makers to consider all possible solutions to a problem before being shown any information about the problem. This should lead to more effective judgments because the decision-maker is open to the existence of opposing arguments (Anderson, 1982). No prior auditing study has considered this technique, consequently we know little of its' effectiveness. The examination of this technique would add to the limited pool of tools available to the auditing profession. In this thesis, I propose that inoculation will reduce the effect of framing so that participants exposed to the inoculation treatment will not display a framing effect, whereas participants not exposed to the treatment will. In addition, I hypothesise that the relation between audit experience and responsiveness to training has an inverted-U shaped value function. The testing of this hypothesis would identify the auditors most responsive to training as there is surprisingly little literature on the relation between audit experience and responsiveness to this. Consequently, this study

⁵ Inoculation theory requires the decision-maker to consider both (or all) alternatives to a problem before being shown any information on the problem.

⁶ Decision-makers are required to consider why the chosen alternative may be incorrect and/or why the alternative rejected may be correct.

examines the effect of inoculation on the framing of internal control evaluations as well as the effect that audit experience has on an auditor's responsiveness to inoculation. Prior studies suggest that in making judgments, people generate causal explanations as a bridge between the information and the problem (Tversky and Kahneman, 1973). These causal explanations once generated cannot be readily changed, and because of their inflexibility, the quality of the resultant judgment is reduced (Hoch, 1984). This shortcoming is further exacerbated because people underestimate how easily causal explanations are generated (Hoch, 1984). Therefore, in the search for ways to improve audit judgments, consideration should be made of whether the technique being considered would regulate the premature formation of causal explanations. Consequently, inoculation is chosen in this investigation of alternative judgment improving tools.

Auditors work as members of an audit group as well as on their own. The use of audit groups/teams is based on the assumption that these groups would improve the quality of the judgment. The increasing use of both audit groups to solve complex problems may be due to the need of public accounting firms to reduce potential litigation, the increased complexities of the reporting requirements and the general increase in the use of specialists in auditing (Schultz and Reckers, 1981). The group process involve decision-makers undertaking a two-part process to arrive at a judgment - a private revision phase and a public weighting phase (Sniezek and Henry 1990; 1989). During the revision phase, the decision-maker revises his/her personal judgment in light of the views presented by other group members, and this may include decomposing the problem. During the weighting phase, group members weigh up the various global judgments

presented by group members and arrive at a judgment. Groups are used in decision-making based on the assumption that, compared to individual decision-makers, they are more effective in making complex judgments as well as making judgments in an uncertain environment (Fisher and Ellis, 1990).

An audit group generally comprises two or more auditors. Audit groups may be interacting or non-interacting. Non-interacting groups are often called "composites" or staticised groups. Two types of interacting groups are found in auditing (Solomon, 1987). According to Solomon, an audit team is hierarchical in nature and the judgment is often made by the most senior team member. Audit groups, on the other hand, comprise auditors of equal rank within the firm's hierarchy and make judgments jointly (Solomon, 1987). This study examines the latter, i.e., interacting audit groups as well as individual auditors in a going concern evaluation task and an error detection task. The purpose of the examination is to discover whether the use of audit groups leads to more effective judgments compared to individual auditors. Consequently, questions that this study will attempt to answer include: Are audit group judgments significantly different from individual auditor judgments, and is consensus higher for the former than for the latter? Are audit groups more accurate than individual auditors, and are audit groups more confident and better calibrated than individual auditors?

The social psychology literature suggests that decision-makers judge an event more likely to occur after they have provided an explanation for its occurrence. This finding in the social psychology literature is rather robust and has been observed in a wide range of decision-makers including auditors (Koonce, 1992). The auditing standards require auditors to explain their judgment by way of providing and documenting audit evidence

to support a judgment. Such a requirement results in an explanation effect. This study considers whether the provision of a counterexplanation would result in more effective judgments by individual and groups of auditors. In counterexplaining, auditors are required to consider why the chosen judgment alternative is incorrect and/or why the rejected alternative is correct⁷. Prior auditing literature suggests that counterexplanation may result in more effective judgments. Counterexplanation results in higher belief revision (Heiman, 1990), is effective in correcting the curse of knowledge (Kennedy, 1995), and reduces the explanation effect (Koonce, 1992). There is evidence in the literature of the increasing use of counterexplanation in practice. For example, Libby and Trotman (1993) find that audit reviewers pay greater attention to evidence inconsistent with the conclusions reached by their audit subordinates in analytical procedures. Before this practice of requiring a counterexplanation becomes more widespread, I see a need to investigate whether there are any negative effects of counterexplanation. Consequently, this study examines possible negative effects of counterexplanation by asking: Will counterexplaining accurate judgments result in less accurate judgments? I conjecture that there are positive as well as negative aspects of counterexplanation, and that the group process may be an effective way of reducing these negative aspects. This study will also examine whether the use of audit groups would mitigate the negative aspects of counterexplaining.

The results contribute to the existing literature by identifying the effects of three methods of improving audit judgments. This adds to the limited pool of techniques

⁷ Counterexplanation is different from inoculation in that the former requires the decision-maker to consider why his/her judgment may be incorrect after the evidence had been viewed and the judgment made, while inoculation requires decision-makers to consider all alternatives to a problem before viewing any information.

available to public accounting firm management, trainers and educators and assists in the education and training of auditors by identifying the conditions under which these techniques are most effective, as well as the conditions under which their use may be counter-productive.

The remainder of this study is organised as follows. Chapter 2 reviews prior literature on inoculation, the group process and counterexplanation and discusses the related theories. Three experiments are carried out, and these are detailed in Chapters 3, 4, and 5. Chapter 3 reports the effect of inoculation on framing as well as the effect of audit experience on inoculation in an internal control evaluation task. Chapter 4 reports the effects of audit groups and individual auditors on going concern evaluation. The experiment reported in chapter 4 tests whether audit group and individual auditor judgments are different, and examines groups' and individuals' information processing as a means of explaining the difference(s). These include an examination of self-reported cue usage, cue selection and cue weighting. In Chapter 5, the group process and counterexplanation are used in a fraud detection task to examine whether group process and counterexplanation would increase fraud detection rates. The robustness of counterexplanation is also tested by requiring participants who make accurate and inaccurate judgments to counterexplain their judgments. A summary and discussion of the limitations of this study as well as areas for future research conclude this thesis in Chapter 6.

CHAPTER 2

TECHNIQUES FOR IMPROVING AUDIT JUDGMENT

INTRODUCTION

Chapter Two discusses various techniques used in both psychology and accounting to improve judgment. It begins with a description of prior literature and theories on inoculation. The next section discusses the theories and prior literature of the group process. The chapter concludes with an examination of prior literature on counterexplanation and discusses the related theories.

INOCULATION THEORY

Inoculation theory is borrowed from medical science where it is used to explain the ability of people to ward off infectious diseases after a mild dose of these diseases is introduced to their bodies. Inoculation theory was introduced into social science as a means of immunising people against persuasion (see McGuire and Papageorgis, 1962; McGuire, 1961; McGuire and Papageorgis, 1961). The aim of inoculation is to immunise people's beliefs against persuasion by pre-exposing them to "weakened, defence stimulating forms of the counterarguments" (McGuire and Papageorgis, 1961, 327). The purpose of this pre-exposure is to strengthen the cognitions of people by providing them with experience in refuting counterarguments and to sensitise them to the existence of counterarguments (McGuire and Papageorgis, 1961).

Another area where inoculation is used is in marketing communications. For example, in the early 70's, the Standard Oil Company was found guilty of deceptive advertising

and forced to correct the deception. Pre-empting negative publicity, Standard Oil, in its subsequent advertising, addressed the issue of the deceptive advertising as well as the positive claims of the company's products (Hunt, 1973). Such advertising had a positive effect on the image of Standard Oil instead of the negative effect anticipated by the company (Hunt, 1973). The reasons for the success of inoculation in marketing studies are based on two aspects of this theory. First, studies by McGuire and Papageorgis found that it is possible to immunise people's beliefs from subsequent counterarguments (attacks) and that a defence which considers both supporting and opposing arguments is more effective than a one-sided argument (McGuire and Papageorgis, 1962; McGuire, 1961; McGuire and Papageorgis, 1961). Second, marketing studies also found that two-sided arguments are more effective in creating resistance to change in the face of subsequent attacks (Crowley and Hoyer, 1994; Kamins and Assael, 1987). Inoculation has also been used in social intervention to inoculate job seekers against setbacks (Caplan et al., 1989), and in decision-making to correct belief-perseverance (Anderson, 1982).

Some studies suggest that causal explanations are generated automatically as a by-product of decision-making (e.g., Hoch, 1984; Koriat et al., 1980), so preventing their formation is not always possible (Anderson, 1982). Causal explanations cause the belief perseverance observed in social psychology and accounting studies (Tutin, 1993; Koonce, 1992; Ahlawat, 1992). One method of correcting belief-perseverance⁸ caused by causal explanations is to regulate the formation of such causal explanations

⁸ Belief perseverance is the tendency of people to hold firm to illogical beliefs that are resilient to "logical and empirical challenges". Such beliefs may become isolated from the data that gave rise to them in the first place (Anderson et al., 1980).

(Anderson, 1982). Anderson et al. (1980) provided their participants with case studies suggestive of either a positive or negative relation between risk-taking and success as a firefighter. Half the participants were required to provide a written explanation of a positive relation between these two factors while the other half explained a negative relation between risk-taking and success as a firefighter. All participants were then thoroughly debriefed of the fictitious nature of that relation. A post-test questionnaire provided a second set of measures on participants' perception of the relation between risk-taking and success as a firefighter. The results show a belief perseverance that survived the debriefing.

Anderson (1982) was the first to introduce inoculation into decision-making as a tool for correcting belief perseverance. Following the arguments used by McGuire and Papageorgis (1961), the aim of Anderson's (1982) inoculation treatment was to sensitise participants to the existence of opposing arguments. In Anderson's experiment, all participants were required to consider the relation (whether positive or negative) between success as a firefighter and risk-taking propensity. As part of the inoculation treatment (inoculation condition), participants were required to explain "both (or all) possible relations between the relevant variables prior to examining any data at all". A second group (explanation condition) explained the relation after examining the case materials. The third group (dual-explanation condition) explained the relation after examining the case materials and counterexplained after the debriefing, while a fourth group did not explain or counterexplain (control condition). All participants were then debriefed of the fictitious nature of this relation before the dependent measures were collected. Participants in the explanation and control conditions displayed more belief

perseverance than participants in the other two conditions. Anderson's results show that the inoculation treatment is effective in making people more cautious about forming causal explanations and leads to a reduction in belief perseverance.

In sum, inoculation has been used to immunise people against persuasion by providing them with experience in refuting counterarguments. This has been observed in both social psychology and marketing communications. Two-sided arguments are more effective than one-sided arguments. Inoculation is also effective in correcting belief perseverance by regulating the formation of causal explanations.

In the search for decision tools, the researcher should consider the match between the target judgment and the tool. For example, in correcting the premature generation of causal explanations, the technique(s) considered should regulate them before they are formed rather than after their formation. People construct causal relations between the information and the target event whether or not they are formally required to do so. Hoch (1985) confirmed this when he found that participants who provided an explanation were not significantly more accurate than participants who did not provide an explanation. This led him to conclude that both groups of participants must have generated causal explanations even though one group was not specifically required to do so. The literature also suggests that people under-estimate the ease by which causal explanations between the information and the target event are formed and when formed, how enduring they may become (e.g., Hoch, 1984; Anderson et al., 1980). One way of improving the quality of audit judgments is to make all evidence (both supporting and opposing) more salient to the auditor.

Inoculation works by pre-empting the formation of causal explanations. Prior to examining any information, the decision-maker is required to consider all possible solutions to a problem⁹. By doing so, first, the decision-maker's cognitive processes are rendered more versatile. As all possible relations between the relevant variables have been considered, the effect of confirmation bias¹⁰ is reduced. Anderson (1982) found that participants in the inoculation condition were more ready to discard their beliefs compared to participants not exposed to the inoculation treatment. Second, all possible relations then become equally available in the memory structures of the decision-maker as belief perseverance is a result of the overt availability of only one set of arguments. Inoculation makes both (or all) sets of arguments available in the memory of participants and leads to the reduction in belief perseverance observed by Anderson (1982). By extension then, inoculation would also effectively correct explanation effects¹¹ through making opposing arguments more salient to the decision maker. Third, by making such relations available, causal explanations are prevented from forming prematurely and belief perseverance is diffused because the decision-maker is sensitised to the existence of opposing arguments. According to Anderson, inoculation's effectiveness is derived at the theory formation stage by improving decision-makers' foresight. Fourth, this causes the decision-maker to exercise greater caution in performing the subsequent task. (Anderson, 1982). Therefore, inoculation works by

⁹ One decision aid considered in the literature is counterexplanation (refer Section 2.3). Counterexplanation works by requiring the decision-maker to consider why his/her judgment may not be correct, and why the alternative rejected may be correct. This is carried out after the decision-maker has already made his/her judgment.

¹⁰ Prior psychology and some auditing studies suggest that decision-makers' cognitive processes are designed to search for information that builds on their perceptions rather than information that contradicts them. This is referred to as confirmation bias (Church 1991; Church 1990).

¹¹ Explanation effect is the tendency of decision makers to judge the explained event more likely to occur than the unexplained event.

demonstrating how easily both (or all) relations can be explained, and how difficult it is to identify which is the correct one (Anderson, 1982).

GROUP PROCESS

Studies comparing the performance of groups against individuals have been carried out in social psychology since the turn of the century with mixed results. Some studies find that group judgments are superior to individual judgments, while others find the opposite to be true (Sniezek and Henry, 1989 and 1990; Kozaiecki, 1981; Einhorn et al., 1977; Janis, 1972; Wallach et al., 1962). By the early 1960's the task of integrating the results of these studies had become impossible due to their proliferation (McGarth and Altman, 1966). Groups studies in auditing did not begin until the mid-1970's (e.g., Watson, 1975; Barrett and O'Malley, 1976). Since that time, the number of auditing studies that examined groups have grown relatively slowly. This is mainly due to the difficulty that auditing researchers have in securing sufficiently large numbers of participants (Solomon 1987).

Various theories in psychology have been used to explain the differences observed between individual and group judgments. For example, Janis (1972) quotes the example of "the Bay of Pigs fiasco" as a negative aspect of group judgments. He attributes the judgments made by the late President Kennedy and his inner Cabinet to the phenomenon of "groupthink"¹² and suggests three reasons why groupthink occurs. Groupthink occurs because of the need to preserve group cohesiveness, the secrecy of the group judgment process, and when the group leader actively promotes his/her views,

¹² "Groupthink refers to a deterioration of mental efficiency, reality testings, and moral judgment that results from in-group pressures" (Janis, 1972, 9).

the rest of the group feels a greater need to conform (1972). Wallach et al. (1962), using a series of general knowledge questions, finds that groups tend to take more risks¹³ than individuals and attributes this to the choice-shift phenomenon. Koziellecki (1981) finds excessive optimism, that is, groups tend to overestimate their chances of success and underestimate their chances of failure. Both Wallach et al. and Koziellecki attribute this to the group members' sense of anonymity. In an experiment on escalation of commitment, Whyte (1993) finds that both individuals and groups escalate their commitment to an earlier investment judgment even when there is no justification for doing so.

However, there is robust evidence supporting the proposition that group judgments are superior to individual judgments. Social psychology studies carried out by Sniezek and her colleague (e.g., Sniezek and Henry, 1989; Sniezek and Henry, 1990; Sniezek, 1992) find that groups perform better than individuals. In a general knowledge task, Sniezek and Henry (1989) find that the group judgments of their participants were more accurate than the mean or median individual judgments, and 30% of their group judgments were more accurate than the most accurate individual judgment of the members of the groups. Similar results are found by Sniezek and Henry (1990). Their participants performed a car purchasing task, and the results show that groups are more accurate than individuals, and in addition, groups outperform their best member. Group member confidence is also higher but such high confidence should not be interpreted as a proxy for judgment quality (Sniezek, 1992).

¹³ For a fuller discussion on risky shift, see Pruitt (1971).

Earlier auditing studies on groups examined the structure of public accounting firms and audit teams (Watson, 1975), the comparison of the review process within and between public accounting firms (Bamber and Bylinski, 1982), the comparison of judgment quality between individuals, interacting groups and composites (Solomon, 1982), and the differences between before- and after-group judgments (Abdel-Khalik et al., 1983). According to Watson (1975), "the chief factor that appears to influence both group and individual behaviour is the limitations upon (or opportunities for) the amount and kind of interaction" (262). Consequently, for simple and certain (concrete) tasks, less structured groups would perform better than structured groups, and in complex and uncertain tasks, structured groups would outperform unstructured groups (Watson, 1975). Watson also finds that audit teams (which are generally more structured than management services teams) operate in a more certain environment than management services teams and the degree of certainty of the task environment is positively correlated to the structure of the teams. The behaviour of the leaders also has an effect on the performance of audit groups (Pratt and Jiambalvo, 1981). Behaviours that encourage staff innovation, take into account staff personal needs, practise frequent positive feedback and reduce negative feedback, use a time budget, and limit the variety of tasks assigned, lead to higher performance.

Due to the constraint of large numbers of participants, most of the prior studies on audit groups (e.g., Trotman, 1985) required their participants to make an initial judgment as an individual. This was used as the dependent variable for the individual condition. Following this, participants were formed into groups where they made the same judgment a second time, and this judgment was used as the dependent variable in the group condition. This treatment may result in a learning effect, which may confound the

results (Iselin, 1991). Hence, the findings of prior auditing studies that used this methodology are inconclusive and should be subjected to further testing.

Various theories have been used to explain the findings in prior accounting and auditing studies. These include choice-shift theory, best member theory, information load theory and group-assisted judgments. The next sections examine these theories.

Choice-Shift Theory

Choice-shift theory states that group judgments tend to be more risky than the judgments of individual group members (see Paese et al., 1993; Wallach et al., 1962). Various psychology studies have observed this phenomenon. It has been observed among students (Paese et al., 1993; Moscovici and Zavalloni, 1969; Doise, 1969), men and women (Wallach et al., 1962), blue-collar workers (Jamieson, 1968) and professionals (Siegel and Zajonc, 1967). Four factors are suggested as possible explanations of the choice-shift observed in group judgments (Pruitt, 1971).

First, is the diffusion-of-responsibility theory suggested by Kogan and Wallach (1967) and Wallach and Kogan (1965)¹⁴. It is based on the assumption that risky judgments create stress and the group process reduces such stress (Pruitt, 1971). Because the responsibility for the judgment is shared among the groups members, stress is reduced and the group is able to make riskier judgments (Pruitt, 1971). In addition, choice-shift also occurs when the group discussion includes "novel" deliberations (Paese et al., 1993). In this case, the judgment tends to be in the direction of these "novel"

¹⁴ This refers the tendency of individuals to shift part of their responsibility to the other group members when they (the individuals) move from the individual decision mode to the group judgment mode.

deliberations. Stress exists in the auditing environment, and this stress is heightened by the complexity of some audit tasks, the severity of the consequences of incorrect judgments, and the competitive nature of the environment (for promotion and monetary and other rewards). Due to these factors and to the conservative nature of auditor training, instead of a risky-shift, a cautious-shift among auditors is a more intuitive conclusion, and is supported by prior studies which find a cautious-shift among auditors (Barrett and O'Malley, 1976; and Reckers and Schultz, 1982).

The second explanation for the observed choice-shift is familiarisation theory (Pruitt, 1971; Bateson, 1966). Because group members are able to interact and discuss the problem on hand, this increases their familiarity with the information contained in the data set. Such familiarity leads to more confidence and results in riskier judgments (Bateson, 1966). Criticisms of this explanation are, first, it is too specific as it does not explain cautious-shift, and second, the familiarisation effect cannot be replicated by other researchers (Pruitt, 1971). This theory is untested in auditing but there is no reason to believe that it would be applicable to auditors, as auditors tend to be under-confident rather than over-confident.

The third explanation for the observed choice-shift is leadership theory. This theory suggests that the more risk-taking members of the group could, with relative ease, persuade the other group members to make more risky judgments (Pruitt, 1971). This explanation has some support in the psychology literature because there is a significant correlation between the first solution proposed by a group member and the final group solution (Dubrovsky et al., 1991; Sniezek, 1989; McGuire et al., 1987). These first

advocates are usually the most talkative members (Stein, 1975), most knowledgeable (Regula and Julian, 1973), have leadership qualities (Hollander, 1978; Hemphill, 1961) and are possibly risk-takers as well (Wallach et al., 1962). On the surface, such observations are not expected among auditors for two reasons. First, due to the hierarchical nature of audit teams, the most senior team member usually makes the final judgment, and second, as conservatism increases with experience (Smith and Kida, 1992), it is unlikely that the team leader would permit a risky-shift in the team's judgment. Experience and anecdotal evidence indicate a different approach. The economic recession of the 1980's witnessed the collapse of many large corporations and resulted in a number of lawsuits against auditors for their alleged failure to adequately report on the financial position of their clients. Not all these cases were without merit given the large number that were settled out of court. These events suggest that the leadership theory is not totally unfounded.

Various accounting and auditing studies have examined group judgments in the context of choice-shift theory. These include Reckers and Schultz (1982), Schultz and Reckers (1981), and Barrett and O'Malley (1976). Barrett and O'Malley (1976) examined the effect of choice-shift among auditing students. They required student participants to estimate the contingent liabilities of a medical clinic as a result of negligent practices. Two types of tasks were manipulated - material and immaterial tasks. Students performed the task twice: once as an individual, and another as a member of a group. The results show that risky-shift is higher among group judgments than individual pre-group judgments for the immaterial tasks and a cautious-shift is observed for the material tasks.

In an attempt to explain the cautious-shift phenomenon observed by Barrett and O'Malley, Schultz and Reckers (1981) used diffusion-of-responsibility theory and social value theory¹⁵ to explain the choice-shift phenomenon observed in the context of group judgments. Auditor participants were required to make two judgments - one as an individual and another as a member of a group - on whether a contingency footnote disclosure was necessary. No significant difference in choice-shifts was observed between individual and group judgments even though group judgments were subjected to less variation. In addition, groups also made use of more information in making their judgment.

In a follow-up experiment to test for choice-shift, Reckers and Schultz (1982) divide their uncertainty disclosure tasks into three parts. In the first part and for each of two tasks, student participants were given an explanation of the task, economic information about the client and the latest financial report of the client. Subsequently, student participants made recommendations on the necessity for uncertainty disclosure. In the second part, they were randomly assigned to groups of four. Participants were required to reach a consensus (if possible) on the issue of disclosure. In the third part, participants were separated and made disclosure judgments on their own. The results show that for high materiality cases, there was a significant cautious-shift from individual to group judgments; and for low materiality cases, there was a significant risky-shift from individual to group judgments. These findings support those of Barrett and O'Malley (1975). The findings of these studies suggest that auditors generally are not subject to the risky-shift observed among participants in the psychology studies

¹⁵ This theory suggests that when individuals move to a group judgment mode, they experience an effect that is revealed in the need to maintain a favourable self-perception and self-presentation.

(except in the low materiality tasks). In fact, for the high materiality tasks, a cautious-shift is observed. This tendency towards conservatism could be explained by the conservative nature of the audit environment.

Best Member Theory

Best member theory is premised upon the assumption that group members are able to identify their best member and follow his/her judgment so that the judgment of the group is more effective than the average judgment of its members (Einhorn et al., 1977).

The use of audit groups as a vehicle for decision-making is predicated on the assumption that group judgments are superior to individual judgments. As the sayings go - "two heads are better than one" and "many hands make work light". But from a review of the literature, it is unclear whether interacting audit groups are definitely better than individual auditors and composites of auditors as decision-making units.

However, audit groups, individual auditors and composites of auditors will continue to be used. There are situations where interacting audit groups would outperform individual auditors and composites of auditors and situations where the reverse is true. The challenge is to identify these situations.

Three fundamental factors determine a group's ability to perform and these are the nature of the task, the relative knowledge of the individual members and the judgment process adopted by the group (Steiner, 1966). Based on these factors, Einhorn et al. (1977) propose that in a judgment task, the final judgment is a result of "weighting and combining of the judgments of the individual members" (159). Therefore, the means by which the group weighs and combines the judgments of individual members and the impact of such weightings and combinations on the final judgments would have an

important effect on the quality of the group's judgment (Einhorn et al., 1977). Thus, group interaction plays an important part in determining judgment quality. Einhorn et al. developed four models to measure "whether groups do misweight in actuality, and with what frequency" (159). In the random model, the group randomly picks a member's judgment as a surrogate for the group's judgment. The judgment quality of this type of group is generally low because it assumes that members are unable to identify their better group members. The group's performance can be improved by simply averaging the judgment of its members. This leads to the second type of group (average model) where each group member's judgment is weighted equally. The disadvantage of this type of group is that if group members simply weigh each other's judgment equally, it would be a lot more economical to take an average of individuals' judgment than to have them come together as a group to make the same judgment. Best member model assumes that group members, through their interaction, are able to identify and use their best member's judgment as a surrogate for the group judgment. While the group judgment may improve using this technique, it may not in all cases. This shortcoming is taken into account in the fourth model, which is called the proportional model. A group may not be able to identify its best member and may mistakenly nominate the judgment of a less-able member. So the probability that the group may not have identified its best member is factored into the proportional model.

According to Einhorn et al. (1977), the best member model outperforms the average model at all levels of standardised bias¹⁶. They find that standardised bias and group

¹⁶ Two types of errors visit upon the judgments of individuals. These are random error and standardised bias (Einhorn et al., 1977). Random error is merely the distance between the true value and the judgment of the individual. This is stated as:

$$b = (x_i - \mu) \quad (1)$$

size affect the quality of the judgment of the random and average models. The average model can reduce the random error or discrepancy of the individuals' judgment and, where standardised bias is low (e.g., among experts), the average model may increase the accuracy of the judgments. But where standardised bias is large (e.g., among novices), the proportional model is preferred to the average model.

Best member theory makes two assumptions. First, where standardised bias occurs, interaction among group members results in the identification of individual member expertise, and second, having individual group members with different levels of expertise, allows group members to follow their best member (Libby et al., (1987). For standardised bias to occur, a task must be reasonably complex (Einhorn et al., 1977), so that the more complex the task, the greater the ability of groups to outperform individuals by identifying their best member. It follows, therefore, that the best member must be significantly better than the average of the group members. If the best member is only marginally superior to the average of the group members, then the group judgment is only marginally better than the average of the group members (Iselin, 1991). In such a case, it may be more economical to simply take the average of members' judgment rather than incur the added expense of assembling a group of interacting decision makers. Libby et al. (1987) and Trotman (1985) find that bank loan officer and auditor groups respectively are able to identify their best members, and consequently outperform the composites. Similarly Yetton and Bottger (1982) find that group members can identify their best member, but Miner (1984) finds that such an

where b is the random error, x_i is the true value and μ is the mean of the individual judgments (Einhorn et al., 1977). Standardised bias, however, is the difference between the mean of a population of individual judgments and the true value and is denoted as:

$$\beta = (x_i - \mu) / \sigma \quad (2)$$

ability is only slightly better than chance. Following best member theory approach may not improve the quality of the judgment (Hall and Watson, 1970) especially when group members do not attempt to identify their best member and resort to a majority judgment, and they confuse experience with relative expertise (Iselin, 1991). Using students in an audit task, Trotman et al. (1983) finds that their groups could not identify their best member.

Various accounting and auditing studies have examined the effect of best member theory on the effectiveness of decision groups. These include Uecker (1982), Chalos (1985), Libby et al. (1987), Solomon (1982), Trotman et al., (1983), Trotman and Yetton, (1985) and Trotman (1985). Student participants were used in Uecker's study, which required them to choose sample sizes to be used for quality control in an information system. In all, groups of three and individuals were required to choose 32 sample sizes. Participants performed the tasks initially as individuals and later, as members of a group. The results show that interacting groups performed better than composites. Even though the group process led to improved judgments, these judgments were not at their best as they were outperformed by the statistical model (Uecker, 1982).

In Chalos (1985), using best member theory, loan officers were required to make loan default judgments using financial ratios and the three latest sets of financial reports. Significant differences between individual and group judgments were noted, that is, group judgments were superior to individual judgments with the groups making fewer errors. These included both Type I and Type II errors. Interacting loan committees,

where β is the standardised bias and σ is the population standard deviation (Einhorn et al., 1977).

however, marginally outperformed composites. In addition, both composites and interacting groups did not perform as well as "the model of man" but they both outperformed the individuals.

Libby et al. (1987) hypothesised that the effectiveness of best member theory is dependent upon the characteristics of the interacting groups. To test this, both ad hoc and practised groups were used in their experiment which required loan officers to make 39 bankruptcy predictions. These predictions were based on the companies' financial profiles which comprised five financial ratios. The practised groups were made up of three loan officers from the same bank while the ad hoc group members were randomly assigned. The same participants provided both the individual judgments and the interacting group judgments. No difference in performance is observed between interacting groups, composites and the most influential member of each group. The results also show that interacting groups are unable to identify their best member. Libby et al. suggest that interacting groups are unnecessary or the group could be disbanded once their best member has been identified. In addition, practised groups do not outperform ad hoc groups. This could be due to the fact that the advantage of practice may be offset by other social pressures (Libby et al., 1987).

In an experiment with volunteer business and government administrators, Iselin (1991) finds that interacting groups do not outperform composites. The experiment manipulated two levels of information load (high and low), three types of decision-making units (individuals, interacting homogenous groups and interacting diverse groups), and two levels of task learning (these were before and after repeated-measure dependent variables). The experimental task required participants to predict the

bankruptcy/loan default of ten companies (comprising seven non-failed and three failed companies). Participants received five financial ratios for the low information load level and the high information load participants received the same five financial ratios plus balance sheets and income statements for two years. Participants performed four predictions (which were not used in the analyses) followed by another six predictions. They received no feed-back on these predictions. Based on the initial four judgments, they were designated as experts or novices, and randomly assigned to homogeneous and diverse interacting groups and individuals. Homogeneous groups comprised either three experts or three novices while diverse groups comprised one expert and two novices. They were then given the second six predictions in a new questionnaire and instructed to improve on their initial predictions. The results indicate that groups are unable to and do not identify their best member.

In Solomon (1982), auditor participants were given information about the client including size and industry, information related to the account balance to be audited, the relevant internal control procedures and results of compliance testing (which included sample size and frequency and nature of errors detected). Participants were required to estimate prior probability distributions of various account balances. The results indicate that interacting groups are marginally superior to individuals but they (interacting groups) are significantly more superior to the composites.

Trotman and his associates carried out a series of auditing studies on group judgments and achieved mixed results (Trotman et al., 1983; Trotman and Yetton, 1985; Trotman, 1985). Trotman et al. (1983) required their student participants to make payroll internal

control evaluations. The students made the same evaluations, first, as individuals and second, as groups. The results, which were measured using consensus (after removal of the diversification effect¹⁷), show that groups did not outperform composites. The authors attributed these results to the student participants' inability to identify the best member in their group and concluded that there was "no justification for the introduction of group judgments to replace individual judgments. Instead, we could simply combine independent judgments to improve performance" (290). The use of student surrogates may have accounted for the inability of the audit groups to outperform the composites as the students were unable to identify their best member.

In a follow-up study, Trotman and Yetton (1985) used audit managers and audit seniors in an audit review task and reached the same conclusion, i.e., interacting groups did not outperform composites. Participants evaluated whether internal controls in the payroll area were effective. In all, they made 15 evaluations. Participants made their evaluations initially as individuals. They were then randomly assigned to groups and were allowed to keep a copy of their initial evaluations. The groups comprised one manager (who performed the review process) and two audit seniors. Both interacting groups and composites were manipulated. The results show that interacting groups did not achieve higher consensus than composites in internal control evaluations. The 15 internal control evaluation tasks used were described as well understood and well structured. This being the case, they did not allow the best member to outperform the rest of the group, i.e., standardised bias was reduced. Prior studies (e.g., Chalos and Pickard, 1985)

¹⁷ Diversification effect refers to the reduction of random errors attributed to individual judgments through the use of composite groups. This results in a tighter distribution of judgments around the mean.

find that groups outperform individuals by suppressing outlier behaviour among individuals. Because the tasks used by Trotman and Yetton were relatively simple, the outlier effect was reduced and consequently, no significant difference between audit group and composite performances was observed.

Instead of using consensus as a surrogate for accuracy [as did Trotman et al. (1983) and Trotman et al. (1985)], Trotman (1985) measured accuracy. His auditor participants were required to estimate the expected dollar error for total inventory and the total required audit hours. Unlike prior studies, Trotman made use of practised groups. Composites comprised one manager and an audit senior from the same firm while interacting groups comprised two seniors from the same firm. Again, the same auditor participants provided both individual and group judgments. The results show that groups outperform composites, but this result could be affected by a learning effect (Iselin 1991).

Information Load Theory

Information load theory, which was proposed by Chalos and Pickard (1985), states that groups process large amounts of information better than individuals. As individuals have limited information processing capacity, large amounts of information would normally cause cognitive strain (Hogarth and Einhorn, 1992). In the case of groups, however, several individuals working together would compensate for the individuals' inadequacies. This would increase the group's corporate processing capacity and lead to more effective judgments. In addition, information load theory states that group judgments are more consistent than individual judgments. Prior psychology studies find that individuals weigh information cues with moderate accuracy but this individual

ability is applied rather inconsistently (Brehmer, 1976). This results in the inconsistent judgments observed among individual decision-makers compared to group judgments (Chalos and Pickard, 1985; Einhorn et al., 1977). This theory finds support in the results of Schultz and Reckers (1981). They find that, compared to individual auditor judgments, audit group judgments are more consistent, and audit groups make use of more information in arriving at their judgment.

For a group to work effectively, group members must pool their processing of information cues and integrate the cues into a solution. This normally occurs during the group interaction process. During this interaction, suggestions and counter-suggestions are put forward by group members. An effective group should be able to identify a suggestion as its best or optimum solution. Failure to recognise this reduces the effectiveness of the group. When group members are unable to share information processing among themselves, they become less effective (Iselin, 1991). There is evidence in the literature to suggest that interacting groups are better able to integrate their information processing to reduce the uncertainty of the task, eliminate redundant information search and improve judgment quality (Howell et al., 1970). To reap the benefits of group interaction, group members must pool and integrate information cues to form a solution (Hill, 1982). However, some studies show that, in general knowledge tasks, the group judgment processes result in a loss of information when information is being accumulated (e.g., Tuckman and Lorge, 1962), combined (e.g., Howell et al., 1970), or used in the decision-making stage (Laughlin and Branch, 1972). When group members are unable to share information processing among themselves, they do not outperform composites (Iselin, 1991).

Prior accounting studies that examine the group process using information load theory include Chalos and Pickard (1985), Iselin (1991) and Stocks and Harrell (1995). To date, no auditing study has examined the difference between audit groups and individual auditors using this theory. In a task involving loan review, Chalos and Pickard (1985) hypothesise that information load theory could be explained by the difference in cue selection, cue weighting, and judgment consistency between groups and individuals. This is based on the assumption that groups process information and integrate knowledge to form a judgment more effectively and efficiently than individuals. Individuals, on the other hand, may interpret information inconsistently, and may miss important cues, so that their judgment may not be the optimum and their processing of information less efficient. Chalos and Pickard's results, however, show no significant difference in cue weighting between groups and individuals, but both these decision units outperform the composites. This suggests that groups and individuals are equally effective in their weighting of information. But, composites outperform both groups and individuals in cue selection. There is a high degree of agreement between both groups and individuals in their choice of cues. On the other hand, the group judgments are more consistent than the individual judgments. This finding supports prior literature (most notably Einhorn et al., 1977) that groups' ability to handle higher information loads would eliminate the inconsistencies of individual judgments. This study, however, suffers from several confoundings (see Iselin, 1991, 1995). As with other group studies, this one also uses the same participants for both individual and group judgments. The second confounding is the manipulation of the information load. In the low load condition, participants were given three financial ratios with which to make a loan judgment while participants in the high load condition received the same financial ratios plus the full set of financial report. Limiting the available information to merely three

financial ratios increases task difficulty and reduces the external validity of the study (Iselin, 1991).

In an attempt to overcome the confoundings of Chalos and Pickard's (1985) study, Iselin retests information load theory with business and government administrators using a bankruptcy prediction task. He finds no significant difference between the judgments of groups and individuals as well as between participants in the different information load conditions. The finding that groups could not process more information than individuals is attributed to the inability of the group members to share the information processing so that they could outperform the individuals (Iselin, 1991). The results of Iselin's study are, however, affected by a limitation, that is, Iselin assumes that providing participants in the high load condition with more information would result in higher information processing. Extant psychology literature shows that this may not be the case. Human information processing makes use of various biases and simplified processing strategies (see Hogarth, 1980). If the five financial ratios provided are the most salient cues, then the provision of the financial report is redundant. Iselin did not measure cue usage as did Stocks and Harrell (1995). Therefore, there is no assurance that the provision of different information loads leads to increased information processing between the two load conditions (especially when the information provided to the low load condition was adequate for bankruptcy prediction). However, perception of information load was measured, and participants in the high load condition rated the information load to be significantly higher than participants in the low load condition. Iselin finds that business and government administrators are unable to share information processing responsibilities. Nevertheless, he does not rule

out the proposition that groups outperform individuals but believes that it would occur under some (yet unidentified) conditions.

In an experiment that manipulates two levels of information load (high load - nine financial ratios; low load - six financial ratios) and two judgment modes (individuals and groups), Stocks and Harrell (1995) required loan officers to make financial distress predictions. The prediction tasks included bankruptcy, loan default, non-payment of preference shares dividend and sale of assets. The dependent variables were cue use, accuracy, consistency and consensus. Compared to individual auditors, audit groups made use of more cues in making judgment for both load conditions. Groups were more accurate than individuals in the high load condition but both groups and individuals were equally accurate in the low load condition. The judgments of the groups were more consistent in both load conditions and the same was true of judgment consensus.

Group-Assisted Judgments

In the audit environment, auditors often have the benefit of consulting with or seeking advice from colleagues on particular audit issues (Reckers and Schultz, 1982 and 1993; Gibbins and Emby, 1984; Solomon, 1987). Such consultation exposes the auditor to the views of others and may result in a change in their prior judgment (Pruitt, 1971). According to Pruitt, through the consultation process, decision makers become aware of information and/or arguments that they may not be aware of when working on their own. This pooling of arguments and counterarguments should allow groups to identify the optimum solution. This suggests that group-assisted judgments are more effective than unassisted judgments. But the findings in the psychology literature are mixed. Some studies find that decision-makers often perceive their own judgment to be better

than it really is as well as to be superior to those of the other group members (Lamm et al., 1972; Myers, 1975). In addition, there is a tendency to perceive the responses of the other group members to be more extreme than their own (Baron et al., 1973). Similarly, the results in auditing studies are mixed and these are discussed below.

There are limited studies in auditing that examine group-assisted judgments and these include Reckers and Schultz (1982 and 1993), Johnson (1994), and Abdel-Khalik et al. (1983). Reckers and Schultz (1982) examined the effect of individual versus group-assisted judgments on the choice-shift phenomenon. They hypothesised that group-assisted judgments would shift towards one extreme of the individuals' prior judgments and this shift would be determined by the direction of the majority's views. In addition, consistency of judgment was also examined. Student participants were required to determine the adequacy of contingency disclosure. The results did not show any choice-shift but consistency of judgment after group-assisted evaluation was found to be higher.

In Abdel-Khalik et al. (1983), the effect of group-assisted planning of external audit programs was investigated together with types of EDP-audit techniques used by internal auditors and levels of internal auditors' responsibilities. The authors hypothesised that individual responses are stable, and therefore, would not change as a result of group-assisted discussion. The results show a high degree of correlation between participants' before- and after-discussion judgments. This indicates that group-assisted discussion does not have an effect on participants' initial judgment.

Reckers and Schultz (1993) examined whether group-assisted discussion has an effect on auditors' ability to identify fraud signals. Failure on the part of auditors to detect material fraud would incur high costs and group-assisted discussion is one way of reducing the cost of litigation (Reckers and Schultz, 1993). It is expected that an auditor's reaction to new information is dependent upon group-assisted discussion as well as his/her perception of the likelihood of fraud. Using an inventory task, Reckers and Schultz (1993) had auditor participants assess the likelihood of fraud in a given case before dividing them into group-assisted participants (group-assisted condition) and individual participants (individual condition). They were then required to make two assessments: first, whether the financial report contained fraud; and second, whether the inventory balance should be written down for obsolescence. Later they reviewed the additional information in either a good news/bad news order or a bad news/good news order. Participants in the individual condition made repeated judgments on the existence of fraud and write-down of inventory due to obsolescence while participants in the group-assisted condition made these judgments after consulting with their group members. The results support the hypothesis and find that auditors' reaction to fraud signals improved with group assistance.

Based on the assumption that an auditor's memory forms a vital part in determining judgment quality, Johnson (1994) carried out an experiment which manipulated two types of judgments - individual and group-assisted judgments, two levels of time delay - one hour and 24 hours, and two memory tasks - recall-then-recognition or recognition-only. Participants were required to perform an audit workpaper review. The workpapers contained descriptions of the client's business, management, and operations as well as industry data, risk assessments of merchandise inventory and long-term debt and

workpapers of work performed by an entry-level audit staff. In part one of the experiment, participants were required to review and comment on the workpapers. In part two, participants were randomly assigned to one of two memory tasks. After a time delay of one hour or 24 hours, part three of the experiment participants were randomly assigned to individual or group-assisted conditions where the tasks in part one were repeated. The results indicate that the number of items recalled and accuracy of recall are higher for the group-assisted condition and the shorter time delay. There was an interaction effect between type of judgment (individual or group-assisted) and time delay on accuracy and confidence. In addition, participants in the group-assisted condition made fewer Type II errors in recognition memory than individual auditors. The group-assisted participants also demonstrated over-confidence.

In sum, choice-shift theory finds that decision-making groups have a tendency to take more risks than individual decision-makers. This finding in the psychology literature is quite robust. However, conflicting results are observed in the auditing literature. Auditing studies find a cautious-shift on material matters and a risky-shift on immaterial matters (Barrett and O'Malley, 1976; Reckers and Schultz, 1982). On the other hand, Schultz and Reckers (1981) find no difference in choice-shift between audit groups and individual auditors.

Best member theory states that through interacting with each other, group members are able to identify their best member and use that member's judgment as a proxy for the group's judgment. This should enable groups to outperform individuals. Prior auditing studies based on best member theory produce mixed results. Some report that audit groups do not outperform individual auditors (Trotman et al., 1983; Trotman and

Yetton, 1985), while others find that audit groups do outperform individual auditors (Solomon, 1982; Trotman, 1985).

Information load theory states that audit groups should be able to process larger amounts of information compared to individual auditors. Because of this, audit group judgments should be more consistent than the judgment of individual auditors. In addition, differences in cue selection and cue weighting should be observed. No prior auditing study has examined audit group judgments in the context of information load theory. Studies in accounting that have been premised on this theory find no difference in cue selection and cue weighting and group members are unable to share the information processing responsibilities (Chalos and Pickard, 1985; Iselin, 1991). However, group judgments are more consistent than the judgments of individual auditors (Chalos and Pickard, 1985). Stocks and Harrell (1995) find that, on the whole, groups are able to process higher information loads more effectively than individuals.

Individual auditors working alone, may not be aware of all the arguments and counterarguments surrounding a particular task. Consultation with colleagues may highlight the existence of such arguments. Group-assisted judgments are based on these assumptions. Prior auditing studies show that group-assistance results in higher consistency (Reckers and Schultz, 1982), improves auditors' detection of fraud signals (Reckers and Schultz, 1993), and increases auditors' recall of information and accuracy of recall (Stocks and Harrell, 1995). Abdel-Khalik et al. (1983) on the other hand, find that group discussion has no effect on individual auditor's judgments.

COUNTEREXPLANATION

Counterexplanation is the provision of reasons that either speak against or provide evidence against the chosen alternative, or speak for or point towards the alternative rejected (Koriat et al., 1980). It is seldom required in decision-making and decision-makers do not habitually counterexplain. Counterexplanation has led to more effective judgments in the social psychology literature. In a general knowledge task, Koriat et al. (1980) used counterexplanation with some success to correct over-confidence in student participants. Half the participants wrote explanations in support of their judgments and the other half wrote counterexplanations. They then rated their confidence in having made an accurate judgment. The results show that counterexplanation is useful in reducing over-confidence in students' judgments, as over-confidence is partly due to the neglect of disconfirming information (Koriat et al., 1980). Hoch (1985) finds similar over-confidence with student participants, and like Koriat et al., hypothesises that such over-confidence could be mitigated by the use of counterexplanation. Final-year university students were required to predict their job prospects at graduation in nine months' time. Participants in the manipulation groups then wrote explanations and/or counterexplanations. Hoch divided his participants into four groups; a control group, a pro group (who wrote explanations), a con group (who wrote counterexplanations), and a pro-con group (who wrote both explanations and counterexplanations). Students, consistent with prior psychology studies, were over-confident in their judgments and this over-confidence was corrected after the provision of a counterexplanation. The control group and the pro group showed similar high confidence which supported Koriat et al.'s prior findings.

The provision of a counterexplanation before an explanation results in lower likelihood judgments while the provision of an explanation before a counterexplanation does not (Hoch, 1984). This is because, in the latter case, the provision of the first set of reasons interferes with decision-makers' ability to consider counterexplanations, and therefore, an explanation effect is observed (Hoch, 1984).

Ross et al. (1975) demonstrate that people's beliefs in social theories are amazingly resilient to change. Such beliefs can be based on very little information and can survive even after the initial information has been totally discredited (Anderson et al., 1980). Anderson (1982) predicts that such belief perseverance can be corrected. The procedures were described above. The results show that, in addition to the effectiveness of inoculation in correcting belief perseverance, counterexplanation is similarly effective in debiasing belief perseverance.

Anderson and Sechler (1986) used counterexplanation to correct belief perseverance in social theories. As stated above, prior studies show that people's beliefs in social theories are surprisingly resilient to change. Exposure to new data may result in even more extreme beliefs, especially for emotive social theories such as capital punishment (Lord et al., 1979). In an experiment by Anderson and Sechler, participants were divided into four groups. In Group One, participants were required to explain the positive relation between success as a firefighter and risk-taking propensity (a nonemotional issue); Group Two explained the negative relation between success as a firefighter and risk-taking propensity. Group Three participants wrote both positive and negative relations and Group Four did not either explain or counterexplain. Participants were then required to consider the suitability of 16 applicants for the position of firefighter.

Four pieces of information were provided for each applicant: gender of applicant, risk preference, intelligence and physical capabilities. These characteristics were combined in a 2^4 factorial resulting in 16 cases. Participants rated each applicant on a seven-point scale anchored by very unacceptable and very acceptable. The results indicate that counterexplanation successfully reduces the explanation effect as the latter is partly due to ignoring alternative arguments.

The studies by Koriat et al. (1980), Anderson (1982), Hoch (1984), Hoch (1985), and Anderson and Sechler (1986) involve the application of counterexplanation in knowledge tasks. The real test for counterexplanation is in domain-specific tasks. Tutin (1993) finds that counterexplanation is only marginally effective in reducing the explanation effect among clinical psychologists. Participants received two detailed patient case histories, each containing approximately 1,400 words. Significant judgment differences were noted between participants who explained/counterexplained a particular case first and participants who explained/counterexplained the same case second. Physiological reasons could have confounded the results. While domain-specific studies are limited, the initial results of the use of counterexplanation among auditors appear encouraging (Heiman, 1990; Koonce, 1992; Kennedy, 1995).

On the premise that counterexplanation is often required for analytical procedures, Heiman (1990) gave all participants case materials containing the current year's unaudited financial report together with three financial ratios. Participants were asked to respond to the following question: "Given the fact that a financial statement error occurred and affected the ratios on the previous page, how likely do you think it is that the error was next period's credit sales being recorded in the current period?" Here,

participants made their first likelihood judgment (pre-test judgment). After which, they were divided into three groups. One group of participants considered five counterexplanations offered by the experimenter (experimenter-provided group). The second group generated their own counterexplanations for the error (self-generated group), while a third group acted as a control. Subsequently, all participants made a post-test likelihood judgment. A within-subject interaction between pre-test/post-test likelihood judgments and source of alternative is observed. As predicted, the between-subject post-test likelihood judgment of the experimenter-provided group was significantly lower than their pre-test likelihood judgment. Significant within-subjects pre-test and post-test likelihood judgment differences were also observed. The results support the proposition that counterexplanation affects likelihood judgments. In a second experiment, for the two manipulation conditions, the number of counterexplanations considered by participants was varied in number and in strength. The results indicate that auditors' likelihood judgments are affected by the number of counterexplanations considered but, contrary to prior studies (most notably Einhorn and Hogarth [1986]), not by the strength of the counterexplanations.

Using an analytical review Koonce (1992) carried out an experiment where the order of explanation and counterexplanation was changed. After reading the case materials, participants either wrote explanations prior counterexplanations or vice versa. A third group wrote only explanations. An explanation effect is observed for the explanation-only group. For the dual-explanation groups, a recency effect was observed, i.e., counterexplanation was effective in reducing the explanation effect but only in the group that explained before counterexplaining. This finding contradicts Hoch (1984). Hoch found a primacy effect instead and used the inference theory to explain his results.

In sum, counterexplanation is effective in correcting over-confidence (Koriat et al., 1980; Hoch, 1985). While explanation leads to increased likelihood judgment for the explained event, counterexplanation results in lower likelihood judgments (Hoch, 1984; Anderson and Sechler, 1986), it also increases decision-makers' readiness to abandon their initial judgment, and hence, moderate any belief perseverance (Anderson, 1982). In an auditing context, counterexplanation similarly reduces the explanation effect (Heiman, 1990; Koonce, 1992), and a recency effect has been observed (Koonce, 1992). Another counterexplanation effect observed is the amelioration of confirmation bias (Chung and Monroe, 1996; Church, 1991). Counterexplanation has also been used to debias the curse of knowledge in an auditing context (Kennedy, 1995).

Apart from Tutin (1993) and Anderson and Wright (1988), prior studies reported in the social psychology literature examined the explanation effect on student participants in general knowledge tasks. It is unclear whether experts, working on a domain-specific task would demonstrate the same explanation bias. The studies by Tutin and Anderson and Wright report that experts are subject to less explanation bias. However, other studies in auditing find the opposite; they find that auditors similarly have an explanation bias (e.g., Ahlawat, 1992; Koonce, 1992). In addition, Koonce finds that counterexplanation is effective in reducing the explanation bias but only if the auditors are required to explain prior to counterexplaining (a recency effect). But her findings may have been affected by a confounding, as a significant task difficulty effect is observed between the two groups (Koonce, 1992). Auditors considering counterexplanations are able to generate more alternative hypotheses to explain the findings of analytical procedures (Heiman, 1990). Auditing researchers have only

recently begun to examine the effects of explanations and counterexplanations on audit judgment. Consequently, the range and depth of these studies are limited and further studies would be able to add substantially to our understanding of explanation and counterexplanation in audit judgment.

Counterexplanation has been observed to work in four ways. First, the explanation bias is the result of causal explanation, therefore, any judgment tool must start with an examination of causal explanation. As pointed out by Koriat et al. (1980), causal explanations are formed automatically as a by-product of decision-making, so preventing it is nigh impossible (Anderson, 1982). Besides, the question of explanation bias is not so much related to causal explanation as it is to our underestimating the ease of its (i.e., causal explanations) creation (Anderson and Sechler, 1986; Ross et al., 1977; Ross et al., 1975). Anderson and Sechler (1986) propose that in the face of two competing solutions to a problem, the solution where causal relations are easier to imagine or recall would be the one adopted. This implies that the explanation bias found in the social psychology studies is a result of ignoring opposing arguments and is easily created. Prior studies find that merely asking participants to explain the hypothetical relation between two variables creates an explanation effect (e.g., Anderson et al., 1980). The provision of an explanation increases the availability of causal relations between the information and the explained event, and leads to an explanation effect. One way of reducing the explanation effect is to increase the availability of opposing arguments. Anderson and Sechler (1986) and Anderson (1982) suggest that the technique used to debias belief perseverance, i.e., counterexplanation, may be effective in debiasing the explanation effect at both the individual and group levels. In their experiments, Anderson and Sechler find that counterexplanation is effective in reducing

the explanation effect, leading them to conclude that the explanation effect is due to the neglect of opposing arguments. They explain that:

“The main error leading to the explanation effect is not in using the availability of plausible causal explanations in judging the probable relation between two variables. Rather, the error seems grounded in people’s inability (or unwillingness) to see that the availability of a particular explanation may have been due to factors unrelated to the truth of the explanation, and that equally plausible causal explanations could be generated for alternative or opposite variable relations” (32).

Second, another way of correcting the explanation effect in audit judgment is to make opposing evidence more salient through the use of counterexplanation (Anderson, 1982). Anderson’s results show that counterexplanation leads to a reduction in belief perseverance and he attributes this to four factors. First, the consideration of counterexplanation increases versatility that may result in lower belief perseverance. Second, belief perseverance may be a result of availability, so the act of considering counterexplanations makes opposing arguments equally available and belief perseverance is reduced. Third, the order by which these arguments are generated does not impede the effectiveness of these procedures (though other studies such as Koonce [1992] find otherwise). Fourth, counterexplanation causes theory revision. However, beliefs in social theories are quite different from beliefs in self or social event occurrence (Anderson and Sechler, 1986) and are presumably even more different from domain-specific tasks such as auditing. Beliefs in social theories are relatively stable, e.g., capital punishment (Lord et al., 1979) and beliefs in self or social events may be formed ahead of the counterexplanation. Although belief perseverance and the explanation effect are dissimilar phenomena, they are both results of causal

explanations. Therefore, counterexplanation focuses on correcting causal explanations as a means of correcting both belief perseverance and the explanation effect.

Third, the findings in the literature of over-confident decision-makers is quite robust but this is generally confined to general knowledge tasks and other professionals (apart from auditors). Hoch (1985) uses theories of availability, confirmation bias and interference to explain over-confidence. It is also suggested that the explanation effect may be caused by over-confidence (Hoch, 1985; Koriat et al., 1980). If a decision maker is confident of an explained target event, then s/he would increase the subjective likelihood judgment of the target event occurring. Correcting over-confidence may lead to a reduced level of the explanation effect. Koriat et al. (1980) first used counterexplanation to improve the calibration¹⁸ of decision-makers. As causal explanation (or the generation of supporting reasons) is a by-product of decision-making, over-confidence¹⁹ results when counterexplanations are ignored (Koriat et al., 1980). Decision-makers such as auditors do not generate counterexplanations automatically and must be specifically instructed to do so (Koriat et al., 1980). Requiring decision-makers to explicitly consider counterexplanations forces them to consider disconfirming information and over-confidence is reduced. Accuracy may also be improved by increasing the consideration of opposing evidence (Hoch, 1985).

Fourth, similar to correcting the curse of knowledge, counterexplanation works by reducing the strength of the causal explanation between the information and the target

¹⁸ Perfect calibration occurs when a decision-maker's rate of accurate judgments matches his/her rate of confidence in those judgments.

¹⁹ Over-confidence results when the decision-maker's rate of confidence exceeds his/her rate of accuracy.

event (Kennedy, 1995). By requiring participants to focus their attention on counterexplaining, an experimenter-prescribed outcome prior to making predictions about other participants' estimate of the outcome, Kennedy predicts that this manipulation would weaken the tendency to construct causal explanations. She finds inconclusive evidence that counterexplanation leads to professional scepticism.

Apart from Heiman (1990) and Koonce (1992), auditing studies do not consider the effectiveness of the resultant judgment. Heiman and Koonce consider the effect of counterexplanation on the explanation effect and find that if auditors provide a counterexplanation for an inaccurate judgment, they are more likely to revise their judgment. It follows that if auditors counterexplain an accurate judgment, they should be able to distinguish an accurate judgment from an inaccurate one. So after counterexplaining an accurate judgment, they should not be persuaded by the opposing arguments to revise their judgment. If auditors provide a counterexplanation for an accurate judgment, and they revise their judgment to a less accurate one, then counterexplanation may be counter-beneficial. Therefore, the real strength test of counterexplanation would be for auditors who make accurate judgments to not revise their judgment in spite of having to counterexplain their initial judgment.

CHAPTER 3

EXPERIMENT ONE - THE EFFECT OF INOCULATION, FRAMING AND AUDIT EXPERIENCE ON THE EVALUATION OF INTERNAL CONTROLS

INTRODUCTION

This experiment examines the effect of inoculation, framing and audit experience on internal control evaluations and is motivated by five factors. First, prior studies suggest that in making judgments, people generate causal explanations as a bridge between the information and the problem (Tversky and Kahneman, 1973). These causal explanations once generated cannot be readily changed, and because of their inflexibility, the quality of the resultant judgment is reduced (Hoch, 1984). This shortcoming is further exacerbated because people underestimate how easily causal explanations are generated (Hoch, 1984). Therefore, in searching for ways to improve audit judgments, consideration should be made of whether the technique being considered would regulate the premature formation of causal explanations. Consequently, inoculation is chosen in this investigation of alternative means of improving audit judgment because of its ability to regulate the formation of causal explanation.

Second, the accounting literature suggests that framing has an effect on the nature of judgments elicited (e.g., Sullivan, 1996; Emby, 1994; McMillan and White, 1993; Kida, 1984), as well as the manner in which information is processed (e.g., Dunegan, 1996). In general, a negatively framed judgment problem results in more comprehensive information processing than a positively framed judgment problem (Dunegan, 1996), and consequently, the judgment is also affected. Greater understanding of this heuristic would lead to the development of more effective training techniques. This study extends

prior auditing research by investigating whether framing effects can be reduced or eliminated by the use of inoculation.

Third, biased decision makers may be divided into two categories - perfectible and incorrigible decision makers (Fischhoff, 1982). The former refers to decision makers whose judgments could be improved by various corrective and educational techniques. The study of this type of decision maker is important because it helps trainers and educators understand how their judgments may be improved. There is a trend in the current auditing literature that investigates ways of improving audit judgments (e.g., Koonce, 1992). Incorrigible decision makers, on the other hand, are decision makers who are not responsive to training. Through the use of inoculation, we may understand the magnitude, predominance and inflexibility of their biases.

Fourth, most behavioural research in auditing focuses on experienced auditors, and consequently, we have limited knowledge of how auditors' knowledge structures and their understanding of audit-related events are developed with experience (Libby and Luft, 1993; Waller and Felix, 1984). In addition to knowledge structures and understanding, the literature suggests that novices and experts process information differently (Schoenfeld and Hermann, 1982). Prior studies have not examined the relationship between audit experience and responsiveness to judgment tools. Consequently, little is known of the nature of this relationship. While a linear relationship is normally assumed, this study postulates that this relationship is actually nonlinear and would take on an inverted-U shape. There are policy implications for public accounting firm management, as identifying this would allow them to direct

training resources at the most receptive group of audit staff. Using auditors with different levels of audit experience, this experiment will model the effect of inoculation on auditors with different levels of audit experience and identify those auditors who are most receptive to training.

Finally, understanding the effect of presentation order of internal control procedures is important because prior studies suggest that changing the presentation order of these procedures may result in different conclusions about control risk (see Hogarth and Einhorn, 1992). Prior studies manipulate the presentation order of information and find a recency effect (e.g., Asare, 1992). Koonce (1992), on the other hand, examines the order effect of writing supporting and opposing arguments (after the participants had examined the information) while holding the presentation order of information constant. She finds a recency effect instead of the hypothesised primacy effect and attributes this to the participants' perception of task difficulty (Koonce, 1992). In this study, the presentation order of information is similarly held constant. However, participants are required to write supporting and opposing arguments prior to viewing any case information. Prior psychology literature suggests that when decision makers are required to write supporting and opposing arguments, a primacy effect results (Hoch, 1984).

Interaction effects between levels of audit experience and the inoculation treatment and between framing and inoculation are expected. It is hypothesised that participants exposed to an inoculation treatment will make judgments that are significantly different from the judgments of participants in the control group at three levels of audit experience. It is also hypothesised that the order by which auditors write supporting and

opposing arguments will have an effect on their judgments. To test these hypotheses, an experiment using an internal control evaluation task was carried out. Participants exposed to an inoculation treatment were required to write supporting and opposing arguments before viewing any case materials. They then reviewed the case materials and evaluated the internal controls of a company in two different frames. Half the participants evaluated the internal controls framed as strength of internal controls, and the other half of the participants evaluated the internal controls framed as risk of internal controls. The control participants did not write supporting and opposing arguments. I find a significant interaction effect between the inoculation treatment and audit experience. Consistent with prior psychology literature (e.g., Hoch 1984), the results support the hypothesis of a primacy effect. However, I do not find an interaction effect between framing and inoculation.

INTERNAL CONTROL EVALUATION

Various proxies have been used to measure judgment quality such as accuracy, consistency of judgment and consensus. Accuracy, although desirable, is not always determinable in certain audit tasks until some subsequent event occurs. One such audit task is the evaluation of internal controls. If the auditor determines internal controls to be effective, the later discovery of material weaknesses will invalidate the initial judgment. On the other hand, if no such material weaknesses are discovered, the auditor assumes that his/her initial evaluation is accurate even though material weaknesses may have gone undetected.

Joyce (1976) used an internal control task in the accounts receivable area and required participants to estimate audit time for five classes of audit procedures. His results show

a lack of agreement among auditors, which was attributed to the nature of the task. Higher agreement is found in tasks involving the rating of internal controls (Ashton, 1974) than in tasks requiring the estimation of audit work required or audit time (Joyce, 1976). Gaumnitz et al. (1982) modified Joyce's task by requiring the participants to evaluate the quality of internal controls over accounts receivable instead of estimating required audit time²⁰. Their results indicate that agreement is positively related to experience. Trotman and Wood (1991) analysed 16 internal control studies using meta-analysis techniques and conclude that there is agreement among auditors in the evaluation of internal controls and that 89% of the different results reported in the literature is due to sampling error.²¹ So a high level of consensus appears to exist among experienced auditors in the evaluation of internal controls.

FRAMING EFFECT

A framing effect suggests that a judgment could be affected by the wording of the problem. Numerous studies have been conducted in psychology to test this theory, most notable are the Asian-disease experiments carried out by Tversky and Kahneman (1981). In these experiments, the solutions offered by Tversky and Kahneman were actually identical except that the solutions were framed differently. One program emphasised gains while the other emphasised losses. Participants' judgments appeared to follow a similar pattern, i.e., "choices involving gains were often risk averse and choices involving losses were often risk taking" (Tversky and Kahneman, 1981, 453). Similarly, Sullivan (1996) carried out a series of experiments that required corporate

²⁰ This is because consensus is higher in the evaluation of internal controls than in estimating audit hours.

²¹ Sampling error is the difference between the sample statistic and the population parameter caused by chance sampling fluctuations (Harrison and Tamaschke, 1984).

financial managers to make investment judgments. The solutions offered to these managers were in the same format as Tversky and Kahneman (1981). In making their judgments, managers followed the same pattern as Tversky and Kahneman's participants, i.e., the managers acted to avoid risk when the choices involved gains and were risk-taking when the choices involved losses. Tversky and Kahneman (1981) proposed prospect theory to explain this perceived anomaly. In a risky situation, rational decision-makers would be risk-averse to maximise their expected utility, and in loss situations, decision-makers would be risk-taking to maximise their expected utility. Thus, the value function, when graphed, takes on an S-shape. If prospect theory holds true, then the manner in which judgment problems are framed determines the nature of the judgments solicited. Sullivan (1996) demonstrated that the manner by which an investment proposal is framed, determines whether it gets the "go-ahead". Similar results are obtained by Bazerman (1984). He required participants to make corporate down-sizing judgments. When the judgment problem was framed as gains (in number of jobs saved), participants' choices were more risk-avoiding. Conversely, when the judgment problem was framed as losses (in number of jobs lost), participants' choices were more risk-taking even though the choice alternatives offered were identical.

Prior auditing studies on framing examined the effects of framing on going concern evaluation (Kida, 1984; Trotman and Sng, 1989), internal control evaluation (Emby, 1994), and the nature of the hypothesis frame (whether error or nonerror frame) (McMillan and White, 1993). Kida (1984), using a going concern task, divided his participants into two groups. One group received a hypothesis framed as viability and the other group received a hypothesis framed as failure. Participants were given 20 information items - half of them indicative of failure and the other half viability, and

asked to list the information cues that were relevant to their judgment. Participants in the viability frame were expected to rate the viable items as more relevant than the failure items. The opposite was expected of participants who received the failure frame. The results, which are in the hypothesised direction, show that framing has an effect on the participants' rating of the relevance of information items.

In an extension to Kida (1984), Trotman and Sng (1989) examined the effects of hypothesis framing, cue diagnosticity and prior expectations on going concern evaluation. It was expected that participants who received strong ratios as prior expectations would be affected by framing, while participants who received weak ratios would not. The results show that when the prior information indicated failure, the relative number of failure and viable items considered were not affected by framing. However, when the prior information suggested viability, framing had an effect on the number of failure and viable items considered by the auditors.

Emby (1994) required participants to make substantive testing judgments based on internal controls framed as risk and strength. The case materials were presented either simultaneously or sequentially. The substantive tests prescribed by participants in the risk/sequential condition were significantly different from both the strength conditions and marginally significantly different from the risk/simultaneous condition. In addition, participants in the strength/sequential condition prescribed substantive tests that were significantly different from both the risk conditions, and marginally significantly different from the strength/simultaneous condition.

Instead of having auditors respond to preset frames as did Kida (1984) and Trotman and Sng (1989), McMillan and White (1993) had their participants frame their own hypothesis - either nonerror or error - to explain the source of fluctuations in the current year's unaudited financial report ratios. The nonerror frame suggested that environmental factors such as the economy, contributed to the fluctuations, while the error frame suggested material errors in the financial report. Auditors who preferred the error frame considered both confirming and disconfirming evidence, while those in the nonerror frame did not. In addition, McMillan and White found that the auditors' information search process is conservative regardless of the frame being tested, and confirmation bias reduces the effects of conservatism for auditors in the nonerror frame.

In investigating whether framing would affect the presentation order of information, Asare (1992) carried out an experiment which tested two frames (viability and failure) and two information orders (positive/negative and negative/positive) in a going concern task. Using the contrast assumption, framing was not expected to be affected by the presentation order of information, because the auditor would focus on contrasting information regardless of its presentation order, and his results support this argument.

Internal control evaluation tasks may be framed in one of two ways (Emby, 1994). Since the S-shaped value function of prospect theory is steeper for losses than for gains, this leads decision makers to emphasise losses more than gains (Emby, 1994). Translated to an internal control evaluation task, auditors who evaluate internal controls framed as a risk evaluation task would emphasise the negative aspects whereas those who evaluate internal controls framed as a strength evaluation task would not emphasise strengths. Consequently, significant differences between the frames would be observed.

ORDER EFFECTS

The psychology literature suggests that the order in which information is presented affects the resultant judgment (Hogarth and Einhorn, 1992). A primacy effect occurs when the information considered first is given more importance in the final judgment²², and a recency effect occurs when the information considered last has a greater bearing on the final judgment. If the order of the presentation of the information is changed, the judgment will also change. There is disagreement in the psychology literature regarding which order effect is more dominant. While the psychology literature traditionally finds a primacy effect, some studies have found a recency effect (e.g., Nisbett and Ross, 1980).

Two treatments of order are observed in the literature. The first manipulates the presentation order of information and for most audit tasks, this results in a recency effect (e.g., Asare, 1992; Ashton and Ashton, 1990 and 1988). The second holds the presentation order constant and manipulates the order of writing supporting and opposing arguments (e.g., Koonce, 1992; Hoch, 1984), and mixed results are observed (Koonce, 1992; Hoch, 1984). In a series of three experiments, Hoch (1984) manipulated hypothesis-generation tasks such as purchase of a video cassette player, will Britain go to war over the Falklands, and will OPEC impose an oil embargo. After viewing the information, participants wrote reasons to explain whether these events would (pro) or would not (con) occur. Half of the participants explained in the pro/con order and the order was reversed for the other half. The results show that writing the first set of

reasons interferes with participants' ability to write the second set of reasons. In all three experiments, a primacy effect is observed. Koonce similarly manipulated the order in which supporting and opposing arguments were written after participants had viewed the information, and a primacy effect was hypothesised. Instead, her results show a recency effect and she attributes this to task difficulty. As the participants wrote supporting/opposing arguments that were experimenter-prescribed, this may have affected the results. This is because the participants may have suspected that the experimenter was pointing them in the right direction, and therefore, a recency effect was observed.

In sum, merely controlling the presentation order of information and requiring auditors to make judgments based on their evaluation of the information should lead to a recency effect. On the other hand, requiring auditors to write supporting and opposing arguments in a specific order should lead to a primacy effect because the writing of the first set of arguments interferes with the auditors' ability to consider the second set of arguments.

THEORY DEVELOPMENT

Framing effects have been observed in both the social psychology and accounting literature. Prospect theory, proposed by Tversky and Kahneman (1981) to explain framing effects, states that decision makers act in a manner that maximises their expected utility. Consequently, decision makers would be risk averse when the judgment problem involves gains or is positively framed, and would be risk-taking

²² However, one would assume the more diagnostic information would be considered important

when the judgment problem involves losses or is negatively framed. Significant differences in internal control evaluation are expected between internal controls framed as risk and strength, as the S-shaped value function of prospect theory is steeper for losses than for gains. Participants in the risk frame are expected to emphasise internal control weaknesses, whereas participants in the strength frame are not expected to emphasise strengths. Inoculation is expected to correct the effects of framing of internal controls. Prior to examining any information, participants are required to consider all possible solutions to a problem, including all aspects of an internal control system. Consequently, the effect of framing is mitigated because all possible solutions are encoded in memory and are equally available. In addition, the real nature of the task becomes more salient to the decision maker, and thus reduces the effect of the frame. Therefore, an interaction effect between framing and inoculation is expected. Framing effect for the control group and no framing effect for the treatment group will be observed, so that the difference between the two frames in the treatment group is reduced. The interaction hypothesis states that:

H₁ There will be an interaction effect between framing and inoculation on the evaluation of internal controls.

Fischhoff (1982) discusses the concept of a perfectible decision maker. S/he is someone whose biased judgments could be corrected by relevant training. Fischhoff (1982) offers five strategies for restructuring the judgments of perfectible decision makers. First, perfectible decision makers may be forced to make judgments on what they "explicitly" know rather than what is available "in the head". Second, they should be encouraged to look for opposing evidence as opposed to searching for only supporting evidence. Third,

regardless of the order of its presentations.

ways of decomposing a complex task should be offered to decision makers, and fourth, they should learn to link the present problem to related situations they have encountered in the past. Last, they should consider possible alternative solutions to the present problem. Perfectible decision makers should also be educated on developing problem-solving skills (Fischhoff, 1982).

Perfectible decision makers' responsiveness to correction is related to their having acquired the necessary schemata. Ashton and Ashton (1988) suggest that auditors' perceptions, accumulation and interpretation of audit evidence are developed through experience and previously acquired knowledge. This indicates that the acquisition of knowledge and the development of auditors' memory structures are gradual rather than immediate. To explain this, Waller and Felix (1984, 386) developed a framework of learning from experience which was premised on the assumption that "learning from experience involves the formation and development of cognitive structures, which organise declarative and procedural knowledge in long-term memory, through interaction with new information generated by experiential action and/or observation". They called this "schemata".

Choo and Trotman (1991) make five observations of schemata which distinguish novice from experienced auditors. First, schemata enable experienced auditors to group, classify and order information, and second, experienced auditors are able to perform this in a more comprehensible manner. Third, the schemata determine what information should be stored in and accessed from memory. Fourth, the nature of the information retrieved from the schemata influences the judgment outcomes, and fifth, the effectiveness of the judgment is related to the amount of information retrieved from

memory. This being the case, it is argued that auditors' responsiveness to a judgment technique is dependent upon the availability of these memory structures, and hence, upon their experience. Novice auditors may not have developed the memory structures to respond adequately. They are perfectible decision makers except they do not have developed schemata and are unable to respond to the learning conditions created by the technique. Domain-specific knowledge and skills have not been encoded and stored in long-term memory so that a search of memory is futile. In addition, they have not acquired the knowledge and skills necessary to support their judgments (Gibbins, 1984). Consequently, their judgments remain relatively unaffected. However, while they are not responsive to judgment tools, they are responsive to education, training and experience. Decision makers with some audit experience may not have fully-developed schemata, but their schemata are sufficiently developed to enable them to respond favourably to training. However, not all experienced auditors are perfectible. Very experienced auditors who have developed sophisticated, effective and stable memory structures may not be responsive to training. In addition, these auditors may have developed "specialised heuristics" which appear to serve them well. For example, very experienced auditors have developed a conservatism heuristic to reduce the substantial risks associated with many audit judgments (Smith and Kida, 1991). Because of these factors, they may not be receptive to training.

This study postulates that the relationship between audit experience and responsiveness to training is an inverted-U shaped value function. To test this, three levels of audit experience are examined (no audit experience, one to 18 months, and 19 to 60 months). The objective of using three groups is to capture the responsiveness of audit experience to a judgment tool. Eighteen months audit experience is used as a cutoff because, from

personal experience and anecdotal evidence, at this stage, many auditors are put in charge of small audits where they would be required to evaluate internal controls. By this time, their schemata (in relation to internal controls) should be more developed. Participants with no audit experience are not expected to respond to the inoculation treatment as they do not have the necessary schemata. Participants with one to 18 months are expected to be most responsive to the inoculation treatment. While their judgments are not yet perfected, they have acquired sufficient knowledge and skills to recognise this and, as their schemata are not fully developed, they are expected to respond to the judgment tool. The most experienced (19 to 60 months) are not expected to respond strongly to the treatment; this is because they have well developed schemata that are unlikely to be affected by a simple judgment tool such as inoculation. Based on this discussion, an interaction effect between inoculation and audit experience on the evaluation of internal controls is expected, and the second hypothesis states:

- H₂ There will be an interaction effect between inoculation and audit experience on the evaluation of internal controls.
- H_{2a} For the participants without audit experience, there will be no significant difference in the evaluation of internal controls between participants given the inoculation treatment and participants not given the inoculation treatment.
- H_{2b} For the participant with one to 18 months audit experience, there will be a significant difference in the evaluation of internal controls between participants given the inoculation treatment and participants not given the inoculation treatment.
- H_{2c} For the participants with 19 to 60 months audit experience, there will be no significant difference in the evaluation of internal controls between participants given the inoculation treatment and participants not given the inoculation treatment.

Prior auditing studies manipulate the order of presentation of information and find a recency effect (e.g., Asare, 1992; Ashton and Ashton, 1988). When the experimental

treatment of order effect is changed from manipulating presentation order of information to requiring participants to write supporting and opposing arguments before making a judgment (i.e., an inoculation treatment), Hoch (1984) predicted a primacy effect. Inoculation requires auditors to integrate and generate information from memory prior to information search in order to explain both (or all) possible solutions to a problem. Hoch (1984) proposes that the provision of supporting (opposing) arguments before making a judgment would affect decision makers' likelihood judgments of the target event, i.e., an explanation effect is observed. This is because the provision of the first set of arguments interferes with decision makers' ability to generate the second (or subsequent) set(s) of arguments (Roediger, 1978). Roediger labels this "output interference". The ability to recall is also affected. More items are recalled from the first set of arguments than from the second (Roediger, 1978). Hoch (1984) suggests that when generating the second set of arguments, the decision maker is unable to forget the first set of arguments. His results show that the first set of arguments is also easier to produce than the second set, and a primacy effect is observed. Apart from Koonce (1992) who found a recency effect, the findings of Hoch (1984) and Roediger (1978) are as yet untested in an audit setting. Similarly, the inoculation treatment is untested in an audit setting.

Hoch (1984) suggests that another explanation for the finding of primacy when the order of writing supporting and opposing arguments is changed, is the theory of anchoring and adjustment (reported by Tversky and Kahneman [1981]). People make judgments by starting from an initial value and adjustments are made to this value as more information is processed. The adjustment continues until a judgment is reached or the information has been exhausted. Prior research finds that this adjustment is usually

inadequate. When auditors are required to write supporting (opposing) arguments, these form their anchor and while they may adjust their anchor when they write the opposing (supporting) arguments, this adjustment is normally insufficient. When auditors subsequently review the data set, they are more aware of the first set of arguments they have written and this makes them more receptive to the cues supporting this argument. Consequently, they encode more cues supporting this position and a primacy effect is observed. In the current task, I expect that auditors who write arguments in the positive/negative order to evaluate internal controls as more effective than auditors who write arguments in the negative/positive order, as the order of writing supporting and opposing arguments affects the resultant judgment.

H₃ The inoculation treatment will result in a primacy effect.

THE EXPERIMENT

Participants

The participants were 139 accountants attending a national training program to prepare them for the auditing section of the Institute of Chartered Accountants' (ICA) Professional Year examinations in Australia. At the time of testing, all participants had completed their first university degree and had work experience with a chartered accounting firm in the areas of auditing, business services, or taxation. The mean work experience was 31 months (minimum six months and maximum 152 months)²³. Participants' audit experience ranged from no experience to 60 months of audit experience (mean 14 months). For those participants with audit experience, the mean

²³ Two participants had less than one year work experience (six and eight months). The normal minimum work experience required to sit the Professional Year examinations is one year. However, the ICA

was 20 months (sd 11 months). The average age of participants was 23.5 years. Participation was on a voluntary basis and the participants received A\$10 for their effort. The experiment was carried out in four groups of 20 to 51 participants.

permits accountants with less than one year work experience to sit their Professional Year examinations provided they have completed an Honours degree.

Case Materials and Administration

Figure 3.1 provides an overview of the experiment.

Figure 3.1 - Experimental Procedures

Step 1	Read preliminary information			
	Inoculation (at three levels of audit experience)		Control (at three levels of audit experience)	
	Risk frame	Strength frame	Risk frame	Strength frame
Step 2	Write supporting/ opposing (opposing/ supporting arguments)	Write supporting/ opposing (opposing/ supporting arguments)	(Do not write arguments)	(Do not write arguments)
Step 3	Evaluate case materials	Evaluate case materials	Evaluate case materials	Evaluate case materials
Step 4	Rate internal controls	Rate internal controls	Rate internal controls	Rate internal controls
Step 5	Complete post- test questionnaire	Complete post- test questionnaire	Complete post- test questionnaire	Complete post- test questionnaire

The experiment tested three levels of audit experience (comprising no audit experience, one to 18 months of audit experience, and 19 to 60 months of audit experience), two frames (risk and strength) and an inoculation treatment group and a control group in a 3 x 2 x 2 full factorial design (Figure 3.1). The case materials were based on the case

developed by Emby (1994) and involved the evaluation of internal controls in the inventory area (Appendix A). The only modifications made were to the name of the client and other proper names to reflect an Australian context. Each participant was given a copy of the experimental materials in a booklet. The cover page contained an assurance of the confidentiality of their responses. Some introductory information was given. Participants were told that the purpose of the experiment was to understand how auditors formulate judgments when evaluating internal controls, and as part of the study participants would be required to evaluate the internal control system. In addition, they were told that there is no right or wrong answer and they were to provide their best judgments. The first variable manipulated was inoculation. Half of the participants received the inoculation treatment (inoculation condition) and the other half acted as control (control condition). To capture framing effects, half the participants received internal controls framed as "strength" (strength frame) and the other half received internal controls framed as "risk" (risk frame). The "inoculation" participants in the risk frame were told that:

Part of an auditor's task is to evaluate the risk of the internal control system in the inventory area. Please write down in the spaces provided below, the situations you think may result in low or high risk in the internal control system. Write down as many situations as you can think of. Phrases and short sentences would be sufficient.

Situations resulting in a low risk internal control system in the inventory area:

Situations resulting in a high risk internal control system in the inventory area:

Half of the "inoculation" participants in the risk frame wrote internal control situations in the low/high risk order and the order was switched for the other half. The "inoculation" participants in the strength frame were told that:

Part of an auditor's task is to evaluate the strength of the internal control system in the inventory area. Please write down in the spaces provided below, the situations you think may result in a weak or strong internal control system ...

Situations resulting in a weak internal control system in the inventory area:

Situations resulting in a strong internal control system in the inventory area:

Half of the "inoculation" participants in the strength frame wrote the situations in the weak/strong order, and the other half in the strong/weak order. Participants in the control group (both strength and risk frames) did not write supporting or opposing arguments.

Following this, the case materials were provided. The case materials comprised a general history and background of the client, a brief description of the inventory system and seven internal control attributes comprising both high and low risk features. The order of these information items was randomised for each participant to control for version effects. After reading the case information, participants evaluated the internal controls. They then completed a post-test questionnaire which solicited some personal data as well as information on participants' evaluation of various aspects of the task. Upon completion, participants were paid and debriefed.

Dependent Variable

Since the evaluation of internal controls may be framed in one of two ways, participants in the risk frame were asked to assess the risk of the internal controls and they recorded their evaluation of the internal controls on a seven-point Likert-type scale anchored by

“low” (1) and “high” (7). Participants in the strength frame recorded their evaluation of the internal controls on a seven-point Likert-type scale anchored by “weak” (1) and “strong” (7)²⁴.

Audit Experience

Consistent with prior studies, the number of months of general audit experience is used as a measure of experience (e.g., Ho, 1994; Choo and Trotman, 1991). This is especially appropriate in the evaluation of internal controls as past experience and anecdotal evidence suggest that this task is mostly assigned to less experienced auditors. My study divided participants into three groups - Group 1 (no audit experience), Group 2 (one to 18 months of audit experience, mean 12 months), and Group 3 (19 to 60 months of audit experience, mean 25 months).

RESULTS AND DISCUSSION

Diagnostic Checks

ANOVA assumes constant variance of all cells. Consequently, Bartlett’s Box F tests and Cochran’s C test are performed for the analyses to test this assumption. In all tests, the homogeneity of variance assumption is not violated.

Checks on participants in the inoculation condition reveal a significant task difficulty effect ($F = 8.27$, $p = .005$ [two-tailed]) (Panel A, Table 3.1). Participants in the inoculation condition found the combination of the tasks of writing supporting and opposing arguments and evaluating the internal controls significantly more difficult than participants in the control condition who merely evaluated the internal controls. Their

²⁴ These scores were reversed for the data analyses.

task difficulty scores are 4.16 (sd 1.70) and 3.20 (sd 1.59) respectively (Panel B, Table 3.1).

In addition, audit experience has a significant effect on task difficulty ($F = 10.87, p = .000$ [two-tailed]) (Panel A, Table 3.1). The task difficulty scores on a nine-point Likert-type scale for the three groups are 4.56 (Group 1), 3.76 (Group 2), and 2.24 (Group 3) (Panel B, Table 3.1). Group 1 (no experience) participants found the task significantly more difficult than participants in Group 2 (one to 18 months experience) and Group 3 participants (19 to 60 months experience). This, of course, is expected because Group 1 participants did not have the relevant domain-specific knowledge to complete the task, whereas the other two groups had. Frame has no significant effect on task difficulty ($F = <1, p >.1$ [two-tailed]) (Panel A, Table 3.1).

Further analyses show no significant difference between the evaluation of internal controls between all the participants of the current study and Emby's (1994) participants²⁵.

²⁵ The overall IC score mean of the present study is 3.43 and Emby's is 3.23.

Table 3.1 - The Effects of Inoculation, Framing and Audit Experience on Task Difficulty

Panel A Source	SS	DF	MS	F	p
Main effects					
Inoculation (I)	19.76	1	19.76	8.27	.005
Frame (F)	.75	1	.75	.32	ns
Audit experience (E)	51.92	2	25.96	10.87	.000
2-way interaction					
I x F	2.77	1	2.77	1.16	ns
I x E	.35	2	.17	.07	ns
F x E	2.20	2	1.10	.42	ns
3-way interaction					
I x F x E	5.67	2	2.83	1.19	ns
Residual	296.25	124	2.39		
Panel B					
Inoculation					
Treatment condition (n = 64)				4.16 (1.70)	
Control condition (n = 72)				3.20 (1.59)	
Audit Experience					
No audit experience (n = 41)				4.56 (1.91)	
1 to 18 months (n = 38)				3.76 (1.63)	
19 to 60 months (n = 60)				2.24 (1.24)	
Frame					
Risk frame				3.74 (1.67)	
Strength frame				3.57 (1.76)	

Task difficulty scores: 1 = "not difficult at all", and 9 = "very difficult".

Analyses and Discussion of H₁

An interaction effect between framing and inoculation is predicted, the inoculation treatment will mitigate the effects of framing. The mean internal control (IC) scores²⁶ are 3.41 (sd 1.18) (risk frame) and 3.45 (sd 1.11) (strength frame) (Table 3.2). ANOVA results²⁷ show no significant main framing effect ($F < 1$, $p > .1$) (Table 3.3)²⁸. Participants in both frames made similar evaluations of the internal controls. A significant main effect for inoculation is observed ($F = 3.76$, $p = .024$) (Table 3.3). The mean IC scores are 3.26 (sd 1.07) and 3.58 (sd 1.19) (Table 3.2). Participants in the treatment condition evaluated the internal controls as significantly less effective than participants in the control condition. This supports the stated purposes of inoculation which are to make the decision maker more aware of opposing arguments and exercise caution in decision-making. Hence, participants in the treatment condition made more conservative judgments. No interaction effect between framing and inoculation is observed and H₁ is not supported. The insignificant result could be caused by, first, "experience preference" (Frisch 1993). In explaining this, Frisch distinguishes between "decision preference" and "experience preference". Decision preference refers to the decision maker's choice, i.e., the frame, while experience preference refers to the frame the decision maker actually experienced as preferable during the judgment process. An auditor may commence the judgment process in a particular frame, e.g., risk, but in the process of making the judgment, the auditor may find the strength frame to be more

²⁶ The participants in the risk frame recorded their evaluation of the internal controls on a scale anchored by low (1) and high (7) and the participants in the strength frame on a scale anchored by weak (1) and strong (7). Because of this, the raw scores for these two frames are not compatible. Consequently, the scores of the participants in the risk frame are reversed to match those in the strength frame.

²⁷ All tests of significance are one-tailed unless stated otherwise.

²⁸ Emby (1994) found a framing effect when he asked participants to prescribe substantive tests based on their evaluation of internal controls. In this experiment, participants were not asked to prescribe substantive tests because the decision to increase or decrease substantive testing is based on other

relevant, i.e., experience preference may reduce the effects of framing. Second, while the auditor may commence with a particular frame, during the judgment process, the framing effect may diminish as the true purpose of the task (that is, the evaluation of internal controls regardless of the frame) becomes more dominant in the auditor's working memory. These reasons may account for the lack of framing effect, and hence, the lack of an interaction effect between framing and inoculation.

Table 3.2
The Effects of Inoculation and Framing on the Evaluation of Internal Controls - Mean (sd)

Treatment	<u>Frame</u>		Overall
	Risk	Strength	
Inoculation	3.29 (1.06) n = 34	3.23 (1.10) n = 33	3.26 (1.07) n = 67
Control	3.51 (1.30) n = 36	3.65 (1.09) n = 36	3.58 (1.19) n = 72
Overall	3.41 (1.18) n = 70	3.45 (1.11) n = 69	3.43 (1.14) n = 139

1 = weak internal controls (high risk), 7 = strong internal controls (low risk).

factors besides the evaluation of internal controls. Consequently, the analyses of H₁ was based on the participants' evaluation of internal controls.

Table 3.3
ANOVA Results: The Effects of Inoculation, Framing and Audit Experience on
Evaluation of Internal Controls

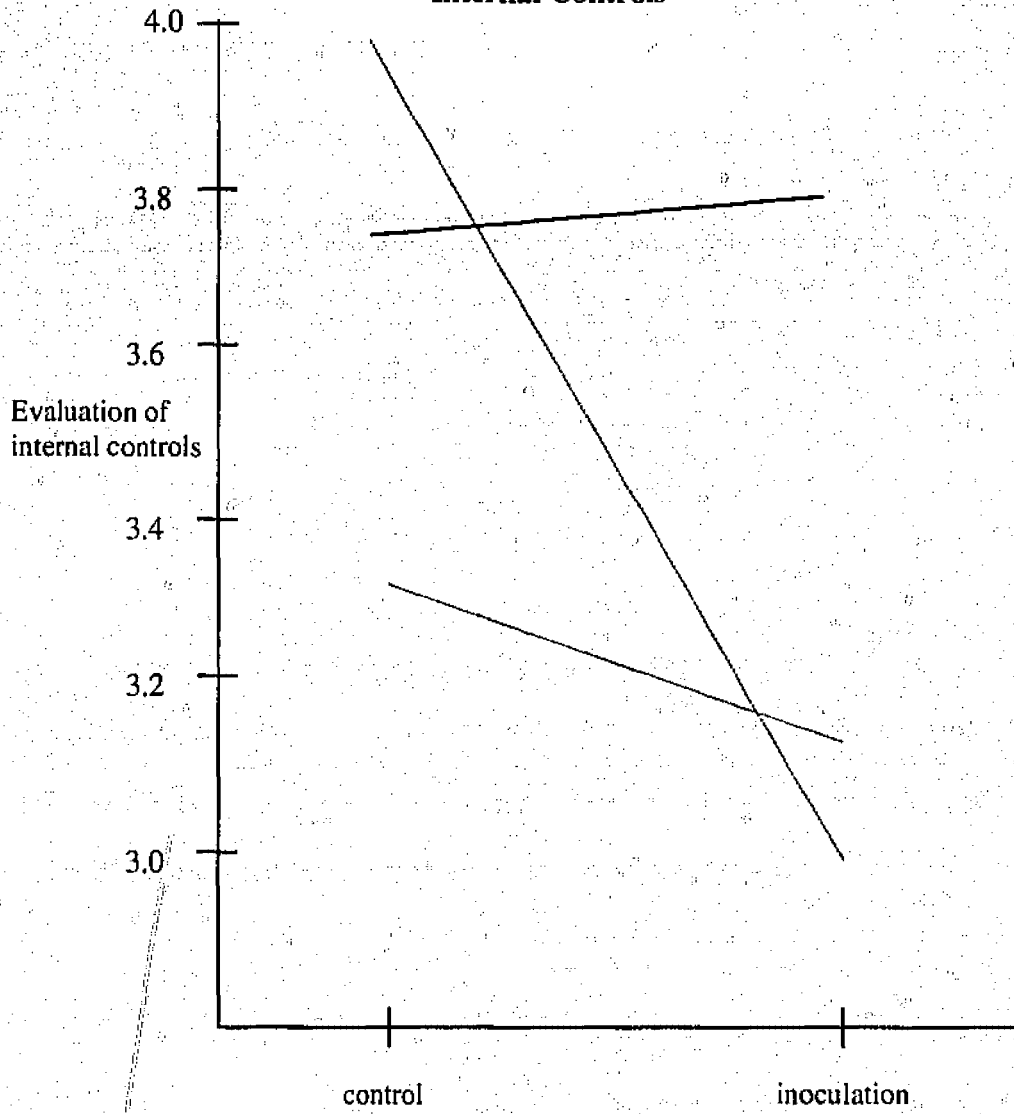
<u>Source</u>	<u>SS</u>	<u>DF</u>	<u>MS</u>	<u>F</u>	<u>p</u>
<u>Main effects</u>					
Inoculation (I)	5.07	1	5.07	3.96	.024
Framing (F)	.03	1	.03	.02	ns
Audit experience (E)	6.89	2	3.44	2.69	.036
<u>2-way interaction</u>					
I x F	.00	1	.00	.00	ns
I x E	4.95	2	2.48	1.94	.074
F x E	1.06	2	.53	.41	ns
<u>3-way interaction</u>					
I x F x E	.63	2	.32	.25	ns
Residual	162.53	127	1.28		

ns = not significant

Analyses and Discussion of H₂

This second hypothesis predicts that the inoculation treatment would result in significant differences in the evaluation of internal controls by participants with audit experience and those without audit experience. Inoculation works by sensitising decision makers to the presence of opposing arguments (Anderson, 1982). It is expected that participants with one to 18 months audit experience would be most affected by the inoculation treatment, whereas the inoculation treatment would not have a significant effect on participants without audit experience because they have not developed the necessary schemata or acquired the relevant domain-specific knowledge. It is not expected to have an effect on participants with 19 to 60 months' experience because their more sophisticated and developed schemata are less susceptible to change.

Figure 3.2
The Effects of Audit Experience and Inoculation on the Evaluation of Internal Controls



_____ 19 to 60 months' experience
_____ 1 to 18 months' experience
_____ no experience

Figure 3.2 shows the interaction effects, and the cell means are shown in Table 3.4. The mean internal control scores (IC scores) are 3.78 (Group 1), 3.35 (Group 2), and 3.25 (Group 3) (Table 3.4), and ANOVA results indicate that these IC scores are significantly different across the three experience levels ($F = 2.69$, $p = .036$) (Table 3.3). The omnibus F-test reveals a marginally significant interaction effect between audit experience and inoculation ($F = 1.94$, $p = .074$) (Table 3.2). Planned comparisons are carried out to test the hypothesis. Inoculation had a significant effect on Group 2 participants ($t = 1.94$, $p = .027$) and no significant effect on Groups 1 and 3 ($t = 1.42$, $p = .157$ [Group 1]; $t = .64$, $p = .267$ [Group 3]), and therefore, H_2 and the alternative hypotheses are supported.

Table 3.4
The Effects of Inoculation and Audit Experience on the Evaluation of Internal Controls (Mean [sd])

Treatment	Group 1	Level of Audit Experience		Overall
		Group 2	Group 3	
Inoculation	3.79 (1.13) n = 19	2.98 (0.98) n = 24	3.13 (0.99) n = 24	3.26 (1.07) n = 67
Control	3.77 (1.02) n = 22	3.93 (1.07) n = 14	3.33 (1.30) n = 36	3.58 (1.19) n = 72
Overall	3.78 (1.06) n = 41	3.33 (1.10) n = 38	3.25 (1.18) n = 60	3.43 (1.14) n = 139

1 = weak internal controls (high risk), 7 = strong internal controls (low risk).

Group 1 participants' mean IC scores are 3.79 (inoculation) and 3.77 (control) and these are not significantly different. Given that the purpose of the inoculation treatment was to sensitise participants to the existence of opposing arguments, this treatment apparently did not have an effect on novice auditors. One explanation is that they may not have

developed the schemata necessary to make domain-specific judgments (Libby and Luft 1993; Choo and Trotman 1991; Waller and Felix 1984). Because of the limitations of their memory structures, they were unable to respond adequately to this treatment. Their mean IC scores were significantly less conservative compared to the other two groups, as these participants had not acquired the conservatism heuristics common to experienced auditors (Smith and Kida, 1991).

A large difference in the mean IC scores of Group 2 is noted (control - 3.93, inoculation - 2.98), and this difference is significant. The inoculation treatment had a significant effect on these participants. Group 2 participants in the inoculation treatment evaluated the internal control as significantly less effective than participants in the control condition. These participants appeared to be the most responsive to the inoculation treatment. One explanation for this is that, while they may have had more developed memory structures compared to Group 1 participants, these structures still lagged behind the more experienced auditors. However, they had acquired sufficient knowledge to make them responsive to the learning conditions created by the inoculation treatment. Fischhoff (1982) refers to them as "perfectible judges". The mean IC scores of the treatment participants were not significantly different from Group 3 participants. Inoculation resulted in more conservative judgments among Group 2 participants.

The mean IC scores of Group 3 participants are 3.13 (inoculation) and 3.33 (control), and these are not significantly different. This result shows that Group 3 participants were also not responsive to the inoculation treatment, but their IC scores were significantly more conservative than participants in Group 1 ($t = 2.35, p = .010$). These results are consistent with the findings reported in Smith and Kida (1991). They report

findings of conservatism in experienced auditors' judgments from a wide range of studies. Generally, experienced auditors are less affected by the general heuristics and biases reported in the psychology literature, as the nature of audit training and experience leads to the development of "specialised" heuristics "that prove highly effective for tasks within their domain of expertise" (486). As Group 3 participants in the control condition would have considered both types of arguments regardless of whether formally required to do so or not, the inoculation treatment did not result in any significant difference in their judgments. In sum, the above analyses show that inoculation was effective on Group 2 participants but was not effective on Groups 1 and 3 participants. The results also show that the relation between audit experience and responsiveness to inoculation is nonlinear and has an inverted-U-shaped value function.

Analyses and Discussion of H₃

In the inoculation treatment, participants were required to write arguments in a positive/negative order²⁹ or a negative/positive order³⁰. As a result of this manipulation, H₃ predicts a primacy effect, that is, participants who wrote arguments in the positive/negative order would evaluate the internal controls as being more effective than participants who wrote arguments in the negative/positive order. The results show this to be the case. The mean IC scores are 3.49 (positive/negative) and 3.03 (negative/positive) respectively (Table 3.5), and the difference between these two orders is significant ($t = 1.8, p = .038$) (Table 3.5). Thus, H₃ is supported. The possibility that these results may be caused by task difficulty is ruled out. The t-results of the effect of

²⁹ The positive/negative order comprises participants who wrote arguments in the low/high order in the risk frame and the strong/weak order in the strength frame.

³⁰ The negative/positive order comprises participants who wrote arguments in the high/low order in the risk frame and the weak/strong order in the strength frame.

order on task difficulty is not significant ($t = .60, p = .55$ [two-tailed]). Similarly, order has no effect on effort expended ($t = .54, p = .59$ [two-tailed])³¹. The results show that the order by which auditors write positive/negative arguments results in a primacy effect, and supports prior psychology studies (Hoch 1984; Roediger 1978), but contradicts Koonce's (1992) results.

Table 3.5 - Order Effects

<u>Source of Variation</u>	<u>Mean (sd)</u>	<u>t</u>	<u>p</u>
Positive/negative (n = 33)	3.49 (.97)	1.80	.038
Negative/positive (n = 34)	3.10 (1.11)		

CONCLUSIONS

The experiment tested the effects of framing, inoculation and audit experience on internal control evaluation. H_1 predicts an interaction effect between framing and inoculation. However, no significant effect is found, and H_1 is not supported. The participants were not affected by the framing of internal control evaluations as risk and strength.

The second hypothesis examined whether audit experience has an effect on participants' responsiveness to inoculation. To test this hypothesis, I manipulated three groups of participants, each group having a different level of audit experience, and inoculation. I

³¹ ANOVA results with task difficulty as a covariate are similarly not significant ($F = 1.85, p = .18$ [two-tailed]).

find support that audit experience affects participants' responsiveness to inoculation and that the relation between audit experience and responsiveness to inoculation is nonlinear. Auditors with one to 18 months' experience were most responsive to training in internal control evaluation. Therefore, training resources should be targeted at this group of audit staff. Unlike participants in Group 2, participants in Group 1 and Group 3 did not respond to the inoculation treatment. Group 3 participants evaluated the internal controls as significantly less effective than Group 1 participants. Group 2 participants in the inoculation condition similarly made more conservative evaluations of the internal controls compared to those in the control condition. This suggests that inoculation is most effective with less experienced auditors as they have yet to develop the skills and expertise of experienced auditors. The inoculation treatment did not affect the judgments of participants without audit experience because they did not have the relevant schemata to react to the treatment.

Finally, I found that the evaluation of internal controls was affected by the order in which participants wrote supporting and opposing arguments. Participants who wrote arguments in the positive/negative order evaluated the internal controls as more effective than participants who wrote arguments in the negative/positive order. This study shows that requiring auditors to write supporting and opposing arguments results in a primacy effect instead of the normally observed recency effect when presentation order of information is manipulated.

LIMITATIONS

There are limitations to the experiments reported here. First, the case materials employed in these experiments comprised relatively less information than what auditors

normally work with. However, it is unclear what effect(s) more detailed case materials would have on the judgment techniques examined and what limitations the use of simplified case materials would have on the external validity of the results of this study. Second, in trying to secure responses that the participants could justify, all participants in this study were paid. This procedure does not, however, correspond to the audit environment where auditors would suffer penalties for not providing justifiable responses. Such penalties include financial losses (in promotion and pay rises as well as litigation costs) and loss of reputation. Third, the participants in this experiment worked alone whereas in practice they often work in groups or have the opportunity to consult with group members. It is unclear what effects having participants work alone would have on the external validity of the results of this study.

CHAPTER 4

EXPERIMENT TWO - AN INVESTIGATION OF INFORMATION PROCESSING BY AUDIT GROUPS

INTRODUCTION

Prior judgment research draws attention to weaknesses in auditor judgment. For example, auditors' going-concern judgment lacks consensus³², and this is observed across a range of experience levels (Ho, 1994). Because of the weaknesses identified, there is a need for judgment research to not only identify and explain sub-optimum judgments but to investigate means of improving them³³. One approach to correcting sub-optimum judgments is the use of groups. Both accounting and psychology studies report that groups outperform individuals over a wide range of tasks.

In practice, auditors often work in teams, but audit groups are sometimes used as well. However, prior auditing research has largely examined individual auditor judgments to the exclusion of groups and teams. Consequently, after two decades of audit group studies, we know relatively little of the impact of audit group decision processes on audit judgments. One reason for this is that group studies require large numbers of participants, a factor that often constrains auditing researchers. The constraint of large numbers of participants is also evidenced by the research methodology adopted in prior studies. Such a research methodology may have confounded the results, as a learning effect occurs when participants make the same judgment first, as an individual, and second, as a member of a group (Iselin, 1991). Prior group studies that examine consensus and accuracy among audit groups and individual auditors have adopted this

³² That is, different judgments are reached based on the same information.

methodology (Trotman et al., 1983; Trotman and Yetton, 1985; Trotman, 1985). Because of this, there is a need for further auditing studies that control for this confounding, as well as a need to retest prior findings.

In this experiment, I compare the going-concern judgment of individual auditors with the going-concern judgments made in an interacting group setting. Interacting audit groups, instead of hierarchical teams, were used for two reasons. First, even though the formal review procedures of public accounting firms suggest the use of hierarchical teams, informal discussions between auditors (often of equal rank) do occur. Second, interacting audit groups encourage more uninhibited interaction and discussion compared to hierarchical teams. This is because the latter are dominated by the most experienced or most senior member, and hence, the opportunity for learning by novice auditors is reduced. During the group interaction, novice auditors are able to put forward their views and receive feedback from their peers. This provides them with not only insight into their own judgments, but also exposes them to the thought processes of other group members. Such interchange may not occur in hierarchical teams.

One purpose of this experiment is to compare the judgment processes of interacting groups and individuals. One method of improving the effectiveness of decision-making is to decompose the problem into smaller judgments (Messier, 1995). The auditor performs each sub-judgment separately, and assembles them into a global judgment (Messier, 1995). In a going-concern task, the sub-judgments comprise the going-concern problem indicators and the mitigating factors. The auditor selects the relevant cues,

³³ Not all attempts to achieve judgment consensus result in their stated aim. For example, providing auditors with non-statistical sampling guidelines leads to less consensus in auditors' recommended

assigns weights to them, and forms his/her global judgment. The judgment quality is, consequently, dependent upon the range of the auditor's cue usage, and his/her cue selection and weighting³⁴ (Stocks and Harrell, 1995; Ashton, 1992; Bonner, 1990; Chewning and Harrell, 1990; Chalos and Pickard, 1985). Prior research suggests that groups can perform these tasks more effectively, and therefore, the quality of their judgment can be improved (Stocks and Harrell, 1995; Chalos and Pickard, 1985; Hill, 1982; Howell et al., 1970). This study tests whether consensus is higher in audit groups' going-concern judgments compared to individual auditor judgments. In addition, information processing methods (comprising self-reported cue usage, selection and weighting) are compared to allow us to understand the cause(s) of the differences between audit groups' and individual auditors' performance.

The study found that, in addition to being more conservative, consensus was also higher among audit group compared to individual auditor judgments. Groups use a wider range of information cues in making their judgments. Differences in cue weighting are also observed, groups assigned higher weights to going-concern problem indicators but both groups and individuals did not weight mitigating factors differently. No significant difference in cue selection between groups and individuals is noted. Audit experience appears to have little impact on these results.

sample sizes (Kachelmeier and Messier, 1990).

³⁴ I use cue usage to describe the volume of cues processed by decision makers in this experiment (Chewning and Harrell 1990). Cue selection relates to the predictability/appropriateness of the cue of the decision on hand (Chalos and Pickard 1985) and cue weighting refers to the relative importance of the cue (Selling and Shank 1989).

GOING-CONCERN ASSESSMENT

Professional pronouncements require auditors to consider the appropriateness of the going-concern assumption for the preparation of the client's financial report and to issue a going-concern opinion when the viability of the client is uncertain. Standard-setters have increasingly set more stringent standards concerning going-concern evaluation (e.g., AUS 708 [AARF, 1996]; SAS No. 59 [AICPA, 1988]; SAS No. 34 [AICPA, 1981]). While this is a much researched topic in auditing, there is, however, limited literature on the effect of auditors' information processing on their going-concern judgments.

A going-concern evaluation task was used because it is one of the most difficult tasks an auditor faces. Training for such judgments should begin early in an auditor's career (McMillan, 1994; Dombrowski, 1993). The use of practical case studies is one way of providing novice auditors with on-the-job experience (Dombrowski, 1993). Such practical experience allows novice auditors to apply basic auditing principles in a practical manner and results in greater understanding and learning. It also provides insight into whether cue usage, selection and weighting have experience effects. A difficulty of making going-concern judgments is that few guidelines are available to auditors. While professional standard setters have attempted to identify going-concern problem indicators as well as mitigating factors (e.g., AUS 708 [AARF, 1996]), the final judgment ultimately rests with the auditor after consultation with colleagues in both interacting and non-interacting groups.

THEORY DEVELOPMENT

As explained in Chapter 2, various explanations have been used in the auditing literature to explain groups' superior performance over individuals'. These include best member theory and information load theory. Best member theory makes two assumptions. First, where standardised bias occurs, interaction among group members results in the identification of individual member expertise, and second, individual group members, having different levels of expertise, allow group members to follow their best member and use that member's judgment as a proxy for the group's judgment (Libby et al., 1987). For standardised bias to occur, the task must be reasonably complex (Einhorn et al., 1977), so that the more complex the task, the greater the ability of groups to outperform individuals by identifying their best member. It follows, therefore, that the best member must perform significantly better than the average of group members. If the best member is only marginally superior, then the group judgment is only marginally better than the average of group members' judgment (Iselin, 1991). In such a case, it may be more economical to simply take the average of group members' judgment rather than incur the added expense of assembling a group of interacting decision makers.

Information load theory, proposed by Chalos and Pickard (1985), states that groups can process large amounts of information better than individuals. As individuals have limited information processing capacity, large amounts of information normally cause cognitive strain among them (Hogarth and Einhorn, 1992). In the case of groups, several individuals working together compensate for individual inadequacies. This increases the group's corporate processing capacity and leads to more effective judgments. In addition, information load theory states that group judgments are more consistent than individual judgments because groups are able to select and weigh cues more effectively.

Prior psychology studies find that individuals weigh information cues moderately accurately, but this individual ability is applied rather inconsistently (Brehmer, 1976). This results in the inconsistent judgments observed among individual decision makers compared to group judgments (Chalos and Pickard, 1985; Einhorn et al., 1977). This theory finds support in the results of Schultz and Reckers (1981) who observed that group judgments are more consistent and groups also make use of more information in arriving at their judgment.

Other explanations of why groups outperform individuals include groups' ability to recall more information (Stasser, 1992). This ability arises in two ways. First, the recall of information is shared among group members. Together, group members may recall what an individual may have failed to recall, so that, even though no single member could recall all the information, the group could recall significantly more information than an individual. Second, this higher ability of recall is only beneficial if group members recall different information items. According to Stasser (1992), as group members may potentially recall different information items, the group discussion is enriched and the group's judgment outcome is more effective than individual judgment.

For a group to work effectively, group members must be able to pool their processing of information cues and integrate the cues into a solution. The pooling of information by group members ensures that groups have a larger information base to support their judgment and reduces the uncertainty of the task (Stasser, 1988; Stasser and Titus, 1985). This normally occurs during the group interaction process. During this interaction, suggestions and counter-suggestions are put forward by group members. An effective group should be able to identify a suggestion as its best or optimum solution.

When group members are unable to share information processing among themselves, they become less effective because group members have unequal information about the task (Iselin, 1991). The literature suggests that interacting groups are able to integrate their information processing to reduce the uncertainty of the task, eliminate redundant information search and improve judgment quality (Howell et al., 1970). To reap the benefits of group interaction, group members must pool the information cues and integrate the cues to form a solution (Hill, 1982).

In addition, group members may bring different anchors to the group process (Sniezek, 1992). As individual group members have different anchors, the range of anchors considered by the group is increased. This widens the perspective of the group. Consequently, the information processing of the group is more encompassing and the resultant judgment more effective. These explanations suggest that groups may be more efficient processors of information than individuals and they (i.e., the groups) may also combine information more effectively. All these explanations (best member theory, information load theory, groups' superior information recall, groups' potential ability to pool individual stores of information, and groups' wider range of anchors) imply that groups can process information more effectively and efficiently compared to individuals in relatively more complex judgment tasks.

This experiment examines group judgment in a going-concern context. This type of task tests the assumptions of this experiment that, due to the nature of audit tasks and audit responsibilities, more efficient and effective information processing are explanations of why audit groups outperform individual auditors. Because groups have the advantage of a wider range of anchors, it is expected that group interaction will facilitate the pooling,

recall and integration of information cues. This, coupled with the auditors' experience with, and acceptance of, group work, leads to the expectation that there will be significant differences between the judgments of audit groups and individual auditors.

Various auditing studies compare the judgment quality of groups and individuals. Several features emerge from these studies. Among these is the conservative nature of audit group judgments. Prior psychology studies find that decision makers tend to assume more risk when they move from making individual judgments to making judgments as a member of a group, that is, a risky-shift is observed (Paese et al., 1993). However, a cautious-shift is observed among audit group judgments for material tasks (Reckers and Schultz, 1982; Barrett and O'Malley, 1976). Brown (1986) explains this using cultural value theory, i.e., if the judgment context is risk-taking (for example, betting on horses), then the group process will result in riskier judgments compared to the group members' individual judgments. If the judgment context is for caution (for example, going-concern evaluation), then more conservative judgments will result. The audit environment demands conservatism and caution in judging because risky judgments are costly, especially in relation to going-concern. Consequently, audit group judgments are expected to be more conservative than individual auditor judgments. Given the above description of the nature of the audit environment, the following hypothesis is formulated:

H₁ Audit group judgments will be more conservative than individual auditor judgments.

Schultz and Reckers, 1981 indicate that judgment consensus is higher among audit groups than individual auditors. Studies by Trotman and Yetton (1985) (using a payroll internal control task), and Trotman et al. (1983) (using an inventory task) show lower consensus among interacting audit groups compared to staticised groups or composites.

The use of student surrogates may account for the inability of audit groups to outperform composites in Trotman et al. (1983). As discussed in Section 2.2, the nature of the tasks in Trotman and Yetton (1985) reduced standardised bias, and hence, no significant difference in performance was observed between composites and interacting groups. Trotman (1985) and Sniezek and Henry (1990; 1989) found that interacting group judgments are more accurate than individual judgments.

A critical aspect of the group process is whether it reduces individual inconsistency or outlier behaviour and results in higher consensus, as consensus is often used as a proxy for judgment quality (Schultz and Reckers, 1981). It is expected that social pressures created by the group process will moderate extreme views held by group members as they work to accommodate each other's views (Fisher and Ellis, 1990). This should result in higher consensus among group judgments compared to individual judgments. Psychology studies, however, observe the opposite to be true; the group process, very often, results in even more extreme views (e.g., Myers and Bishop [1971] found that groups became more racially-intolerant than the individuals). In domain-specific tasks such as auditing, such behaviour is neither expected nor acceptable. Auditors often work in groups and teams and so are used to interacting with group members. The importance of this ability is made even more significant by the nature of audit tasks and audit responsibilities and the time pressure under which auditors function. First, because of the complexities of some audits, auditors are aware of the importance of group work and

the need to share and integrate expertise (Schultz and Reckers 1981). Second, unlike the participants observed in psychology studies, auditors work with extended legal liabilities and incorrect judgments can incur severe penalties. This may cause them to work more effectively as groups and teams to reduce such eventualities. Third, auditors usually work under tight time constraints, so they are aware of the benefits of dividing the processing responsibilities among group members. Holding divergent views would be opposed to the majority of the group, and time and social pressures would interact to modify such views. Therefore, unlike the participants observed in psychology studies, the audit group process is expected to reduce individual inconsistencies, so that consensus among audit group judgments is expected to be higher than among individual auditor judgments. Thus, H₅ states:

H₂ Consensus will be higher among audit group judgments than among individual auditor judgments.

Judgment quality is often determined by agreement among decision makers (consensus) and/or agreement over a number of judgments by the same decision maker (consistency), and both are dependent upon cue usage, cue selection and cue weighting (Stocks and Harrell, 1995; Ashton, 1992; Bonner, 1990; Chewing and Harrell, 1990; Chalos and Pickard, 1985³⁵). Judgment quality often suffers when decision makers use of information is limited. For example, studies in psychology find that a hypothesis-confirming strategy often leads to erroneous judgments, and a more comprehensive information processing approach should lead to more effective judgments. Chewing

³⁵ To overcome the uncertainty of a going-concern task, consensus or agreement among auditors is often used as a proxy for judgment quality (Wright, 1988; Ashton, 1983; Libby, 1981). Whilst Einhorn (1974) cautions against over-reliance on consensus, in some audit tasks such as going-concern evaluation, consensus continues to be used as a measure of judgment quality. In addition, Ashton (1985) finds significant correlation between consensus and accuracy.

and Harrell (1990) find that increased cue usage leads to increased judgment quality. In addition, judgment quality is determined by cue selection, that is the ability of decision makers to differentiate the predictability of the cues processed (Chalos and Pickard, 1985). An inability to identify the salient cues reduces the effectiveness of the judgment. Another determinant of judgment quality is cue weighting (Stone and Dilla, 1994; Bonner, 1991, 1990; Chalos and Pickard, 1985). As these three factors are established determinants of judgment quality, this experiment compares conservatism and consensus among audit group and individual auditor judgments using these three items as explanatory factors. To explain the hypothesised differences between audit groups' and individual auditors' judgments, I examine cue usage, cue selection and cue weighting. In this experiment, self-reporting of these variables is to be used.

Audit groups' pooling and recall of information are only possible if group members interact, share processing responsibilities, compensate for individual members' inability to recall information and are able to identify the best or optimum solution put forward by group members. Due to the nature of the audit environment where the penalties for errors may be high, and where time pressure is an integral part of the environment, it is expected that auditors are able to recognise the importance of these skills and function more effectively as groups. In addition, the higher rate and greater diversity of recall would cause audit groups to attend to more information cues than individual auditors. Therefore, the following hypothesis states:

H₃ Compared to individual auditors, audit groups will attend to more information cues.

A further explanation of the expected difference(s) between audit group and individual auditor judgments is the difference between their selection of information cues. As groups are more comprehensive information processors than individuals, their judgments can also be explained by their ability to identify the predictive ability of the information cues they use (Chalos and Pickard, 1985); this is reflected in their cue selection. Since audit group judgments are predicted to be significantly more conservative than individual auditor judgments, audit groups are expected to select significantly more negative cues and fewer positive cues than individual auditors.

H₄ Compared to individual auditors, audit groups will select more going-concern problem indicators and less mitigating factors.

Having identified the relevant cues, the next task for decision makers is to weigh and combine these cues to form a judgment. While individuals can usually weigh cues with relative accuracy, they do not perform this task consistently (Brehmer, 1976). Groups are able to perform this task with greater consistency because misweighting by a group member will be corrected by other group members. The expected conservatism in audit group judgments compared to individual judgments may also be explained by the nature of their cue weighting. The more conservative judgment may be a result of audit groups giving more weight to negative information cues (i.e., going-concern problem indicators) and less weight to positive information cues (i.e., mitigating factors) compared to individual auditors.

H₅ Compared to individual auditors, audit groups will assign more weight to going-concern problem indicators and less weight to mitigating factors.

THE EXPERIMENT

Participants

The participants were 172 accountants attending a national training program that prepared them for the Institute of Chartered Accountants' Professional Year examinations in auditing³⁶. Participants were paid A\$10 for their effort.

Table 4.1 - Descriptive Statistics

Descriptive Statistics of Participants	
Average age	24 years
Range of audit experience	0 to 48 months
Average audit experience	11 months
Percentage with audit experience	55%

The Task

The going-concern task used was the same as that used by Chung and Monroe (1996). In addition to general information about the company, eight information cues (comprising four going-concern problem indicators and four mitigating factors) were provided and these were randomised for each group and individual participant³⁷. Participants were required to make a judgment on whether or not the company was a going-concern, and to identify and rate the information cues that were most relevant to their judgment. The four going-concern problem indicators were:

1. Legal proceedings have been started against the company that may result in a financial judgment that could not be met.
2. The fire department may make roller shutter doors and windows illegal in residential properties because of the risk they pose in the event of a fire.

³⁶ Table 1 shows the descriptive statistics.

³⁷ Cue diagnosticity is not tested as this has already been performed by Chung and Monroe (1996).

3. A fire destroyed the warehouse that may be uninsured.
4. Some of the company's fixed-term borrowings are approaching maturity and the company may not have realistic prospects of renewal or repayment.

The four mitigating factors were:

1. The company may have idle assets which it can sell for cash.
2. Management may take steps to reduce overhead and administrative expenses.
3. Associates of the company may take up a new issue of the company's shares.
4. Current shareholders may agree to contribute additional capital.

A complete set of the questionnaire can be found in Appendix B.

Experimental Procedures

The experiment was carried out during one of the training sessions. The procedures are detailed in Figure 4.1. Participants were randomly assigned to individual and group conditions. In all, there were 43 individual participants and 43 groups of three participants³⁸. In the group condition, participants formed groups of three at the start of the experiment. All participants were given the case materials in a booklet (Appendix B). The start of the booklet provided introductory information that explained the purpose of the study and provided an assurance of confidentiality. In the group condition, participants were asked to discuss the case materials and arrive at a consensus on whether or not the company was a going-concern. In the individual condition, participants read the case materials and made their judgment. Participants in both conditions were then required to place the case materials and their judgment into the envelope provided. The envelope contained a message which forbade them from removing materials that had been placed in it. The participants in both conditions then

³⁸ Two groups and one individual did not complete the task, this resulted in 41 group and 42 individual usable responses.

wrote the information cues that they considered were relevant to their judgment³⁹ and rated the relevance of these information cues to their judgment. Finally they completed a post-test questionnaire and were debriefed.

³⁹ Upon examining the cues listed by participants in both conditions, some decision units (both groups and individuals) wrote cues that could not be traced to the case materials. In these cases, I made the decision to eliminate these cues from the total number of information cues listed for the particular group or individual.

**Figure 4.1
Experimental Procedures**

	Group	Individual
Step 1	Read introductory information and case materials	Read introductory information and case materials
Step 2	Group discussion	(No discussion)
Step 3	Make going-concern judgment	Make going-concern judgment
Step 4	Write information cues relevant to going-concern judgment	Write information cues relevant to going-concern judgment
Step 5	Rate relevance of information cues	Rate relevance of information cues
Step 6	Complete post-test questionnaire	Complete post-test questionnaire

Group Size

The independent variable of decision mode comprised two levels - individuals and groups. There are different definitions of groups in the literature. Trotman et al. (1983) define groups as comprising two or more decision makers, whereas the psychology literature often defines groups as comprising three or more decision makers (Fisher and Ellis, 1990). Consequently, prior auditing studies used various group sizes (e.g., Trotman and Yetton [1985] used two-member groups; Trotman et al. [1983], Solomon [1982], and Uecker [1982] used three-member groups; and Chalos [1985] used four-member groups). Trotman et al. (1983) found three-member groups to be superior to two-member groups. Therefore, three-member groups were used in this study.

Dependent Variables

The case materials contained a brief description of a client together with eight information cues. Four of these cues were going concern problem indicators and four were mitigating factors. Participants were required to make a going-concern judgment on a 9-point Likert-type scale anchored by "a going-concern" (1) and "not a going-concern" (9). This is used to test H_1 and H_2 .

Extant psychology literature examines the processing, encoding and recalling of information from memory (Cohen, 1981; Snyder and Cantor, 1979; Snyder and Uranowitz, 1978; Zadney and Gerard, 1974). These studies find that decision makers selectively process information during the encoding stage of the judgment process and selectively recall information from memory (Zadney and Gerard, 1974). They tend to encode into memory and retrieve from memory mainly information consistent with their hypothesis (Snyder and Cantor, 1979). Therefore, the information recalled represents the information that is initially processed and encoded into memory (Rothbart et al., 1979). When participants are required to write the information cues they consider relevant to their judgments, they write those cues that are actually processed and encoded in memory. To capture this, when the participants in this study had made their going-concern judgment, they were required to place the case materials and their judgment into the envelope provided. The next task required participants to write the information cues they considered were relevant to their judgment. The number and nature of information cues written are used to test H_3 and H_4 respectively.

After the participants had written the information cues relevant to their judgment, they proceeded to the next page. Instructions on this page required them to turn back to the previous page and rate the relevance of the listed information cues to their judgment on 7-point Likert-type scale anchored by "slightly relevant" (1) and "highly relevant" (7)⁴⁰. Two scores were obtained, a negative score comprising the total assigned to going-concern problem indicators and a positive score comprising the total assigned to mitigating factors. These are used to test H₅.

RESULTS

Diagnostic Checks

Participants in both conditions were required to record the amount of effort they expended on the task as well as their perception of task difficulty on 9-point Likert-type scales. The effort scale is anchored by "very little effort" (1) and "a great deal of effort" (9). The mean scores are 5.90 (SD = 1.00) for the individuals and 5.75 (SD = .84) for the groups. These scores are not significantly different ($t = .71$, $p = .48$ [two-tailed]). This shows that participants in both conditions expended reasonable amounts of effort in performing the tasks. The task difficulty scale is anchored by "not at all difficult" (1) and "very difficult" (9). The mean scores are 4.40 (SD = 1.78) for individuals and 4.15 (SD = 1.12) for groups, and these scores are not significantly different ($t = .79$, $p = .43$ [two-tailed]). This indicates that participants in both conditions found the task to be

⁴⁰ In many experimental studies, self-reported data are a major source of evidence for explaining how decision makers judge or wish to judge. While there are reservations expressed about the use of such data (e.g., Nisbett and Wilson, 1977; Slovic and Lichtenstein, 1971), self-reported data or self-ratings continue to be used in judgment research. This is because prior studies find that in some circumstances, self-ratings can be useful (e.g., Goldstein and Mitzel, 1992; Reilly and Doherty, 1991 and 1989; Anderson and Zalinski, 1990; Surber 1985). For example, Surber (1985) find that self-ratings of cue importance reflect relatively accurately the impact a cue has on the judgment outcome. In spite of their limitations (Nisbett and Wilson, 1977; Slovic and Lichtenstein, 1971), self-ratings do

somewhat difficult and that task difficulty and effort expended did not contribute to the subsequent differences noted.

To test the success of the group manipulation, participants in the group condition were asked to rate their satisfaction with their group's judgment on a 7-point Likert-type scale where one indicates dissatisfaction and seven indicates satisfaction. The mean score is 5.51 (SD = .88). In addition, participants in the group condition were required to rate the amount of interaction in their group on a 9-point scale anchored by "no interaction at all" (1) and "a lot of interaction" (9). The mean score is 6.76 (SD = .89). Both these scores indicate that the group manipulation was successful.

Analyses of H₁

H₁ predicts that the group judgments will be more conservative than the individual judgments. Table 4.2 shows the going-concern judgments of groups and individuals⁴¹. A perusal of the table reveals that the average score for both groups and individuals is greater than 5 (the mid-point). The mean scores are 5.68 (individuals) and 6.30 (groups). T-results show this difference to be significant ($t = 1.72, p = .044$)⁴², and there is support for H₁. The group judgments are significantly more conservative than the individual judgments. In their review of the literature on heuristics and biases, Smith and Kida (1991) find "a pervasive, overriding concern by auditors for negative outcomes (i.e., conservatism)" (484) and attribute it to the asymmetrical loss function that auditors are exposed to. The results of this test show that the judgments of groups

provide insight into decision makers' preferences (Goldstein and Mitzel, 1992; Reilly and Doherty, 1992 and 1989).

⁴¹ A nine-point Likert-type scale anchored by "a going-concern" and "not a going-concern" was used.

⁴² All tests of significance are one-tailed unless stated otherwise.

take on an element of conservatism, while the judgments of individual auditors are more risky.

Table 4.2
Going-Concern Judgments of Groups and Individuals

<u>Going-Concern Evaluation*</u>	<u>Mean</u>	<u>SD</u>	<u>Range</u>	<u>t</u>
Individuals (n = 42)	5.68	1.87	2 to 9	1.72 (p = .044)
Groups (n = 41)	6.30	1.33	3 to 9	

* Going-concern evaluation scale: 1 = a going-concern, 9 = not a going-concern

To identify whether audit experience has an effect on the results, further analyses were carried out. Average experience is used as a proxy for the groups' experience⁴³. The ANCOVA result shows that audit experience does not have an effect on the results⁴⁴ (F = .088, p = .383) (Table 4.3). The main effect for groups/individuals remains significant at the .05 level.

Table 4.3
The Effects of Groups/Individuals and Audit Experience on Going-Concern Judgments

<u>ANCOVA Results</u>	<u>SS</u>	<u>MS</u>	<u>df</u>	<u>F</u>	<u>p</u>
Covariate - Audit experience	.24	.24	1	.09	.383
Main effect - Groups/individuals	8.17	8.17	1	3.02	.043
Residual	214.05	2.71	79		

⁴³ Total group experience was not used because it would cause the group experience to be overstated compared to the individual experience.

Analyses of H₂

Given that the case materials examined by participants are a hypothetical case comprising equal numbers of going-concern problem indicators and mitigating factors, it is not possible to determine judgment accuracy. Instead, H₂ examines inter-participant agreement (consensus) to determine judgment quality.

H₂ states that consensus will be significantly higher among audit groups than among individual auditors. That is, the group process will reduce differences between individuals. The standard deviations are 1.87 (individuals) and 1.33 (groups) (Table 4.2). This shows that consensus among groups is higher than among individuals. The F-test shows the standard deviations to be significantly different ($F = 6.88, p = .01$). These results suggest that, compared to individual auditor judgments, audit group judgments are subject to fewer fluctuations and inter-group agreement is higher. On the other hand, individual judgments are subject to more fluctuations and consensus is lower. This result supports H₂ as well as the findings of prior studies that groups will reduce outlier behaviour among individuals (Schultz and Reckers, 1981).

Analyses of H₃

H₃ predicts that audit groups will attend to significantly more information cues than individual auditors. Self-reported cue usage is examined. The mean numbers of information cues listed by individuals are 5.24 (SD = 1.69) and 6.02 (SD 1.85) by groups. Table 4.4 shows the results. Groups attended to significantly more information cues than individuals ($t = 2.02, p = .023$). Stasser (1992 and 1988) and Stasser and Titus (1985) suggest that, because of their collective information processing capabilities

⁴⁴ T-results show difference in audit experience between groups and individuals ($t = 2.63, p > .1$)

(where the inability of individual members to recall information is compensated by other members), groups attend to significantly more information cues than individuals.

This finding supports H₃.

Table 4.4
The Effect of Groups/Individuals on the Number of Information Items Used (sd)

<u>Source</u>	<u>Audit groups</u>	<u>Individual auditors</u>	<u>t</u>	<u>p</u>
Number of information cues attended to	6.02 (1.85)	5.24 (1.69)	2.02	.023

Audit experience does not have an effect on the mean number of cues used ($F = .72, p = .200$) (Table 4.5, panel A). The main effect for groups/individuals remains significant at the .05 level.

Table 4.5
The Effects of Groups/Individuals and Audit Experience on the Number of Cues Used

<u>Panel A - ANCOVA Results</u>					
	<u>SS</u>	<u>MS</u>	<u>DF</u>	<u>F</u>	<u>p</u>
Covariate -					
Audit experience	2.22	2.22	1	.72	.200
Main Effect -					
Groups/individuals	12.51	12.51	1	4.04	.024
Residual	244.56	3.10	79		

Analyses of H₄

As audit group judgments are expected to be more conservative than individual auditor judgments, H₄ predicts a significant difference in cue selection between groups and individuals. Groups are predicted to select more going-concern problem indicators and fewer mitigating factors than individual auditors. To test H₄, the nature of self-reported

cue selection was examined. The number of going-concern problem indicators and mitigating factors listed by participants were tallied (Table 4.6). Both audit groups and individual auditors reported significantly more going-concern problem indicators than mitigating factors ($t = 4.69$, $p = .000$ [groups]; $t = 5.57$, $p = .000$ [individuals]). H_7 is tested by comparing the number of going-concern problem indicators and mitigating factors between groups and individuals. Audit groups and individual auditors listed 2.46 (sd 1.60) and 2.00 (sd 1.40) mitigating factors respectively, and the difference is marginally significant ($t = 1.41$, $p = .082$) (Table 4.6). Regarding going-concern problem indicators, audit groups listed 3.56 (sd .74) items and individuals 3.24 (sd .76) items, and these are significantly different ($t = 1.96$, $p = .027$) (Table 4.6). This is opposite to the predicted direction. Thus, H_7 is not supported.

Table 4.6
Cue Selection

<u>Panel A - Mean Number of Cue Selected (SD)</u>				
	<u>Audit</u> <u>groups</u>	<u>Individual</u> <u>auditors</u>	<u>t</u>	<u>p</u>
Mitigating factors	2.46 (1.60)	2.00 (1.40)	1.41	.082
Going-concern problem indicators	3.56 (0.74)	3.24 (0.76)	1.96	.027

ANCOVA analysis (with audit experience as a covariate) shows that audit experience does not have an effect on the selection of going-concern problem indicators ($F = .73$, $p = .198$) (Table 4.7, panel A). Audit experience similarly does not have an effect on the selection of mitigating factors ($F = .08$, $p = .387$) (Table 4.7, panel B). The main effect

for groups/individuals remains significant at $p < .10$ (mitigating factors) and $.05$ (going-concern problem indicators).

Table 4.7
The Effects of Group/Individuals and Audit Experience on Cue Selection

<u>ANCOVA Results</u>					
	<u>SS</u>	<u>MS</u>	<u>df</u>	<u>F</u>	<u>p</u>
(A) Going-concern Problem Indicators					
Covariate -					
Audit experience	1.62	1.62	1	.73	.198
Main effect -					
Groups/individuals	4.22	4.22	1	1.89	.086
Residual	176.48	2.23	79		
(B) Mitigating Factors					
Covariate -					
Audit experience	.05	.05	1	.08	.387
Main effect -					
Groups/individuals	2.20	2.20	1	3.85	.026
Residual	45.08	.57	79		

Analyses of H₅

To investigate the cause(s) of the difference(s) between group and individual responses, the nature of self-reported cue weighting is examined in H₅. Two sums are obtained for each group and individual - a positive score comprising the total rating of the mitigating factors and a negative score comprising the total rating of the going-concern problem indicators. As the number of cues used by groups and individuals was different, average weight per cue was calculated by dividing the scores by the number of cues used. The average weight assigned to going-concern problem indicators was 5.26 (sd 1.73) (individuals) and 5.67 (sd 1.24) (groups), and these scores are not significantly different ($t = 1.26, p = .164$) (Table 4.8). The average weight assigned to positive cues was 5.08 (sd 2.85) (individuals) and 4.64 (sd 1.24) (groups). These scores are not significantly

different ($t = .81$, $p = .210$) (Table 4.8) and H_5 is not supported. These results show that the significantly more conservative nature of audit group judgments over individual auditor judgments is not due to the difference in cue selection or cue weighting, rather it is due to the groups' use of more information.

Table 4.8
The Effect of Groups/Individuals on Cue Weighting

	<u>Audit</u> <u>Groups -</u> <u>[Mean (sd)]</u>	<u>Individual</u> <u>Auditors -</u> <u>[Mean (sd)]</u>	<u>t</u>	<u>p</u>
Average weight of going-concern problem indicators	5.67 (1.24)	5.26 (1.73)	1.26	.164
Average weight of mitigating factors	4.64 (1.24)	5.08 (2.85)	.81	.210

Additional analyses of the effect of audit experience on cue weighting were carried out. The average weight assigned to going-concern problem indicators was not significantly affected by audit experience ($F = .88$, $p = .175$) (Table 4.9, panel A). Audit experience similarly did not have an effect on the average weight assigned to mitigating factors ($F = .31$, $p = .291$) (Table 4.9, panel B). The main effect of groups/individuals remains insignificant at the .1 level for both types of cues.

Table 4.9

The Effects of Groups/Individuals and Audit Experience on Cue Weighting

Panel A - ANCOVA Results for Going-Concern Problem Indicators					
<u>Source</u>	<u>SS</u>	<u>MS</u>	<u>df</u>	<u>F</u>	<u>p</u>
Covariate -					
Audit experience	1.70	1.70	1	.88	.175
Main effect -					
Groups/individuals	2.26	2.26	1	1.17	.141
Residual	152.23	1.93	79		
Panel B - ANCOVA Results for Mitigating Factors					
<u>Source</u>	<u>SS</u>	<u>MS</u>	<u>df</u>	<u>F</u>	<u>p</u>
Covariate -					
Audit experience	1.52	1.52	1	.31	.291
Main effect -					
Groups/individuals	2.96	2.96	1	.60	.221
Residual	306.95	2.13	79		

DISCUSSION AND CONCLUSIONS

This experiment investigates the relative performance of interacting audit groups and individual auditors in a going-concern task. The confounding observed in prior studies was controlled by using different participants for the group and individual judgments. The findings support the hypothesis that differences in information processing contribute to the differences observed between group and individual judgments. These findings show that the judgments of audit groups are significantly different from those of individual auditors. The analyses revealed that audit groups make significantly more conservative judgments than individual auditors, which is consistent with prior auditing research (Reckers and Schultz 1982; Barrett and O'Malley 1976). There is a higher degree of agreement among audit groups than there is among individual auditors. The findings of this experiment suggest that individual inconsistency may be reduced by an interacting group process, and support the prior findings of Chalos and Pickard (1985).

Using consensus as a surrogate for judgment quality, it may be said that the group process leads to more effective judgments.

H₃ states that audit groups will attend to more information cues than individual auditors. The analysis involves self-reported cue usage, and the result shows that at the overall level, groups attend to significantly more information cues than individual auditors. The group process increases the range of information cues that individual auditors would normally attend to⁴⁵. When self-reported cue selection is compared, there is a marginally significant difference between the groups' and individuals' selection of mitigating factors. Contrary to expectation, audit groups selected more going-concern problem indicators than did individual auditors. This could be attributed to the groups' general use of more information cues. The analyses of self-reported cue weighting similarly show no significant difference between the weights assigned by groups and individuals to going-concern problem indicators and mitigating factors. This suggests that the ineffectiveness of individual auditors' judgment is not due to the nature of their cue selection and cue weighting, but rather it is due to the limited range of their cue usage.

Audit experience does not appear to have a significant effect on the results. It does not result in significantly more conservative judgments. Experienced auditors do not attend to more cues than individual auditors. While audit experience does have a significant effect on the selection of individual cues, this significant effect is reduced when the going-concern problem indicators and mitigating factors are considered as a group.

⁴⁵ Note that this was not a memory test. Participants were required to write down the information cues that were relevant to their decision. To prevent them from merely copying the information from the case materials, the case materials and their prior decision had been placed in an envelope and were not available for this task.

Similarly, experienced auditors do not weight these two categories of cues significantly differently from inexperienced auditors.

These results indicate that the group process leads to higher consensus among groups generally, and the consideration of a wider range of information cues by individual groups. While it is costly to assemble a group of interacting auditors, this experiment suggests that the benefits may outweigh the costs. There is evidence to suggest that auditors are able to work more effectively as groups than other participants studied in the psychology literature.

LIMITATIONS

There are limitations to the experiment reported here. The first two limitations of Experiment One as detailed in chapter 3 also apply here. In addition, due to the large numbers of participants that were required for Experiment Two, I was constrained to use auditors who were relatively inexperienced as well as participants without audit experience. It is unclear whether experience had an effect on the results.

CHAPTER 5

EXPERIMENT THREE - COUNTEREXPLANATION AND AUDIT GROUPS

INTRODUCTION

Auditors' ability to detect fraud is generally low (Bernardi, 1994; Pincus, 1991 and 1990). Such failure may be attributed to many factors, and prior auditing research has attempted to identify some of these. Bernardi (1994), Johnson et al. (1993), Reckers and Schultz (1993) and Pincus (1991; 1990) examined the effects of personality differences on fraud detection rates. The scope of these studies is limited in that they merely examine the effects of various individual differences and personality traits on fraud detection rates and do not examine the psychology of the judgment. This experiment examines the psychology of the judgment by considering whether the provision of a counterexplanation improves fraud detection. In counterexplaining, the decision-maker is required to consider why his/her chosen judgment alternative may be incorrect and/or why the rejected alternative(s) may be correct (Koriat et al., 1980). Prior auditing literature suggests that counterexplanation may result in more effective judgments⁴⁶. These studies required auditors to counterexplain experimenter-prescribed hypotheses (e.g., Koonce, 1992; Kennedy, 1995). Another treatment of this technique in the psychology literature (e.g., Koriat et al., 1980) required participants to counterexplain their own hypotheses, and found that it reduces over-confidence. Using Koriat et al.'s treatment of explanation/counterexplanation, this experiment will provide evidence on counterexplaining by audit groups and individual auditors and serves two purposes.

⁴⁶ For example, Koonce (1992) found that it reduces explanation bias.

First, Koriat et al. (1980) used counterexplanation to reduce over-confidence, but as auditors are generally under-confident, there may be occasions where counterexplanation may result in counter-productive belief revision. There is evidence in the literature of the increasing use of counterexplanation in practice. For example, Libby and Trotman (1993) found that audit reviewers pay greater attention to evidence inconsistent with the conclusions reached by their audit subordinates in analytical procedures. Before this practice of requiring a counterexplanation becomes widespread, there is a need to investigate whether there is(are) any negative effect(s) of counterexplanation. The effectiveness of counterexplanation as a judgment tool should be more rigorously tested, that is, researchers should question whether there are conditions under which requiring an auditor to counterexplain would be counter-productive. One such question this study asks is: Will counterexplaining accurate judgments lead to less accurate judgments? If the answer to this question is yes, then counterexplanation should be used with caution.

Second, given the possible negative aspect(s) of counterexplanation, its effect on audit groups is also examined. The motivation for this is three-fold. First, group accuracy and calibration are compared to individual accuracy and calibration. The purpose of this comparison is to determine which is the more effective judgment unit. Second, I consider whether audit groups enhance the positive aspects of counterexplanation and lead to higher fraud detection rates. Third, I consider whether any potential negative aspects of counterexplanation can be mitigated through the use of audit groups.

A modified version of the case materials used by Pincus (1991), which contains material management fraud, was used in this experiment. Both individual and groups of auditors

evaluated the case materials and made a judgment on whether or not the inventory balance is fairly stated. They also recorded their confidence in their judgment. After which half the participants provided an explanation for their judgment, while the other half provided a counterexplanation. They then made a final judgment of the fair statement of the inventory balance and rated their confidence a final time. The results found that not only were audit groups more accurate and more confident, they were also better calibrated than individual auditors. Counterexplanation led to higher belief revision than explanation. While counterexplaining inaccurate judgments led to marginally higher accuracy than explaining inaccurate judgments, counterexplaining accurate judgments resulted in significantly lower accuracy than explaining accurate judgments. Counterexplaining also led to lower confidence while explaining led to higher confidence, and there was also an interaction effect between explanation mode (explanation and counterexplanation) and groups/individuals on confidence revision. Finally, causal explanations caused higher interference for those who counterexplained as well as for individual auditors, and groups counterexplained more efficiently than individuals.

EXPLANATION EFFECT

When making judgments, decision makers such as auditors are often required to explain why particular conclusions are reached. Explanation is the provision of reasons that either speak for, or provide evidence for, the chosen alternative, or the provision of reasons that speak against, or point against, the alternative rejected (Koriat et al., 1980). The need to provide explanations is required by the professional standards. For example, AUS 502 - Audit Evidence (AARF, 1996) requires auditors to obtain sufficient appropriate audit evidence to be able to draw reasonable conclusions on

which to base the audit opinion, and AUS 208 - Documentation (AARF, 1996) requires auditors to document matters which are important in providing evidence to support the audit opinion, as well as evidence that the audit was carried out in accordance with Australian Auditing Standards.

A review of the social psychology literature suggests that the provision of an explanation may affect the subsequent judgment of the decision maker. For example, having explained the possibility of the target event occurring, the decision maker is likely to judge an explained event as more likely to occur than an unexplained event. This is referred to as an explanation effect. Explanation also results in belief perseverance. After explaining a causal relation between the information and the impression, such an impression is surprisingly impermeable to change, even after the causal relation between the information and the impression is subsequently discredited (Ross et al., 1975). Explanation effect and belief perseverance are both explanation consequences, but they are dissimilar phenomena. Both will be discussed in this section as they are consequences of causal explanations.

Ross, Lepper and their colleagues carried out a series of experiments to test belief perseverance (Ross et al., 1975; Ross et al., 1977; Anderson et al., 1980). In Ross et al. (1975), two experiments were carried out. In the first experiment, student participants were required to perform a task for which they received feedback indicating their performance in relation to an average student. After a time delay, students were told of the fictitious nature of the feedback they had received, and that the feedback had no relation to their actual performance. Participants subsequently completed a questionnaire which solicited their perception of their own performance. The purpose

was to determine the extent to which the initial fictitious feedback had survived the debriefing process. As expected, "above-average" participants rated their past and future performances significantly higher than "average" and "below-average" participants, even though all of them were debriefed on the fictitious nature of the feedback. The results show a perseverance effect which survived the debriefing procedures. The second experiment followed similar procedures with three exceptions. First, a control group (no-debriefing condition) was added. In addition, a discussion of the perseverance process and its effects was added to the standard debriefing for half the debriefing participants (process debriefing condition), while the remainder received the standard debriefing (standard debriefing condition). Participants in the process-debriefing condition were told of the fictitious nature of the experiment plus the purpose of the study. Second, participants in each condition were divided into two groups - actors (who took part in a dummy physiological experiment) and observers. Each actor was paired with an observer. Third, instead of providing feedback, participants rated their own performance. The task required participants to predict a person's suicidal tendency from a note. At the end of the experiment, the actors completed a questionnaire identical to the one used in Experiment One and the observers were required to rate the performance and abilities of the actor they observed. The results showed that the perseverance effect survived the debriefing procedures especially for participants who performed above- or below-average for both actors and observers. Participants in the process-debriefing condition were of interest, as actors in this group demonstrated belief perseverance, but at a reduced level, while observers continued to show distinct belief perseverance.

While causal explanation is implied in Ross et al. (1975), Ross et al. (1977) tested explicitly for it. Belief perseverance may be due in part to the need of decision makers

to explain the relation between the information and the impression (judgment) and consequently, the impression persists even after the information has been thoroughly discredited (Ross et al., 1975). The first experiment of Ross et al. (1977) was designed to test for this phenomenon. Ross et al. (1977) required student participants to explain the later behavioural patterns of two clinical patients based on the case histories participants had previously read. In the control group, participants read one of two case histories with "attention to information about the patient's background, formative experiences and symptomatic problems that might help predict later events in the patient's life" (819). They then were required to rate the likelihood of occurrence of five possible events - suicide, financial contribution to the Peace Corps, participation in a dangerous medical experiment, alcoholism and volunteer work in a political campaign. In the manipulation group, participants also read one of the two case histories except they were asked to take the role of a clinical psychologist who had just obtained information about a subsequent event in the patient's life. Half of the participants were told that the subsequent event was suicide and the other half financial contributions to the Peace Corps. Participants then wrote the reasons that would explain the action of the patient. Afterwards, they summarised the main reasons for the patient's action without consulting the notes. They were then debriefed and told that "no information is available about whether he (she) committed suicide, made financial contributions to the Peace Corps, or followed any other particular course of behaviour" (820). Participants finally rated the likelihood of the explained event occurring. The results showed that the perseverance effect was caused in part by the explanation given, that is, after explaining a patient's later behaviour based on the case history, participants rated that behaviour as a more likely consequence even after being debriefed.

In Experiment Two, half the participants received the same instructions as those in the manipulation group in Experiment One. The other half of the participants were informed right from the start that the event to be explained was fictitious and that the experimenter had no information of its occurrence or non-occurrence. The rest of the procedures were identical to Experiment One. As the observations in Experiment One may have been a result of the participants overlooking the debriefing message in the case materials, this possibility was negated by placing the debriefing message up-front. In general, the results of Experiment Two supported the findings of Experiment One. Experiment Three followed similar procedures as the previous two except the explained events were divided into critical (hit-and-run driver) and non-critical (contribution to Amnesty International) events. The results demonstrated a belief perseverance, but only for the critical event.

Anderson et al. (1980) carried out two experiments to test the effect of explanation on the participants' judgments. Student participants were required to explain the relation that existed between success in the occupation of firefighter and risk-taking propensity before being told that such a relation is entirely fictitious. The first experiment was a 2 x 2 full factorial design where half the participants explained a positive (negative) relation between risk-taking propensity and success as a firefighter and half the participants were (were not) debriefed on this fictitious relationship. Analysis of the no-debriefing participants' responses showed overwhelmingly that those who explained a positive (negative) relationship saw risk-taking (non-risk-taking) propensity as indicative of later success as a firefighter. What was surprising was that participants who explained the positive (negative) relation in the debriefing group also reached the same conclusions.

This experiment showed a belief perseverance among participants (even though their initial beliefs were based on very minimal data).

In the second experiment, instead of randomly assigning participants to positive or negative relation groups, participants were asked to discover the relation (whether positive or negative) for themselves. One half of the participants explained the discovered relation and the other half did not. The fictitious nature of the relation was later told to half the participants, and the remainder acted as a control group. As predicted, the control group participants saw risk-taking propensity as indicative of future success, while those who explained the negative relation believed the opposite to be true. Such beliefs endured even after participants had been thoroughly debriefed in both the debriefing-explanation and the debriefing-no-explanation participants. But among the no-explanation participants, those who were debriefed showed significantly less belief perseverance than those who were not debriefed. The results show that explanation leads to belief perseverance. This belief may survive even after the participants have been thoroughly debriefed of the fictitious nature of the relation explained, and debriefing does not adequately reduce belief perseverance.

These findings are not confined to Ross, Lepper and their colleagues. Other researchers have found similar results (e.g., Sherman et al., 1981; Sherman et al., 1983; Tutin, 1993). Sherman et al. (1981) sought to replicate Ross et al. (1977) with two extensions. First, the studies by Ross, Lepper and their colleagues required participants to explain hypothetical future events concerning people unknown to them. Sherman et al. (1981) required participants to explain their own hypothetical future performance. This extension was based on the premise that people's perceptions of themselves are ordered

and stable, and therefore, their explanations of hypothetical events do not have a significant impact on their probability estimates of those events. Second, the Ross et al. studies provided participants with detailed clinical cases, thereby increasing the perceived subjective likelihood of the event explained. Sherman et al. (1981), on the other hand, required participants to perform an anagram task. At the start of the experiment, participants completed a personality checklist. They were subsequently divided into three groups. Prior to performing the task, participants in two of the groups were required to imagine that they had already participated in the anagram task and had either passed or failed the test. They were then required to explain their hypothetical success or failure in the task. They were also told that the experimenters had no knowledge of how well or poorly they would perform, so they were aware that what they were explaining was purely hypothetical. The third group acted as a control group. All participants then performed the anagram task which required them to rate whether 12 traits and 11 moods applied to them. The results of the anagram task were used as the dependent measures. The results showed that participants who explained a hypothetical success rated their chances of success in the anagram task higher, while those who explained failure rated their chances of success lower. In addition, those who explained success actually outperformed those who explained failure. These findings confirm those of Ross, Lepper and their colleagues as well as extending the findings to include actual performance.

In the above experiment, the same type of task was used for the hypothetical explanation (the hypothetical task) and the actual explanation (the actual task). Sherman et al. (1981) proposed that belief perseverance should survive even when different tasks are used for the hypothetical and actual tasks. In their second experiment, they required participants

to first explain the hypothetical outcome of an anagram task before performing a word generation task. The results showed that participants who explained their success in the anagram task, expected to perform similarly in other word tasks. Actual performance was similarly affected. Those who explained performing well in the anagram task, performed equally well in the word generation task.

Instead of belief perseverance, Sherman et al. (1983) examined the explanation effect. They proposed that explaining a hypothetical future event causes decision makers to judge the event more likely to occur. In a 2 x 2 full factorial experiment, Sherman et al. required half the participants to read detailed information about an upcoming gridiron game (between Team A and Team B) with a view to recalling the information later on. The other half of the participants were told to read the information in order to form an impression on the game. One half of the participants were told that they would subsequently be required to explain the hypothetical victory of one of the teams. The other half were not told of the explanation requirement. After reading the information, all the participants explained in detail the hypothetical victory of one of the teams. Finally, they judged which team was more likely to win the game. The results showed that participants who explained the victory of Team A judged Team A more likely to win, while those who explained the victory of Team B judged Team B more likely to win. This was observed for all the groups except for the group which was told to form an impression and not told of the need to provide an explanation. Because this group of participants was to form an impression, the information was not encoded in the individuals' memory. When asked to make a final judgment on the outcome of the gridiron game, their judgment was not affected by the explanation. The second experiment added a recall variable to the first experiment. Participants were required to

recall the information either right before or right after their final judgment. The findings of the first experiment were replicated in the second experiment. In addition, the explanation of the hypothetical event significantly affected what specific information was recalled.

As an extension of Sherman et al.'s (1983) study, Hirt and Sherman's (1985) study used both knowledgeable and naive participants in a gridiron task. In Experiment One, students were tested to determine their knowledge of gridiron before being assigned to knowledgeable and naive groups. Half the participants were told of the requirements to explain the hypothetical victory of a particular team in a forthcoming gridiron game, prior to reading the information (before-group). The other half of the participants were not told of the explanation requirement until they had read the information (after-group). Half of the after-group participants were told to read the information with a view to recalling the information afterwards (after-recall-group), and the other half of the participants were told to read the information with a view to forming an impression (after-impression-group). All the participants then wrote explanations on the hypothetical victory of one of the teams. Finally, they made a likelihood judgment on which team would win and completed a free recall task (which was either taken right before or right after the judgments). As expected, an explanation effect was observed for all participants except for the knowledgeable participants in the after-impression condition. It was, however, possible that participants' judgments were based on prior information about the teams rather than on the experimental information. Nevertheless, the results indicated a difference between the judgment of knowledgeable and naive participants. Only knowledgeable participants took part in Experiment Two. Unlike Experiment One, where the gridiron game involved teams familiar to the participants, in

Experiment Two participants were required to explain the hypothetical victory of a game involving teams they were unfamiliar with. The experimental procedures were identical to those in Experiment One. The results showed an explanation effect. The reason for this difference was that participants in Experiment One were not only knowledgeable, but were familiar with the information (the teams). In Experiment Two, while participants were knowledgeable, they were unfamiliar with the information.

The provision of an explanation has other effects besides affecting the decision maker's likelihood judgments. Hoch (1984) proposed that the provision of an explanation for a target event may interfere with the decision maker's ability to consider other perspectives which will, in turn, affect his/her likelihood judgments of the target event. Three experiments were carried out to test this proposition. In Experiment One, participants were required to write reasons supporting and opposing a hypothetical future event. They then made likelihood judgments of the occurrence on the future event. Each participant explained three events - purchase of a video cassette recorder, will Britain and Argentina go to war over the Falklands, and will OPEC impose an oil embargo. The results provided limited support that interference (the provision of reasons for and against a future event) affected participants' likelihood judgments. However, the results could also be affected by other factors. At the time of the experiment, the possible Falklands War and the OPEC oil embargo were emotive issues. In this experiment, Hoch used the number of reasons generated as a covariate to test for interference, and found that generating supporting reasons causes interference. In Experiment Two, he proposed that interference should make it more difficult for participants to generate the second set of reasons, therefore, ease of generation would be a more suitable measure of the effects of interference. Time taken was used as a proxy

for ease of generation. In this experiment, only the video purchase and OPEC cases were used. At least three reasons for and against the cases were to be generated and each time a new reason was generated, participants pressed a key. Finally, they made likelihood judgments of the events occurring. The results, with time taken as a covariate, showed that longer times taken to generate supporting reasons led to lower likelihood judgments and shorter times taken led to higher likelihood judgments. This supports the proposition that interference (in the form of reasons for and against the cases) affects participants' likelihood judgments. The third experiment aimed to test whether time delay would remove the effects of interference. The results show that time delay may reduce the effects of interference, and confirm that interference affects likelihood judgments.

All the studies cited above made use of student participants in general knowledge tasks. It is important to examine whether such findings extend to experts performing domain-specific tasks. Tutin (1993) made use of both student and clinical psychologists. The task required participants to anticipate a future event in a patient's life - either suicide or involvement in a social cause, and was discussed in chapter 2. Analyses of students' results for both tasks confirmed the prior findings of Ross, Lepper and their associates and Sherman and his associates of an explanation effect. The expert judges showed an explanation effect on the social cause case. Their results on the suicide case also showed an explanation effect, but at a reduced level compared to the students, and partially support the proposition that expertise mitigates the explanation effect. The difference between the two groups was not as large as expected. There were two explanations for this. First, there was a significant judgment difference between the psychologists who explained the suicide case first and those who explained it second. This difference was

observed across all three groups. Physiological reasons could have confounded the results as clinical psychologists are normally very busy (Tutin, 1993). Second, because psychologists frequently encounter suicide cases, they may not be motivated to carry out the experiment (Tutin, 1993).

In sum, the provision of an explanation leads to belief perseverance, and such belief perseveres even after the information that formed them initially has been discredited (Ross et al., 1975). The basis of such belief may be very little data (Anderson et al., 1980). Explanation also leads to an explanation effect. Explaining a future event biases decision makers' likelihood judgment in favour of that event occurring (Ross et al., 1977). Additionally, participants who explained their future success on a particular task, have higher expectations of success, and indeed, do perform better on the task compared to participants who explained their future failure (Sherman et al., 1981). The explanation effect persists, regardless of whether the need to provide an explanation is introduced to participants before or after they read the information, but does not occur when participants read the information with a view to forming an impression (Sherman et al., 1983). The effects are reduced when participants are both knowledgeable and familiar with the information, but are equally persistent when knowledgeable participants are unfamiliar with the information (Hirt et al., 1985). Besides affecting likelihood judgments, the provision of an explanation also interferes with participants' ability and time taken to generate counterexplanations (Hoch, 1984). There is limited evidence to show that experts may be less subject to the explanation effect than novices (Tutin, 1993). Even though her results were only partially significant, Tutin proposes that experience should mitigate the explanation effect.

As a follow-on to Tutin's study, the next task is to examine the explanation effect on auditors performing domain-specific tasks. Professional pronouncements require auditors to develop an explanation for reaching a particular conclusion. The explanation may affect auditors' subsequent behaviour in two ways. First, explanation effect may cause auditors to subjectively increase the likelihood of the explained event occurring. Such a bias may impair the effectiveness of the audit, and consequently, the audit opinion issued. Second, the explanation may also lead to belief perseverance among auditors. Church (1990) proposed that the need to maintain cognitive consistency is high among auditors as, very often, a consistent auditor is viewed as a competent auditor. The need to maintain cognitive consistency may exacerbate belief perseverance. Belief perseverance is, however, untested in auditing.

Anderson and Wright (1988) used experienced auditors (mean audit experience 3.23 years) and novices (one undergraduate course in auditing) in an internal control evaluation task. The control group was asked to judge the likelihood of material misstatement in an account balance. There were two treatment groups. In one group, participants were told of the discovery of material misstatement (misstatement condition), and the other group was told that no fraud was discovered (no-misstatement condition) prior to making the likelihood judgment. All participants then explained how the material misstatement of an account balance could have occurred, after which they wrote the procedures that would have prevented the misstatement. The dependent measures were the ratings of the internal control system's strength and risk. The results indicated an explanation effect among the students (but this was caused more by the recall-availability process rather than the process of causal construction), and more

importantly, experience mitigated the explanation effect as no explanation effect was observed among the experienced auditors.

Other auditing studies have found an explanation effect even among experienced auditors (Church, 1991; Koonce, 1992). Church (1991) proposed that auditors who are committed to their hypothesis would resort to a hypothesis-confirming strategy. Auditor participants (mean audit experience 2.5 years) were divided into two groups - high commitment and control. In the task, analytical procedures highlighted a material misstatement and the cause of the misstatement could be in either the sales or purchases cycles. The high commitment participants had to provide a written explanation to justify why they selected a particular cycle. They were also required to provide their name and address in case they were required to discuss their judgments. Participants in the control group did not explain or provide their name and address. The results showed that the high commitment participants preferred confirming to disconfirming information, whereas participants in the control group did not. Chung and Monroe (1996) required auditing students to explain inherited and self-generated hypotheses and found that explanation leads to a confirmatory behaviour. The participants were required to rate the importance of confirming and disconfirming information cues. Consistent with prior literature, participants who inherited their hypothesis were hypothesis-confirming whereas those who generated their own hypothesis were not. This was because their participants were novices.

Other studies similarly found an explanation effect among experienced auditors (e.g., Koonce, 1992). Koonce divided auditor participants (mean audit experience 2.8 years) into two groups - a control group and an explanation group. Her case study was divided

into three parts. Part A contained financial and non-financial information about the client and the client's industry. Part B contained detailed discussions with the client identifying the change in sales mix as the cause of fluctuations detected by the analytical procedures, and Part C contained the debriefing questionnaire. After reading Parts A and B, participants performed a familiarising task (Part C). The explanation participants explained why the detected fluctuations could be caused by the change in sales mix. The control participants did not explain. Subsequently, all participants assessed the likelihood that the fluctuations were caused by the change in sales mix. Using both between- and within-subjects t-tests, Koonce found that the provision of an explanation leads to higher likelihood judgments. This finding is consistent with prior psychology studies and Church's (1991) study, but is contrary to Anderson et al. (1988) and Tutin (1993). The manipulation group's explanation was in the direction of the case materials, and this may have resulted in a confounding effect. The explanation effect subsequently observed may not be an explanation effect, but may simply have occurred because participants knew what the right answer was. On the other hand, if participants explained a(n) different (incorrect) cause of the detected fluctuations, and an explanation effect was observed, the results would be more conclusive.

In audit decision-making, auditing standards require auditors to provide explanations for reaching a particular judgment or conclusion. Even if the explanation is not specifically required, explanations or causal reasoning are generated as a by-product of decision-making (Koriat et al., 1980). Various psychology and accounting studies suggest that the provision of an explanation leads to an explanation effect (Koonce, 1992; Church, 1991; Anderson and Wright, 1988; Anderson and Sechler, 1986). But explanations do not just

affect the likelihood judgments of auditors, they may affect their ability to consider alternative explanations for an error-cause (Church and Schneider, 1993).

Auditing research generally indicates that an explanation effect is observed among experienced auditors. There are several conditions under which it is observed, and these include requiring participants to recall information after an explanation has been provided (Anderson and Wright, 1988), heightening commitment by requiring participants to provide their names and addresses (for a possible post-test interview) after they have provided an explanation (Church, 1991), and providing an inherited hypothesis prior to the explanation requirement (Koonce, 1992).

Many theories have been suggested for the explanation effect phenomenon. The central theme of these theories is that of causal attribution. Tversky and Kahneman (1973) suggest the availability heuristic as a rationalisation for the explanation effect. In decision-making, people construct causal relations between the information and the target event. Such causal reasoning goes on whether it is formally required or not. By explaining the target event, the causal relation between the information and the target event becomes imprinted in long-term memory and is readily available (Carroll, 1978; Tversky and Kahneman, 1973). "A person is said to employ the availability heuristic whenever s/he estimates frequency or probability by the ease with which instances or associations could be brought to mind" (Tversky and Kahneman, 1973, 208). These instances or associations are essentially self-generated and are integrated into a single predictive judgment (Hoch, 1984). Because the information is available in memory and easily recalled, it may explain why the provision of an explanation for a target event increases the subjective likelihood of the target event occurring (an explanation effect).

As a causal explanation is automatically developed when the initial judgment is made, the requirement to provide an explanation is merely a reinforcement of what is already available in memory. Repetitive recall may strengthen the association between the information and the judgment (Tversky and Kahneman, 1973). That is partly why people react quickly to routine questions or tasks. Answers or solutions to nonroutine questions and tasks are not so easily accessible and do not come readily to mind. A search of memory is futile because the relevant information has not been imprinted in memory (Hoch, 1984). However, the availability heuristic is not only affected by frequency of recall. For example, it may be affected by the memorability of a particular event. As a result, the use of availability heuristic may lead to systematic biases (Tversky and Kahneman, 1973). Ineffective judgments may result as decision makers construct scenarios that link the present audit situation to memorable or unusual previous events. On the one hand, if the construction of such scenarios is relatively easy, decision makers will consider the occurrence of the present situation to be as likely as the past event. On the other hand, if such scenarios cannot be easily constructed, decision makers will judge the occurrence of the present situation to be lower. Other factors that affect the availability heuristic include saliency, inadequate memory search, concrete versus abstract contexts, and imaginability to name a few (Tversky and Kahneman, 1974). This heuristic, irrespective of its cause(s), results in an explanation effect (Ross et al., 1977; Anderson and Wright, 1988; Church, 1991; Koonce, 1992; Tutin, 1993) and belief perseverance (Ross et al., 1975; Anderson et al., 1980).

Other theories have been suggested to explain the explanation effect. First is the "simulation heuristic" which is similar but not identical to the availability heuristic. When decision makers are required to explain a future event, the information and their

relation to the explained event become readily available and accessible in their memory. Kahneman and Tversky (1982) call this "simulation heuristic". This heuristic suggests that decision makers will build a series of steps that connect the information (evidence) to the judgment. This cognitive heuristic is premised on the assumption that "downhill events" (a term borrowed from cross-country skiing), that is, events that are easier to construct, are deemed more likely to occur than "uphill events", that is, events that are less easy to construct (Kahneman et al., 1982). There is a need to distinguish between two types of causal explanations: explanations in the form of a restatement of facts, and explanations involving the provision of a causal account to explain the information (Anderson et al., 1980). In the former, the information is not encoded in long-term memory and should not result in an explanation effect. In the latter, the information is encoded in long-term memory, making it more readily available to the decision maker, and results in an explanation effect.

Second, decision makers selectively review the information available and encode only those cues that are consistent with the outcome explained (Sherman et al., 1983). Snyder and his associates (Snyder and Campbell, 1980; Snyder and Cantor, 1970; Snyder and Swann, 1978) found that people generally prefer information that confirms their impressions to information that disconfirms them. Consequently, people selectively review the information set for confirming information and ignore disconfirming information. The former may be more readily available in either the information set and/or in memory. This increases the availability of confirming information to the decision maker and may result in an explanation effect (Church, 1991). Human decision processes are designed to, and people do as a rule, avoid disconfirming information (Snyder and Campbell, 1980). Consequently, during the decision-making process,

people may not have encoded this type of information into memory or if they have, may have difficulty retrieving it (Hoch, 1985), and this results in an explanation effect.

Third, the explanation effect may also be explained by what Hoch (1984) refers to as "interference". According to him, when a decision maker generates a set of reasons for (against) the occurrence of an event, this act interferes with his/her ability to generate reasons against (for) the occurrence of the event. The alternative explained first becomes encoded in long-term memory, whereas the encoding in long-term memory of the latter-explained alternative may be less prominent. This is because the initially-explained alternative interferes with the decision maker's ability to consider alternative explanations. Consequently, the alternative initially-explained becomes more readily available to the decision maker and influences his/her likelihood judgments. Two factors support this interpretation of the explanation effect. First, research has found that the mere act of imagining or explaining a hypothetical future event results in an explanation effect (Sherman et al., 1983; Ross et al., 1977). Second, based on Hoch's proposition, if a participant explains before counterexplaining an event's occurrence, then an explanation effect should result. Conversely, if s/he counterexplains before explaining, then an explanation effect should not result. Hoch's (1984) findings confirm this⁴⁷. Changing the order of the explanation/counterexplanation manipulation does not improve calibration, the participants remain over-confident (Koriat et al., 1980).

In the above discussion, four hypotheses are offered for the explanation effect. These include the availability heuristic, simulation heuristic, confirmation bias and

interference theory. It should be understood that the psychology literature does not consider these as mutually exclusive theories. There is reason to believe that these theories operate together. For example, Hoch (1985) explained that decision makers are generally over-confident because they tend to generate more confirming than disconfirming information, and this imbalance in information generation may be explained by three theories. First, confirming information is more available in memory than disconfirming information because feedback is more likely to be confirming rather than disconfirming, and the ease of remembering confirming information is higher. Second, the generation of one set of reasons may interfere with the generation of a second set of reasons. As people are more likely to generate confirming information first, this interferes with the generation of disconfirming information. Third is the question of control: decision makers find information about controllable situations easier to generate than information about situations they do not control and this leads to higher confidence. Nisbett and Ross (1980, 182) attribute belief perseverance to two theories - confirmation bias and availability. Hoch's (1985) and Nisbett and Ross's (1980) explanations show that these theories do not operate in isolation of each other. So it is reasonable to expect the theories discussed above to operate jointly to create the explanation bias.

Various conclusions can be drawn from prior studies on the effect of explaining. First, explanations are a by-product of decision-making, that is, explanations are generated automatically in the judgment process (Koriat et al., 1980). Second, the provision of an explanation leads to an explanation bias. The event explained is subjectively judged

⁴⁷ However, Koonce (using explanation before counterexplanation) did not find a primacy effect, instead she found a recency effect.

more likely to occur as the explanation increases the saliency of the causal relations between the information and the explained event and increases its availability. Third, not all social theories are subject to the same explanation effect (Anderson and Sechler, 1986). It appears that an explanation effect is weaker for judgments that require extreme initial beliefs such as views on television violence. Also, the explanation effect is less pronounced if the generation of opposing arguments is relatively easy (Sherman et al., 1985). When opposing arguments are easy to generate, the availability of supporting arguments is either reduced or the opposing arguments become equally available. Fourth, the provision of an explanation also leads to belief perseverance as has been observed in the various psychology studies quoted above.

HYPOTHESES DEVELOPMENT

The first part of Experiment Three examines the group process in relation to accuracy, confidence and calibration. The second part of this experiment examines explanation and counterexplanation by groups and individuals.

The theories that support the hypothesis that groups will outperform individuals were discussed in chapter 4. In a complex task (such as the one manipulated in this study), where the standardised bias is large, interacting audit groups are expected to outperform individual auditors. This study manipulates a consensus group where group members interact unrestrained in a face to face discussion to arrive at a single consensus judgment. Prior studies find that this type of group outperforms an average individual (Sniezek and Henry, 1989; Einhorn et al., 1977). This discussion leads to the following hypothesis:

H₁ Audit group judgments are more accurate than individual auditor judgments.

According to the revision and weighting model postulated by Sniezek and Henry (1990; 1989), consensus group judgments are reached in a two-stage process. At the revision stage, individually-held judgments are voluntarily revised in the light of views and information presented during group interaction. This revision continues until new ideas are exhausted or time has run out. At this stage, if group members' individually-held judgments are still divergent, some means of combining these divergent judgments has to be adopted. Regardless of the approach adopted (e.g., average members' judgment, select the most common judgment or randomly select a member's judgment as proxy for the group's), some form of weighting of members' views occurs. While the revision process is mostly private, the weighting process is public as group members negotiate their individual judgments to form a single group judgment. Confidence, in a group process, is important because confidence determines when the judgment process terminates and the judgment is made (Pincus, 1991). Sniezek (1989) proposed that in "group interaction it is confidence that is important because confidence, and not accuracy, can be made explicit at the time the (judgments) are made" (172).

When a group member revises his/her individual judgment, it implies that s/he has higher confidence in the revised one (Sniezek and Henry, 1990). However, a revision in confidence may not involve a corresponding revision in judgment. For example, a group member may have lost confidence in his/her own judgment. In the group weighting process, when group members adopt another's judgment as a proxy for the group judgment, it implies that they (the group members) have higher confidence in that

person's judgment than in their own (Sniezek and Henry, 1990). Group members with lower expertise and lower confidence should give way to members who have more expertise and higher confidence (Sniezek, 1989). Sniezek and Henry (1990; 1989) found that the group process leads to higher group confidence compared to the members' confidence in their pre-group judgment. During the weighting process, when the group negotiation is conducted publicly, there may be social pressures on group members to verbalise higher confidence as lower confidence is inconsistent with the group's choice of a particular judgment and may undermine group cohesiveness. Consequently, it is expected that group confidence will be higher than individual confidence. Therefore, the next hypothesis states:

H₂ Audit groups will have higher judgment confidence compared to individual auditors.

While the social psychology literature resonates with examples of over-confident decision makers (e.g., Butterworth, 1988; Christensen-Szalanski and Bushyhead, 1981; Fischhoff et al. 1977; Oskamp, 1965), auditing studies suggest that auditors may be under-confident (e.g., Mladenovic and Simnett, 1994; Pincus, 1991). Reasons offered to explain this finding include the conservative nature of auditor training, prior negative experience with inaccurate judgments, and most importantly, the extended nature of auditors' legal liability. While this is not as dangerous as being over-confident, there are, nevertheless, implications for the accounting profession. Under-confident auditors may take more time and/or accumulate more evidence than is necessary to reach a judgment. Auditors might not be as efficient as they should be, and this inefficiency could be passed on to clients in the form of higher fees.

Decision makers generally have difficulty quantifying their confidence level so that their confidence rating alone is not an adequate measure of confidence (Gibbins, 1984). Matching accuracy level to confidence rating is a more appropriate measure of confidence (Lichtenstein et al., 1982). High confidence is only desirable if it is accompanied by high accuracy. Groups are expected to be better calibrated than individuals because, first, groups have the advantage of a broader information base to support their judgment through their ability to recall more information and the pooling of their individual store of information; and second, when group members weigh and revise their judgment during the group interaction process, they also adjust their confidence level. The end result of this process is better calibration. Therefore, audit groups should have more appropriate confidence levels compared to individual auditors who do not have these advantages. This discussion leads to the following hypothesis:

H₃ Audit groups are better calibrated than individual auditors.

Auditing standards require auditors to document explanations for audit findings (e.g., AARF, 1996; AICPA, 1982). As mentioned before, documenting explanations reinforces causal explanations already generated by the auditor. When the auditor considers that sufficient causal explanations have been generated and his/her confidence level is sufficient, the judgment process ceases, sometimes prematurely. Koonce (1992) required auditors to explain an inherited hypothesis, whereas auditors often explain a self-generated hypothesis. Different results may have been observed if Koonce's participants had explained their own hypothesis, as an auditor's commitment to a hypothesis is dependent upon its source (Church, 1990).

Explanation bias is the result of causal explanations, therefore, any corrective technique must start with an examination of causal explanations. Anderson and Sechler (1986) proposed that in the face of two competing solutions to a problem, the solution where causal relations are easier to imagine or recall would be the one adopted. One way of countering the effect of causal explanations in decision-making is to make opposing evidence more salient through the use of counterexplanation (Anderson, 1982). Counterexplanation leads to a reduction of the explanation effect because, first, consideration of counterexplanation increases the versatility of a decision maker's judgment processes, and this should result in lower explanation bias. Second, explanation bias may be a result of availability, so the act of considering counterexplanations makes opposing arguments equally available, and the explanation bias is reduced. Third, counterexplanation causes belief revision. Therefore, it is expected that belief revision of auditors who counterexplained will be higher than belief revision of auditors who explained.

H₄ Belief revision of auditors who counterexplained will be higher than belief revision of auditors who explained.

Koriat et al. (1980) first used counterexplanation to improve the calibration of decision makers. As causal explanation (or the generation of supporting reasons) is a by-product of decision-making, over-confidence results when counterexplanations are ignored (Koriat et al., 1980). Decision makers such as auditors do not generate counterexplanations automatically and must be specifically instructed to do so (Koriat et al., 1980). Requiring decision makers to explicitly consider counterexplanations forces them to consider disconfirming information and over-confidence is reduced. Similarly, accuracy may also be improved by increasing the consideration of opposing evidence

(Hoch, 1985). Confirmation bias results when decision makers ignore disconfirming information and encode only information consistent with their hypothesis. Counterexplanation requires decision makers to consider disconfirming information, i.e., information that has been ignored initially. Therefore, decision makers who make inaccurate judgments are forced to examine information supportive of an accurate judgment. This should prime them of the inaccuracy of their judgment and should result in more accurate judgments.

In testing the effect of counterexplanation, unlike Koonce (1992), this study required participants to explain or counterexplain their own hypothesis (Hoch, 1985; Koriat et al., 1980). Consistent with practice in prior psychology literature (e.g., Koriat et al., 1980), no explanation was required prior to counterexplaining⁴⁸. As counterexplaining makes opposing arguments more salient to the auditor, those who have counterexplained will re-examine the data set for opposing arguments. Consequently, auditors who have counterexplained inaccurate judgments should switch to more accurate judgments. They should also be more accurate than those who have explained inaccurate judgments. Consequently, the following hypothesis states:

H₅ Auditors who counterexplained inaccurate judgments will be more accurate than auditors who explained inaccurate judgments.

Revising a judgment after providing a counterexplanation is only desirable when the switch is from an inaccurate judgment to an accurate one. Auditors who have made accurate judgments should not revise them when they are required to counterexplain.

⁴⁸ This is because causal explanations are generated as a natural process of decision-making (Hoch, 1985; Koriat et al., 1980).

Koriat et al. (1980) and Hoch (1985) show that counterexplaining results in reduced confidence among over-confident participants. Auditors, on the other hand, are under-confident for reasons previously discussed (see also Mladenovic and Simnett, 1994; Solomon et al., 1982; Tomassini et al., 1982). Therefore, caution should be exercised in applying any judgment tool on auditors in case it reduces their confidence further. This could result in such uncertainty that counter-productive belief revisions result, i.e., when accurate judgments are substituted for less accurate ones. Consequently, auditors who counterexplained accurate judgments may be less accurate than auditors who explained accurate judgments. Explaining reinforces causal explanations, removes uncertainty, and should not result in negative belief revision. On the other hand, counterexplaining creates uncertainty, and may result in negative belief revision. The following hypothesis indicates that the use of counterexplanation on auditors may be counter-productive.

H₆ Auditors who counterexplained accurate judgments will be less accurate than auditors who explained accurate judgments.

Given that audit groups have the advantage over individual auditors of a larger information base to support their judgments due to their increased ability to recall and process information, an interaction effect between groups/individuals and explanation mode is expected. In arriving at their initial judgment, it is likely that audit groups will have considered arguments which both support and oppose their judgment and will decide in favour of the supporting arguments. Individual auditors will pay greater attention to mainly supporting arguments, and may not have processed the opposing arguments. Therefore, counterexplaining an accurate judgment will have a more negative effect on individual auditors than on audit groups, as it will cause individual auditors to re-examine the data set and to process information that has not been

processed previously. This additional information - whether relevant or irrelevant - creates uncertainty. Explaining accurate judgments will have a more positive effect on audit groups than on individual auditors. This is because the explanation reaffirms the groups' higher amount of causal explanations. Additionally, groups may have already processed both supporting and opposing arguments, and are therefore less likely to be affected by the counterexplanation. The interaction effect hypothesis is stated as:

H₇ There will be an interaction effect between groups/individuals and auditors who explained/counterexplained accurate judgment on judgment accuracy.

As discussed above, causal explanations are generated automatically between the information and the target event (Hoch 1985; Koriat et al. 1980). On the one hand, requiring a decision maker to subsequently explain a judgment reinforces these causal explanations and serves to increase the decision maker's confidence. Counterexplanation, on the other hand, requires the consideration of opposing arguments. This essentially forces the decision maker to consider arguments that may not have been considered previously. Contrasting or conflicting information causes cognitive strain and, as Koriat et al. discovered, leads to a lowering of confidence. It is expected that auditors who provide an explanation will increase their confidence while auditors who provide a counterexplanation will decrease their confidence.

H₈ Provision of an explanation will increase confidence whereas provision of a counterexplanation will decrease confidence.

There should be an interaction effect between explanation mode (whether explanation or counterexplanation) and groups/individuals on confidence revision as audit groups and individual auditors do not react in the same way to explanation and counterexplanation.

Audit groups, being more efficient processors of information (Chalos and Pickard, 1985) and more confident, will increase their confidence more than individual auditors when required to explain. On the other hand, counterexplaining by audit groups will result in larger negative confidence revision compared to individual auditors because (as discussed earlier) audit groups are less over-confident than individual auditors. Individual auditors, who are over-confident, will be expected to have higher confidence than individual auditors when required to counterexplain. Therefore, the interaction hypothesis is stated as:

H₉ There will be an interaction effect between explanation mode and groups/individuals on confidence revision.

Very often in decision-making, opposing arguments are neglected for many reasons. Interference theory suggests that the provision of the first set of reasons reduces the ability of decision makers to generate the second set of reasons (Hoch 1984). That is, the first set of reasons is easier to generate than the second. As people automatically ascribe causal relations between the information and their judgment, when they are subsequently required to explain, they are able to do so with relative ease. However, generating counterexplanations requires more effort for many reasons. First, people do not normally counterexplain when they make judgments, therefore counterexplanations are not readily available from memory. If decision makers are required to produce counterexplanations, they would have to deconstruct the causal attributions already formed in their memory and reconstruct opposing arguments such that these opposing arguments become imprinted in memory, and therefore, available for performing the task. Second, disconfirming information causes cognitive strain (Snyder and Swann 1978) as decision makers do not, as a rule, consider disconfirming information. Human

cognitive processes are designed to avoid cognitive strain. In addition, the consideration of disconfirming information may lead to the need to change one's initial judgment. Very often, changing one's judgment in an audit environment is viewed as being inconsistent and incompetent (Church 1990). Third, causal relations for "downhill" events are relatively easier to construct than "uphill" events (Kahneman and Tversky 1982). Downhill events in the form of causal explanations are already constructed as part of decision-making. Fourth, interference makes it relatively more difficult for decision makers to consider opposing arguments. Therefore, it is predicted that counterexplanations are more difficult to provide than explanations, and auditors who explained will provide significantly more explanations while those who counterexplained will provide significantly fewer counterexplanations. This discussion leads to the following hypotheses:

- H₁₀ Auditors who counterexplained will rate the task as more difficult than auditors who explained.
- H₁₁ The number of explanations provided by auditors who explained will be more than the number of counterexplanations provided by auditors who counterexplain.

Audit groups should, however, be able to provide explanations and counterexplanations relatively easier than individual auditors can. As groups are able to process more information and because of their corporate efforts in recalling and pooling information, they will have considered both supporting and opposing arguments in their decision-making process. Through their sharing of information processing and their interaction, alternative solutions to the problem will have been discussed and considered. Consequently, individual auditors are able to explain and counterexplain with more difficulty than audit groups. The former can also provide fewer explanations and

counterexplanations than the latter. Based on this discussion, the following hypothesis is formulated:

- H₁₂ Individual auditors will rate the task as more difficult compared to audit groups.
- H₁₃ Individual auditors will provide fewer explanations/counterexplanations than audit groups.

THE EXPERIMENT

Case Materials

Inventory is often the most elaborate and time-consuming section of the audit for several reasons (Arens et al., 1992, 654). First, inventory is often a material item on the balance sheet, and very often, it makes up the largest current asset. Second, the difficulty of an inventory count is compounded when the client's inventory is spread over a large geographical area. Third, the wide range of inventory items also creates counting and valuing problems for the auditor. To overcome this, auditors may require the assistance of experts. Fourth, the difficulty of valuing inventory is complicated by the different methods available, and the need to determine obsolescence. Consequently, the audit of inventory is not always as effective as it should be. In a study that compared 55 accounts receivable audits with 26 inventory audits performed by a large public accounting firm, substantially more errors were found in the inventory audits than in the accounts receivable audits and the error rates were higher for the larger accounts and for accounts with the larger line items (Johnson et al., 1981). The errors in inventory audits include both over- and under-statement of balances, material misstatements, and the failure of the auditors to detect fraud (Johnson et al., 1981). For the reasons cited above, the audit of inventory often involves a group of auditors.

The inventory audit task used was a modified version of the task used by Pincus (1991). This task contained a material misstatement of the inventory balance by management. The modifications made include the conversion of weights and measurements to metric, the use of current dates and food items and proper names to reflect an Australian environment. Forty-eight information items out of the original 70 were used and the information for the latest three years were given⁴⁹. The dollar values were not changed as such changes were not considered necessary because the type of restaurant had been changed from a Mexican restaurant to a fast-food restaurant. The final version of the case was examined by two experts in the hospitality and restaurant industry for realism. Both experts were of the opinion that the case was realistic and reflective of current Australian restaurant practice. Even though the case had been simplified, it contained sufficient information to successfully test the variables manipulated in this study.

The case provided information on the background of the company as well as current year's and prior years' inventory and other financial data. The information provided included both positive and negative indicators of the fair statement of the inventory balance so the judgment was not clear cut. The results of the current year's audit procedures for inventory were also available together with selected client records. The fraud was committed by management by overstating ending inventory quantities and values. This can be detected via the various risk indicators such as excessive ending inventory, inventory count procedures, management control, over-ride of internal control procedures by management, and inconsistent opening and closing inventory

⁴⁹ The items deleted were not relevant to the participants' decisions and were deleted in the interest of time.

amounts for the year being audited. A complete set of the case materials for the group/counterexplanation condition is found in Appendix C.

Participants

Participants were 271 accountants attending a national training program that prepared them for the Institute of Chartered Accountants' Professional Year examination in auditing. The average age of the participants was 24 years (sd 2.4 months) and the average working experience was 34 months (sd 27 months). Sixty-one percent of participants had audit experience, which ranged from one to 72 months. The average audit experience for those with audit experience was 11.25 months, and the overall level of audit experience was 10.15 months. The average number of inventory audits performed was three. All the participants had completed at least 40 hours of audit training, which included a four-hour session on auditing inventory together with practical exercises. Therefore, the non-auditors also had some training in this area. Participants were paid A\$20 each for their efforts.

Experimental Procedures

The experiment was carried out during one of the training sessions. The 2 x 2 full factorial experiment tests two explanation modes - explanation and counterexplanation, and interacting groups' and individuals' judgment process. Table 5.1 shows the experimental design and Figure 5.1 shows the experimental procedures. In the group condition, participants formed groups of three at the start of the experiment while those in the individual condition worked alone. All participants received the case materials in a booklet. The front of the booklet provided introductory information which explained the purpose of the experiment and provided an assurance of confidentiality. In both

group and individual conditions, participants were asked to familiarise themselves with the case materials. They read the initial information of the case materials which contained a general description of the operations of a chain of fast-food restaurants. The instructions to all participants read:

Instructions

In order for you to perform an audit of Tucker House's inventory balance, the following information items are available. You should only evaluate those information items you consider are necessary for forming an opinion on whether or not the inventory balance is fairly stated.

Table 5.1 - Experimental Design

Groups/individuals	Explanation Mode*		Total
	Explan- ation	Counter- explanation	
Audit groups	34	31	65
Individual auditors	39	39	78
Total	73	70	143

*Numbers indicate cell size.

The instructions were followed by the content pages, which showed the types of information available, followed by the 48 information items. After acquainting themselves with the information, participants in the group condition were instructed to assume that they had been assigned as a member of an audit group for the audit of the inventory balance of a client. They were also told to discuss the case with the other group members. A large envelope was provided and participants were required to insert each completed task into the envelope before proceeding to the next task. A message on the envelope stated "Do not remove materials that have been inserted into the envelope". Each task was stapled separately and arranged in sequence. All pages of each

set of case materials bore a unique identifying number. Individual auditors received the same instructions except they worked alone.

Figure 5.1
Experimental Procedures

Step 1	Read background information			
Step 2	Read case information			
	Individuals		Groups	
Step 3	Make initial judgment (J1) (likelihood and confidence judgments)		Make initial judgment as a group (J1) (likelihood and confidence judgments)	
Step 4	(A) Write explanation	(B) Write counter-explanation	(C) Write explanation as a group	(D) Write counter-explanation as a group
Step 5	Make final judgments (J2) (likelihood and confidence judgments)	Make final judgments (J2) (likelihood and confidence judgments)	Make final judgments as a group (J2) (likelihood and confidence judgments)	Make final judgments as a group (J2) (likelihood and confidence judgments)
Step 6	Assess difficulty of generating explanation/counterexplanation		Assess difficulty of generating explanation/counterexplanation as a group	
Step 7	Complete post-test questionnaire			

For the groups, the first task required participants to discuss and work on the case as a group to arrive at a consensus on whether or not the inventory balance was fairly stated, after which they rated their confidence. The procedures for individual auditors were similar except they performed the tasks on their own. The results from this part of the

experiment were used to test H_1 , H_2 , and H_3 , which address differences between groups and individuals on accuracy, confidence and calibration.

In the second task, the explanation mode was manipulated. Half the participants in the group and individual conditions were randomly assigned to the explanation condition while the other half were assigned to the counterexplanation condition. Participants in the explanation condition were required to consider the reasons why their answer in task one may have been correct, and/or why the alternative they rejected may have been incorrect. Participants in the counterexplanation condition considered the reasons why their answer in task one may have been incorrect, and/or why the alternative they reject may have been correct. Participants in both conditions then recorded their explanations or counterexplanations in the space provided as part of task three. The number of explanations or counterexplanations recorded was used as the dependent variable to test H_{11} and H_{13} . The groups recorded these after discussion among group members.

In task four, all participants were required to make a final judgment on whether or not the inventory balance was fairly stated. The groups made a consensus judgment while the individuals made the judgment on their own. This was used to test H_4 , H_5 , H_6 , and H_7 , which addressed group and individual belief revision. They recorded their judgment and confidence on similar scales as those reported in task one. The difference between the two confidence scores was used to test H_8 and H_9 which addressed confidence revision. In addition, participants in the explanation condition were required to indicate the ease/difficulty of generating explanations. Participants in the counterexplanation condition rated the ease/difficulty of generating counterexplanations on a similar scale. These difficulty scores were used to test H_{10} and H_{12} . In the post-test questionnaire, all

participants provided various demographic as well as some diagnostic information. They were then paid. In all, the participants took between 40 to 60 minutes to complete the experiment.

Independent Variables

Explanation and Counterexplanation

Consistent with Koriat et al. (1980), participants explained (explanation participants) and counterexplained (counterexplanation participants) their own hypothesis. The counterexplanation participants were not required to provide an explanation prior to counterexplaining. The instructions to participants read:

There are many ways by which people can check the appropriateness of their decision. One of these is to require you to consider why your decision in Task I may be correct (incorrect).

Please consider the reasons why your previous answer may be correct (incorrect) and/or why the alternative you rejected may be incorrect (correct).

Groups/Individuals

The independent variable of groups/individuals comprised two levels - individuals and groups. There are different definitions of groups in the literature. Trotman et al. (1983) define groups as comprising two or more decision makers, whereas the psychology literature often defines groups as comprising three or more decision makers (Fisher and Ellis, 1990). Consequently, prior auditing studies used various group sizes (e.g., Trotman and Yetton [1985] used two-member groups; Trotman et al. [1983], Solomon [1982], and Uecker [1982] used three-member groups; and Chalos [1985] used four-

member groups). Trotman et al. (1983) found three-member groups to be superior to two-member groups. Therefore, three-member groups were used in this study.

Dependent Variables

Accuracy

The case materials contained a material misstatement of inventory by management, and consequently, the inventory balance was not fairly stated. Participants were required to make judgments on whether this balance was fairly stated. Two scales were used. The first was a 9-point Likert-type scale anchored by "fairly stated" (1) and "not fairly stated" (9). Belief revision was represented by the absolute difference between the two accuracy scores. The second measure was a dichotomous scale where participants indicated their judgment by checking either "the inventory balance is fairly stated" or "the inventory balance is not fairly stated". They made these judgments twice - initial and final judgments.

Confidence

Consistent with prior research (e.g., Mladenovic and Simnett, 1994; Pincus, 1991), a self-reported confidence measure was used in this experiment. After each individual auditor or audit group had made a judgment, they were required to rate their confidence on a 101-point Likert-type scale which was anchored by "not confident at all" (0%) and "completely confident" (100%). As participants made accuracy judgments twice, confidence rates were also measured twice - once for the continuous judgment and another for the dichotomous judgment. Confidence revision was obtained by taking the relative difference between the initial and final confidence rates.

Number of Explanations or Counterexplanations

After making the initial judgment, participants were required to consider why their judgment may have been correct (incorrect). In the next task, they were required to write down these explanations (counterexplanations), and the number of explanations (counterexplanations) provided were tallied. Explanations (counterexplanations) provided by participants that were not contained in the case materials were eliminated from the tally.

Difficulty of Providing Explanations and Counterexplanations

After making the final judgment, participants were required to record the difficulty of providing explanations (counterexplanations), and they recorded these on 9-point Likert-type scale anchored by "very easy" (1) and "very difficult" (9) (difficulty score).

RESULTS AND DISCUSSION

Diagnostic Checks

ANOVA assumes that the variance across all cells are equal. Consequently, Bartlett's Box F tests and Cochran's C test were performed in each analysis to test this assumption. In all cases (except in the analyses of confidence revision), the homogeneity of variance assumption was not violated. In examining the confidence revision assessment, the homogeneity of variance assumption was violated (Bartlett Box F: $p < .07$ and Cochran's C: $p < .01$). As the design involved the manipulation of confidence revision with two independent variables, to partly compensate for the violation of this assumption, two one-way ANOVAs were performed. The results of these ANOVAs are consistent with those of the original two-way ANOVA analysis.

All participants were required to record the amount of effort they expended on the task and their perception of task difficulty on nine-point Likert-type scales⁵⁰. The effort scale was anchored by "very little effort" (1) and "a great deal of effort" (9). The mean scores for effort expended were 5.98 (groups) and 5.82 (individuals). These scores are not significantly different ($F = .54$, $p = .46$ [two-tailed]) and show that participants in both group and individual conditions expended similar and reasonable amounts of effort. Explanation mode did not have a significant effect on effort expended. Individual auditors found the task to be marginally more difficult than audit groups did (individuals - mean 5.79, sd 1.39; groups - mean 5.35, sd 1.32) ($F = 2.75$, $p = .10$ [two-tailed]). This is to be expected given the nature of the task, and the fact that the individuals worked alone. Explanation mode did not have a significant effect on task difficulty ($F = 2.19$, $p = .14$ [two-tailed]).

To test the success of the group manipulation, participants in the group condition were required to rate their satisfaction with their group's judgments on a seven-point Likert-type scale where one indicated dissatisfaction and seven indicated satisfaction. The mean satisfaction score was 5.48 (sd - 1.14). In addition, participants in the group condition were also required to rate the amount of interaction among the members of their group on a nine-point Likert-type scale anchored by "no interaction at all" (1) and "a lot of interaction" (9). The mean score was 6.50 (sd 1.26). Both these scores show that the group manipulation was successful. Group interaction and discussion did not have an effect on the accuracy of the groups ($t = -.33$, $p = .744$ and $t = .45$, $p = .658$ respectively).

⁵⁰ The group scores comprise the average score of group members.

The mean audit experience for the audit groups was 1.14 (sd .58) and for the individual auditors it was .72 (sd .85). T-test shows this difference to be significant ($t = 3.38$, $p = .001$ [two-tailed]). To control for this, ANCOVA analyses with audit experience as a covariate were carried out.

Group Versus Individual Performance⁵¹

H_1 predicts that audit groups will be more accurate than individual auditors. As the case materials contained material misstatements by management, the inventory balance was not fairly stated. Since participants recorded their initial evaluation of the inventory balance on a scale anchored by "fairly stated" (1) and "not fairly stated" (9), the higher score represents the more accurate participant(s). The mean accuracy scores were 5.54 (sd 2.07) for groups and 4.26 (sd 1.97) for individuals. Because audit groups were more experienced compared to individual auditors, the analyses are carried out with audit experience⁵², task difficulty and effort expended as covariates. The results indicate that the difference between the accuracy scores is significant ($F = 13.28$, $p < .001$)⁵³ (Table 5.2), and H_1 is supported. ANCOVA results show that audit experience had a marginally significant effect on accuracy ($F = 2.52$, $p = .053$) (Table 2). This result is expected as audit experience is an integral part of the audit environment. Accuracy was not affected by task difficulty and effort expended.

⁵¹ One individual did not complete this section of the task, so H_1 , H_2 and H_3 are analysed with 65 groups and 77 individuals.

⁵² Average audit experience is used as a measure of group experience.

⁵³ All tests of significance are one-tailed unless stated otherwise.

Table 5.2
The Effect of Groups/Individuals on Accuracy

	<u>SS</u>	<u>MS</u>	<u>df</u>	<u>F</u>	<u>p</u>
Covariates:					
Audit experience	10.17	10.17	1	2.52	.053
Task difficulty	5.87	5.87	1	1.46	.125
Effort expended	2.64	2.64	1	.65	.210
Main effect:					
Groups/individuals	53.59	53.59	1	13.28	.000
Residual	544.95	4.04	137		

Similar results were reflected by the dichotomous judgments. Thirty-seven out of 65 (57%) groups made accurate judgments, whereas only 19 out of 77 (25%) individuals made accurate judgments. The difference between these is significant ($z = 3.90, p = .0001$). These results are not directly comparable to Pincus' (1991) results due to the modifications made to the case materials. However, the results do show that audit groups outperform individual auditors very significantly.

Table 5.3
Groups/Individuals Dichotomous Accuracy Judgment

	<u>Groups</u>	<u>Individuals</u>
Fairly stated	28	58
Not fairly stated	37	19

H₂ predicts that audit groups will be more confident than individual auditors. Participants recorded their confidence on a 101-point Likert-type scale anchored by "not confident at all" (0), and "completely confident" (100). The mean confidence scores for the initial continuous judgment were 74.5% (sd 15.3) for groups and 67.5% (sd 15.3) for individuals. The analyses are carried out with audit experience, task difficulty and effort

expended as covariates. The results indicate that whether auditors are working as groups or individuals have a significant effect on confidence ($F = 5.21, p = .004$) (Table 5.4). Audit groups are significantly more accurate and more confident than individual auditors. Therefore, H_2 is supported. Further analyses using the confidence scores associated with the dichotomous judgment was carried out. The mean confidence scores were 68.22 (sd 14.27) for groups and 65.77 (sd 16.45) for individuals. The difference between these two scores is not significantly different ($t = .94, p = .174$). Thus, when participants are required to make dichotomous judgments, audit groups are not more confident than individual auditors, and there is no support for H_2 . ANCOVA results show that audit experience has a significant effect on confidence ($F = 5.21, p = .012$) (Table 5.4). This result is expected given that audit experience is an important part of the audit environment. Task difficulty and effort expended did not significantly affect confidence. The effect of groups/individuals on confidence remains significant at the $p < .01$ level.

Table 5.4
The Effects of Groups/Individuals on Confidence*

	<u>SS</u>	<u>MS</u>	<u>df</u>	<u>F</u>	<u>p</u>
Covariate:					
Audit experience	745.06	745.06	1	5.21	.012
Task difficulty	141.44	141.44	1	.62	.216
Effort expended	358.14	358.14	1	1.56	.106
Main effect:					
Groups/individuals	1194.18	1194.18	1	5.21	.004
Residual	34563.33	245,13	137		

*This represents the initial confidence of participants expressed in relation to their judgments on the 9-point Likert-type accuracy scale.

High confidence is desirable only if it is accompanied by high accuracy. Consequently, calibration was examined. Calibration measures the correlation between accuracy and confidence so that a perfectly calibrated decision maker is as accurate as s/he is confident. To determine calibration, correlation analyses between accuracy and confidence scores were carried out⁵⁴. The correlation coefficients are shown in Table 5.5. Groups are marginally well calibrated, that is, an increase in accuracy is accompanied by a corresponding increase in confidence. A marginally significant inverse correlation between accuracy and confidence indicates that individual auditors are miscalibrated. Thus, H₃ is supported.

⁵⁴ Prior calibration studies (e.g., Mladenovic and Simnett, 1994; Dilla et al., 1991; Yates, 1991) use either the Brier model (Brier, 1950, cited in Yates, 1991, 38-45) or the Oskamp model (Oskamp, 1962, cited in Mladenovic and Simnett, 1994, 188-189) to derive the calibration score. These models assume that each individual participant performs many repetitions of the task. In Mladenovic and Simnett (1994) and Dilla et al. (1991), the participants performed 48 and 40 going concern predictions respectively. When used in a single task situation, however, the results are unstable (Yates, 1991, 38-40). To overcome this problem, Oskamp (1965) and Pincus (1991) analysed the significance of and the direction of the difference between the individual's accuracy and confidence scores. As the participants in this study performed only one task, this latter model is considered more suitable.

Table 5.5
The Correlation Between Accuracy and Confidence
Using the Continuous Accuracy Scale

	Correlation Coefficient	p
Audit groups (n = 64)	.20	.055
Individual auditors (n = 75)	-.15	.100

Calibration between accuracy and confidence was also examined for the dichotomous judgment. The mean confidence scores are shown in Table 5.6. Consistent with Pincus (1991), t-tests were used to compare the confidence between cells. The results show that accurate audit groups are marginally more confident than inaccurate audit groups ($t = 1.53$, $p = .066$), and the former are also marginally more confident than accurate individual auditors. However, accurate individual auditors are not more confident than inaccurate individual auditors as well as inaccurate audit groups. These results also support H_3 .

Table 5.6
The Relation Between Accuracy and Confidence
Using the Dichotomous Scale

	Groups	Individuals
Inaccurate	65.14 (15.28)	66.34 (15.99)
Accurate	70.54 (13.14)	64.32 (18.48)

Effects of Counterexplanation

H₄ states that belief revision for auditors who counterexplained will be higher than for auditors who explained. To test this hypothesis, belief revision is calculated by taking the absolute difference between initial and final accuracy judgments that were measured using the 9-point scale. The average belief revision for counterexplanation was higher than belief revision for explanation (.73 [sd .98] [counterexplanation] and .49 [sd .78] [explanation]) (panel A, Table 5.7). ANOVA analyses reveal the difference to be marginally significant ($F = 2.59, p = .055$)⁵⁵. There is no significant main effect for groups/individuals and no significant interaction effect.

⁵⁵ One group and three individuals did not complete this section of the experiment, so the rest of the hypotheses are analysed with 64 groups and 75 individuals.

Table 5.7
The Effects of Explanation Mode and Groups/Individuals on Belief Revision*

<u>Panel A - Mean (SD)</u>	<u>Explanation Mode</u>		<u>Total</u>
	<u>Explanation</u>	<u>Counter-explanation</u>	
<u>Groups/individuals</u>			
Audit groups	.62 (.81) n = 34	.69 (.91) n = 30	.65 (.85) n = 64
Individual auditors	.38 (.75) n = 37	.76 (1.04) n = 38	.61 (.92) n = 75
Total	.49 (.78) n = 71	.73 (.98) n = 68	.57 (.89) n = 139

<u>Panel B - ANOVA Results</u>					
	<u>SS</u>	<u>MS</u>	<u>df</u>	<u>F</u>	<u>p</u>
Main effects:					
Groups/individuals (A)	.23	.23	1	.29	.274
Explanation mode (B)	2.02	2.02	1	2.59	.055
Interaction effect:					
A x B	.84	.84	1	1.32	.151
Residual	3.09	1.03	135		

*Belief revision is measured as the absolute difference between first and last accuracy judgments using the 9-point Likert-type scale.

ANCOVA analyses show that audit experience, effort expended and task difficulty (used as covariates) did not have a significant effect on belief revision (Table 5.8). Task difficulty and effort expended similarly did not have an effect on belief revision. The main effect for explanation mode remains significant at the $p < .1$ level, and H_4 is supported.

Table 5.8
ANCOVA Analyses of Belief Revision*

	<u>SS</u>	<u>MS</u>	<u>df</u>	<u>F</u>	<u>p</u>
Covariate:					
Audit experience	.20	.20	1	.26	.307
Effort expended	.49	.49	1	.62	.217
Task difficulty	.23	.23	1	.29	.296
Main effects:					
Groups/individuals (A)	.26	.26	1	.32	.286
Explanation mode (B)	1.88	1.88	1	2.37	.063
Interaction effect:					
A x B	.72	.72	1	.90	.174
Residual	104.10	.80	132		

*Belief revision is measured as the absolute difference between first and last accuracy judgments using the 9-point Likert-type scale.

To test the robustness of counterexplanation, H_5 , H_6 and H_7 are tested. H_5 states that auditors who counterexplained inaccurate judgments are more accurate than auditors who explained inaccurate judgments. To test this hypothesis, I consider the final judgments of only those auditors who made inaccurate judgments by dichotomising the data into two groups (accurate participants and inaccurate participants) using the mean accuracy score (4.85)⁵⁶. The mean of the final accuracy score of the inaccurate participants who explained was 2.97 (sd 1.32) (initial judgment is 2.77 [sd .85]), and the mean of the final accuracy scores of the inaccurate participants who counterexplained was 3.84 (sd 1.45) (initial judgment is 3.19 [sd .64]) (panel A, Table 5.9). ANCOVA analyses with the initial accuracy score as a covariate were carried out. The results show that explanation mode had a marginally significant effect on the accuracy of auditors ($F = 2.52$, $p = .064$) (panel B, Table 5.9). Inaccurate participants who explained are marginally less accurate than inaccurate participants who counterexplained;

⁵⁶ Dichotomising the data set using the median (4) yields similar results. This is because no participants have accuracy scores of between 4 and 4.95. Diagnostic check reveals no significant difference in

counterexplaining inaccurate judgments leads to more accurate judgments, and there is weak support for H₅. No significant main effect for groups/individuals and no significant interaction effect were observed.

Table 5.9
The Effects of Explanation Mode and Groups/Individuals on Inaccurate Judgments*

<u>Mean (SD)</u>	<u>Explanation Mode</u>		<u>Total</u>
	<u>Explanation</u>	<u>Counter-explanation</u>	
<u>Groups/Individuals</u>			
Audit groups	3.08 (1.64) n = 13	3.82 (1.54) n = 12	3.42 (1.61) n = 24
Individual auditors	2.29 (1.14) n = 24	3.85 (1.44) n = 23	3.37 (1.37) n = 47
Total	2.97 (1.32) n = 37	3.84 (1.45) n = 36	3.39 (1.44) n = 71

Panel B - ANCOVA Results

	<u>SS</u>	<u>MS</u>	<u>df</u>	<u>F</u>	<u>p</u>
Covariate:					
Initial judgment	48.0	48.0	1	34.05	.000
Main effects:					
Groups/individuals (A)	.5	.5	1	.53	.235
Explanation mode (B)	3.6	.36	1	2.52	.064
Interaction effect:					
A x B	.1	.1	1	.06	.400
Residual	93.0	.14	66		

*Belief revision is measured as the absolute difference between first and last accuracy judgments using the 9-point Likert-type scale.

Additional analyses were carried out with audit experience, effort expended and task difficulty as covariates. These three variables did not have a significant effect on the

initial accuracy between those who explained and those who counterexplained (explanation = 4.8, counterexplanation = 4.9) ($t = .32$, $p = .75$ [two-tailed]).

judgments of inaccurate participants (Table 5.10). The main effect for explanation mode remains significant at the $p < .10$ level.

Table 5.10
The Effects of Explanation Mode, Groups/Individuals and Audit Experience on
Inaccurate Judgments*

ANCOVA Analyses of Accuracy					
	SS	MS	df	F	p
Covariates:					
Initial judgment	35.22	35.22	1	23.70	.000
Audit experience	.00	.00	1	.00	.487
Effort expended	.41	.41	1	.28	.300
Task difficulty	.03	.03	1	.02	.445
Main effects:					
Groups/individuals (A)	.28	.28	1	.19	.333
Explanation mode (B)	3.95	3.95	1	2.66	.054
Interaction effect:					
A x B	.24	.24	1	.16	.346
Residual	92.13	.49	62		

*Belief revision is measured as the absolute difference between first and last accuracy judgments using the 9-point Likert-type scale.

In testing judgment techniques, it is important to consider both their potential negative effects as well as their positive effects. To examine the potential negative effect(s) of counterexplanation, H_6 examines whether counterexplaining accurate judgments will lead to a reduction in accuracy. To test this hypothesis, I consider only the accurate participants who explained with those who counterexplained. The mean accuracy scores of participants after explaining and counterexplaining are 7.21 (.89) and 6.41 (1.34) respectively (panel A, Table 5.11). ANCOVA analyses, using the initial accuracy score as a covariate show this difference to be significant ($F = 7.49, p = .004$)⁵⁷ (panel B, Table 5.11). Counterexplaining accurate judgments leads to a significant downward revision in accuracy so that participants who counterexplained are significantly less accurate than participants who explained, and H_6 is supported.

Table 5.11
The Effects of Explanation Mode and Groups/Individuals on Accurate Judgments*

<u>Panel A - Mean (SD)</u>	<u>Explanation Mode</u>		
	<u>Explanation</u>	<u>Counter-explanation</u>	<u>Total</u>
<u>Groups/individuals</u>			
Audit groups	7.44 (.77) n = 21	6.67 (1.29) n = 19	7.08 (1.11) n = 40
Individual auditors	6.85 (.99) n = 13	6.07 (1.37) n = 15	6.43 (1.25) n = 28
Total	7.21 (.89) n = 34	6.41 (1.34) n = 34	6.81 (1.20) n = 68

<u>Panel B - ANCOVA Results</u>					
	<u>SS</u>	<u>MS</u>	<u>df</u>	<u>F</u>	<u>p</u>
Covariate:					
Initial judgment	46.8	46.8	1	68.56	.000
Main effects:					
Explanation mode (A)	5.1	5.1	1	7.49	.004
Groups/individuals (B)	.5	.5	1	.73	.197
Interaction effect:					
A x B	1.6	1.6	1	2.29	.067
Residual	43.0	.7	63		

*Belief revision is measured as the absolute difference between first and last accuracy judgments using the 9-point Likert-type scale.

H₇ predicts that counterexplanation will have a more negative effect on individual auditors who make accurate judgments than on audit groups who make accurate judgments, so that audit groups which explained will be more accurate than individual auditors who explained. No significant main effect for groups/individuals is observed (F = .73, p = .197) (panel B, Table 5.11). A marginally significant interaction effect between groups/individuals and explanation mode is observed (F = 2.29, p = .067)

⁵⁷ Dichotomising the data set using the median (4) yields similar results.

(panel B, Table 5.11). The interaction effect is shown on Figure 5.2. Planned comparisons show a marginally significant difference between the group/counterexplanation condition and the individual/counterexplanation condition ($t = 1.32$, $p = .097$), and a significant difference between the group/explanation and individual/explanation conditions ($t = 1.96$, $p = .029$). Therefore, audit groups that explained are significantly more accurate compared to individual auditors who explained, and audit groups that counterexplained are marginally more accurate than individual auditors who counterexplained. Thus, H_7 is supported.

Additional analyses with audit experience, effort expended and task difficulty as covariates are also carried out. The audit experience and task difficulty effects are not significant at the $p < .1$ level (Table 5.12). Effort expended had a significant effect on the final judgment for these participants. Greater effort should lead to higher accuracy. The main effect for explanation mode and the interaction effect remain significant at the $p < .05$.

Figure 5.2
Interaction Effect Between Explanation Mode and Groups/Individuals on Accuracy

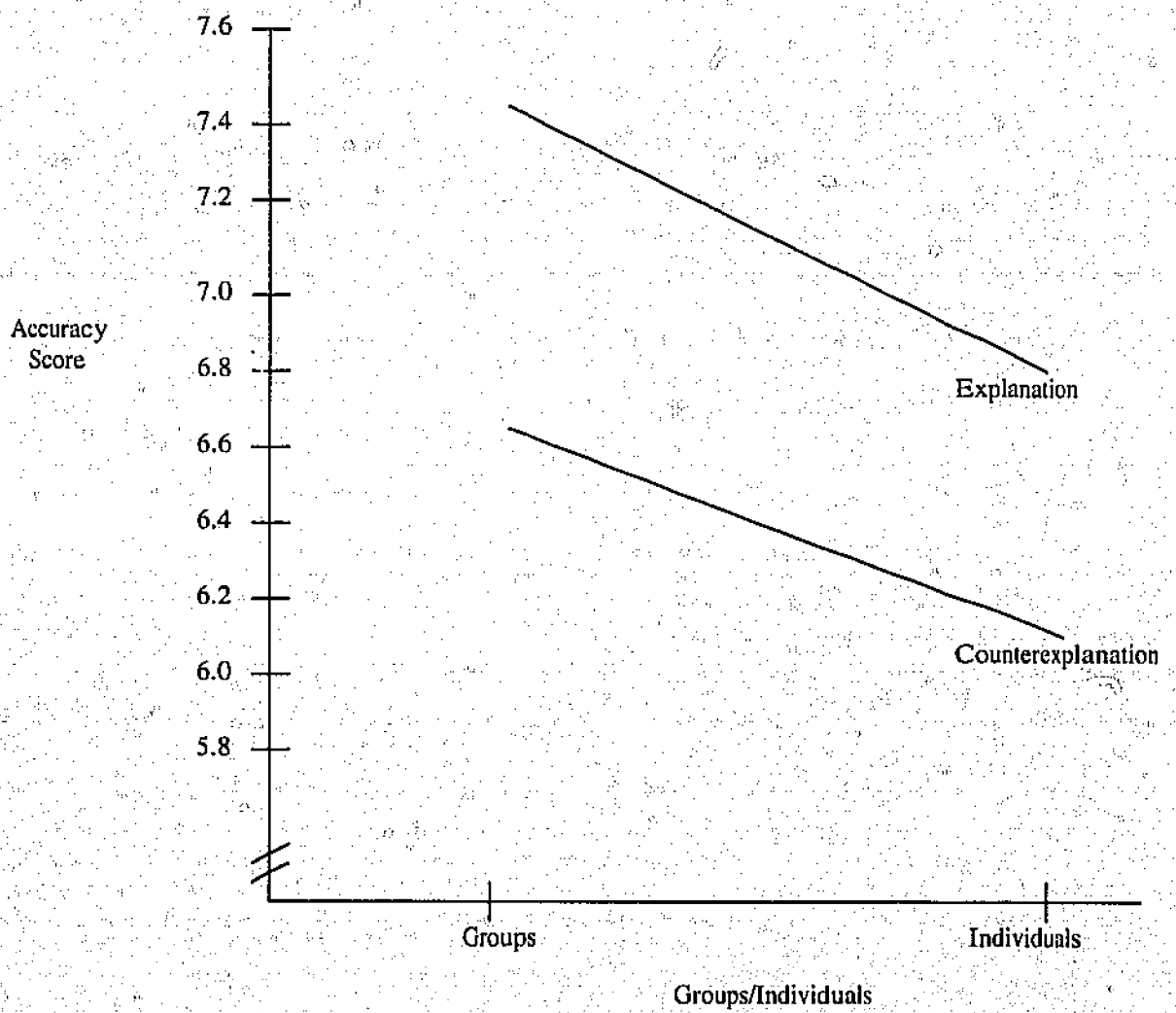


Table 5.12
The Effects of Explanation Mode, Groups/Individuals and Audit Experience on
Accurate Judgments*

ANCOVA Analyses of Accuracy					
	SS	MS	df	F	p
Covariates:					
Initial judgment	30.71	30.71	1	45.42	.000
Audit experience	.06	.06	1	.10	.378
Effort expended	2.01	2.01	1	2.98	.044
Task difficulty	.47	.47	1	.69	.205
Main effects:					
Explanation mode (A)	.19	.19	1	.28	.300
Groups/individuals(B)	2.07	2.07	1	3.07	.042
Interaction effect:					
A x B	1.88	1.88	1	2.78	.050
Residual	40.57	.68	60		

*Belief revision is measured as the absolute difference between first and last accuracy judgment using the 9-point Likert-type scale.

Analyses of belief revision measured on the dichotomous scales are also carried out. When audit groups were required to explain, two inaccurate groups became more accurate (Table 5.13), and this change is marginally significant ($z = -1.41, p = .073$). The accurate groups did not alter their judgment. Four inaccurate audit groups which counterexplained made accurate final judgment while the accurate groups did not alter their final judgment after counterexplaining ($z = -1.73, p = .041$). Individual auditors who were required to explain their judgment did not alter them. When individual auditors counterexplained their judgment, only one individual changed his/her judgment, and this is from an inaccurate judgment to an accurate one ($z = -1.00, p = .158$). These results show that both explanation and counterexplanation are more effective on audit groups than on individual auditors. These results support H_5 but do not provide support for H_6 . There is partial support for H_7 as only groups made significant positive revision when required to explain for counterexplain.

Table 5.13
Belief Revision Using the Dichotomous Scale

<u>Source</u>	<u>Type of Judgment</u>	<u>Initial Judgment (n)</u>	<u>Final Judgment (n)</u>	<u>z</u>	<u>p</u>
Group/ explanation	Inaccurate	15	13	-1.41	.073
	Accurate	19	21		
Group/ counterexplanation	Inaccurate	13	9	-1.73	.041
	Accurate	18	21		
Individual/ explanation	Inaccurate	26	26	.000	1.00
	Accurate	12	12		
Individual/ counterexplanation	Inaccurate	32	31	-1.00	.158
	Accurate	7	8		

Confidence Revision

H_8 states that explanation mode has an effect on confidence revision, so that participants who explained will revise their confidence upwards whereas those who counterexplained will revise their confidence downwards. To test H_8 , confidence revision is calculated by taking the percentage change (relative difference) between the final confidence and the initial confidence so that a negative score indicates a decrease in confidence and a positive score indicates an increase in confidence. The confidence scores used for this analyses represent the confidence expressed by participants in relation to their judgment on the 9-point Likert-type accuracy scales. The average confidence revision score for participants who explained was 1.19 (sd 11.90) and -4.70

(sd 12.63)⁵⁸ (panel A, Table 5.14)⁵⁹ for those who counterexplained, and this difference is significant ($F = 9.64$, $p = .001$) (panel B, Table 5.14). Those who explained revised their confidence upwards, and the opposite was true for those who counterexplained, and H_8 is supported. This result supports prior findings of Hoch (1985) and Koriat et al. (1980).

⁵⁸ There is a significant difference between initial and final confidence for all participants ($F = 8.18$, $p = .005$ [two-tailed]).

⁵⁹ Diagnostic check reveals no significant difference in initial confidence between those who explained and those who counterexplained (71.6% [explanation]; 69.8% [counterexplanation]) ($t = .68$, $p > .10$).

Table 5.14
The Effects of Explanation Mode and Groups/Individuals on Confidence Revision*

<u>Panel A - Mean (SD)</u> <u>Groups/Individuals</u>	<u>Explanation Mode</u>		<u>Total</u>
	<u>Explanation</u>	<u>Counter-explanation</u>	
Audit groups	5.38 (14.17) n = 34	-5.83 (14.81) n = 30	.13 (15.45) n = 64
Individual auditors	-2.46 (8.05) n = 37	-3.82 (10.70) n = 38	-3.14 (9.43) n = 75
Total	1.19 (11.90) n = 71	-4.70 (12.63) n = 78	-1.67 (12.67) n = 139

<u>Panel B - ANOVA Results</u>					
	<u>SS</u>	<u>MS</u>	<u>df</u>	<u>F</u>	<u>p</u>
Main effects:					
Explanation mode (A)	1386.7	1386.7	1	9.64	.001
Groups/individuals (B)	298.2	298.2	1	2.07	.076
Interaction effect:					
A x B	852.0	852.0	1	5.92	.008
Residual	19853.6	143.9	135		

*Confidence revision is the percentage change (relative difference) between initial confidence and final confidence of participants expressed in relation to their judgments on the 9-point Likert-type accuracy scales.

An interaction effect between groups/individuals and explanation mode on confidence revision is predicted in H₉. Individual auditors, being less confident, will react differently to explaining and counterexplaining compared to audit groups. Audit groups who explained will become more confident than individual auditors who explained, whereas audit groups who counterexplained will become less confident than individual auditors who counterexplained. The latter is because counterexplaining would result in groups processing more opposing arguments than counterexplaining by individual auditors. There is a marginally significant main effect of groups and individuals on

confidence revision ($F = 2.07, p = .076$) (panel B, Table 5.14). The interaction effect is significant ($F = 5.92, p = .008$) (panel B, Table 5.14). Figure 5.3 shows the interaction effect. Planned comparisons indicate that when required to counterexplain, the mean confidence revision of audit groups is -5.83 (sd 14.87) and -3.82 (sd 10.70) for individual auditors, but the difference is not significant ($t = .51, p = .305$). When required to explain, the mean confidence revision of audit groups was 5.38 (sd 14.17) and -2.46 (sd 8.05) (panel A, Table 5.11) for individual auditors and this difference is significant ($t = 2.92, p = .002$). Audit groups increased their confidence whereas individual auditors decreased their confidence when required to explain, and H_9 is partially supported.

Additional ANCOVA analyses show that audit experience did not have an effect on confidence revision ($F = .34, p = .282$) (Table 5.15). Effort expended similarly did not have a significant effect on confidence revision. However, task difficulty had a significant effect on confidence revision. Those who found the task more difficult made a larger negative revision to their confidence rating compared to those who found the task less difficult. The former were probably less certain of their judgment compared to the latter. The main effect for explanation mode and the interaction effect on confidence revision remains significant at the $p < .01$ and $.05$ levels respectively.

Figure 5.3
Interaction Effect Between Explanation Mode and Group/Individuals on Confidence Revision

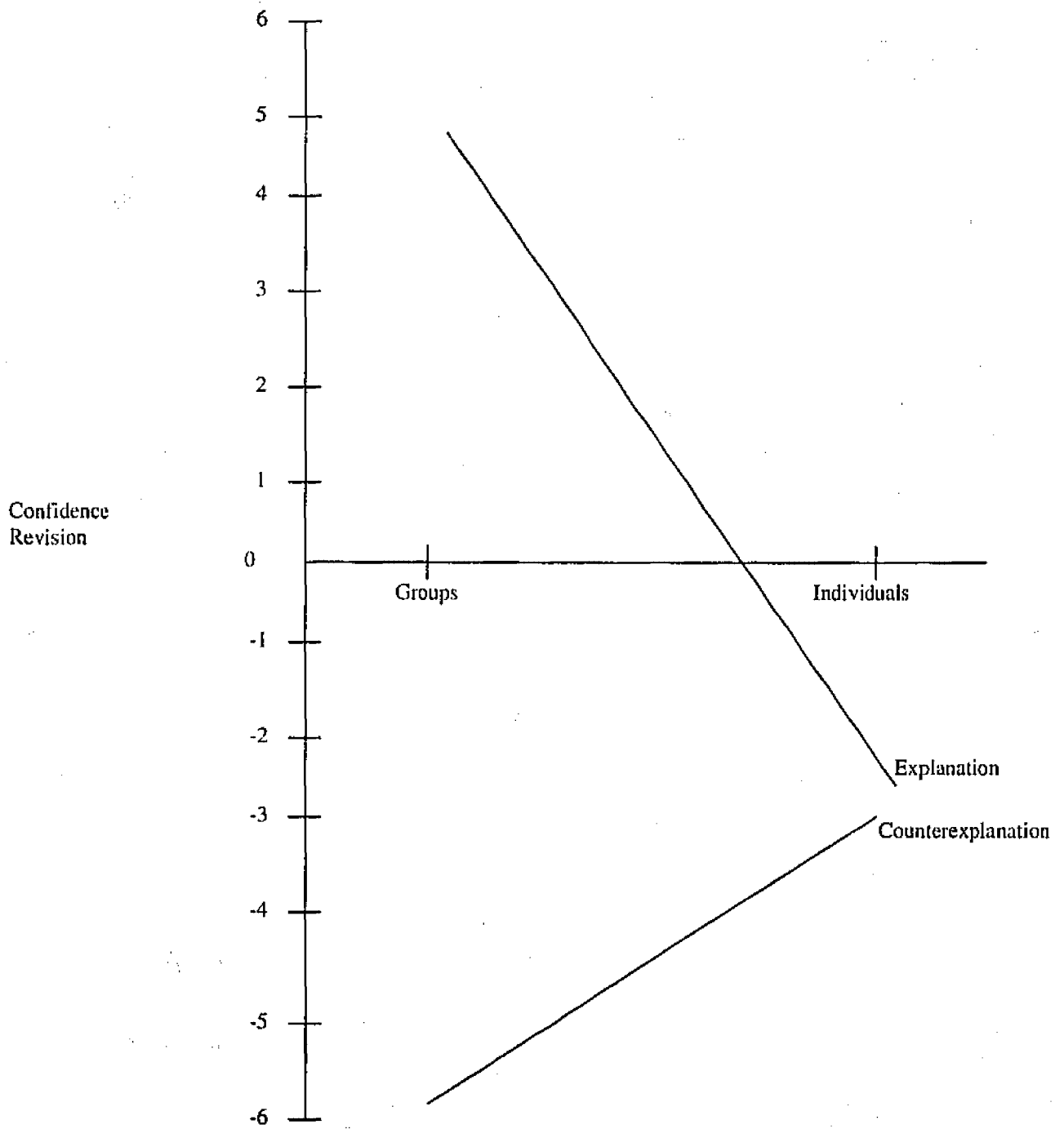


Table 5.15
The Effects of Explanation Mode, Groups/Individuals and Audit Experience on
Confidence Revision*

ANCOVA Analyses of Confidence Revision					
	SS	MS	df	F	p
Covariate:					
Audit experience	48.02	48.02	1	.34	.282
Effort expended	43.68	43.68	1	.31	.291
Task difficulty	420.30	420.30	1	2.93	.044
Main effects:					
Groups/individual (A)	128.06	128.06	1	.89	.173
Explanation mode (B)	1180.92	1180.92	1	8.23	.002
Interaction effect:					
A x B	927.33	927.33	1	6.47	.006
Residual	19220.76	143.44	134		

*Confidence revision is the percentage change (relative difference) between initial confidence and final confidence of participants expressed in relation to their judgments on the 9-point Likert-type accuracy scales.

Interference

H₁₀ predicts that interference created by causal explanations would cause auditors to rate the task of counterexplaining as significantly more difficult than the task of explaining. The mean (sd) difficulty scores are 5.12 (1.81) (explanation) and 5.74 (1.88) (counterexplaining) (panel A, Table 5.16). ANOVA analyses show that the difference between the two difficulty scores is significant (F = 3.91, p = .025) (panel B, Table 5.16). The task of counterexplaining was perceived as significantly more difficult than the task of explaining, and H₁₀ is supported.

Table 5.16
The Effects of Explanation Mode and Groups/Individuals on Difficulty of
Explaining/Counterexplaining

<u>Panel B - Mean (SD)</u>	<u>Explanation</u>	<u>Counter- explanation</u>	<u>Total</u>
Group	5.15 (1.89) n = 34	5.55 (1.43) n = 30	5.00 (1.94) n = 64
Individuals	5.62 (1.27) n = 37	5.96 (1.49) n = 38	5.76 (1.73) n = 75
Total	5.12 (1.81) n = 71	5.74 (1.88) n = 68	5.42 (1.86) n = 139

<u>Panel B - ANOVA Results for Difficulty Score</u>					
	<u>SS</u>	<u>MS</u>	<u>df</u>	<u>F</u>	<u>p</u>
Main effects:					
Groups/individuals (A)	18.51	18.51	1	5.60	.019
Explanation mode (B)	12.90	12.90	1	3.91	.025
Interaction effect:					
A x B	.59	.59	1	.18	.336
Residual	445.91	3.30	135		

H₁₁ indicates that auditors who explained will generate more explanations than auditors who counterexplained. The explanation participants generated 4.86 explanations (sd 2.64), while those who counterexplained generated fewer counterexplanations (mean 4.00, sd 1.69) (panel A, Table 5.17). ANOVA analyses show the difference between these two scores to be significant (F = 5.65, p = .009) (panel B, Table 5.17), and there is support for H₁₁.

Table 5.17
The Effects of Explanation Mode and Groups/Individuals on the Number of Explanations and Counterexplanations Generated

<u>Panel A - Mean (SD)</u>	<u>Explanation</u>	<u>Counter-explanation</u>	<u>Total</u>
Groups	5.91 (1.29) n = 34	6.07 (1.55) n = 30	5.05 (2.35) n = 64
Individuals	5.85 (1.29) n = 37	5.79 (1.22) n = 38	3.95 (2.98) n = 75
Total	4.86 (2.64) n = 71	4.00 (1.69) n = 68	4.44 (2.26) n = 139

<u>Panel B - ANOVA Results for Number of Reasons</u>					
	<u>SS</u>	<u>MS</u>	<u>df</u>	<u>F</u>	<u>p</u>
Main effects:					
Groups/individuals (A)	39.28	39.28	1	8.33	.002
Explanation mode (B)	26.61	26.61	1	5.65	.009
Interaction effect:					
A x B	5.89	5.89	1	1.25	.133
Residual	650.37	4.71	135		

H₁₂ and H₁₃ compare the effects of groups and individuals on interference. H₁₂ predicts that individuals will have higher difficulty scores than groups. The mean (sd) difficulty score for individuals was 5.76 (1.73), and the score for groups was 5.00 (1.94) (panel A, Table 5.16), and these two scores are significantly different (F = 5.60, p = .019) (panel B, Table 5.16), and H₁₂ is supported. In H₁₃, individual auditors are predicted to generate fewer explanations or counterexplanations than audit groups. Individual auditors generated an average of 3.95 (2.98) explanations or counterexplanations while audit groups generated an average of 5.05 (2.35) explanations or counterexplanations (panel A, Table 16). The difference between these two scores is significant (F = 8.33, p = .002), and there is support for H₁₃. The results for H₁₀, H₁₁, H₁₂ and H₁₃ support

Hoch's (1984) proposition that causal explanation creates interference, and makes it more difficult for decision makers to consider alternative explanations.

Additional analyses with audit experience, and effort expended as covariates are carried out with the difficulty score as the dependent variable. ANCOVA analyses show that these two variables had no significant effect on the difficulty score (panel A, Table 5.18). The main effects of explanation mode and groups/individuals remain significant at the $p < .05$. level (panel A, Table 5.18).

Additional ANCOVA analyses (with audit experience, task difficulty and effort expended as covariates) with the number of explanations and counterexplanations as the dependent variables were carried out. Audit experience did not have a significant effect on the number of items generated (panel B, Table 5.18). Effort expended and task difficulty did. An increase in the number of items generated caused participants to rate the perception of effort expended higher. Task difficulty and the number of explanations and counterexplanations generated were positively related. When the number of items generated increased, the perception of task difficulty also increased. The main effects of groups/individuals and explanation mode on the number of items generated remain significant at the $p < .05$ level.

Table 5.18
Difficulty of Explaining/Counterexplaining and Number of Explanations or Counterexplanations Generated

Panel A - ANCOVA Analyses for Difficulty Score					
	<u>SS</u>	<u>MS</u>	<u>df</u>	<u>F</u>	<u>p</u>
Covariate:					
Audit experience	.00	.00	1	.75	.489
Effort expended	4.98	5.98	1	1.45	.112
Main effects:					
Groups/individuals (A)	13.44	13.44	1	4.05	.023
Explanation mode (B)	15.27	15.27	1	4.60	.017
Interaction effect:					
A x B	.34	.34	1	.10	.325
Residual	431.99	3.32	133		
Panel B - ANCOVA Analyses for Number of Explanations or Counterexplanations					
	<u>SS</u>	<u>MS</u>	<u>df</u>	<u>F</u>	<u>p</u>
Covariate:					
Audit experience	1.63	1.63	1	.39	.276
Effort expended	23.74	23.74	1	5.65	.009
Task difficulty	72.70	72.70	1	17.29	.000
Main effects:					
Groups/individuals (A)	22.46	22.46	1	5.34	.011
Explanation mode (B)	17.95	17.95	1	4.27	.020
Interaction effect:					
A x B	6.39	6.39	1	1.52	.110
Residual	554.92	4.20	132		

CONCLUSIONS

This experiment examines the effects of audit groups and counterexplanation on fraud detection rates. The purpose of such an examination is to determine whether audit fraud detection rates can be improved. An experiment that manipulated groups/individuals and two explanation modes was carried out. Participants were required to evaluate inventory case materials which contained material misstatements by management. The judgment was made twice with a(n) intervening explanation or counterexplanation.

The first section of the experiment tests the relative performance of audit groups and individual auditors. This experiment hypothesises that audit group judgments are more

accurate than individual auditor judgments, and that groups are not only more confident, but they are also better calibrated. In sum, the analyses addressing accuracy found that audit groups are significantly more accurate than individual auditors. Translated to fraud detection, groups are more effective in detecting fraud than individuals. This result contradicts prior auditing literature which found that audit groups do not outperform individual auditors, e.g., Trotman and Yetton (1985) and Trotman et al. (1983) and supports Trotman (1985). Given the confounding in Trotman (1985), this result provides better support for the superiority of audit groups over individual auditors in the detection of fraud. While interacting audit groups are costly to assemble, the results suggest that, in some cases, the benefits may outweigh the costs. The benefits of requiring entry level auditors to work in groups are two-fold: first, their judgments are made more effective by the group process, and second, such audit groups are relatively low cost. The results also show that audit experience has a positive effect on the accuracy and confidence of the participants.

Audit groups in this experiment are more confident than individual auditors. As high confidence is only desirable when it is accompanied by similarly high accuracy, calibration is examined. The results provide evidence that audit groups are marginally better calibrated than individual auditors. Among audit groups, high confidence is matched with high accuracy, whereas individual auditors are miscalibrated. The least accurate are more confident while the more accurate are less confident. Confidence determines subsequent action (Norman, 1975), so that a confident auditor acts on his/her judgment and a less confident auditor does not. When confidence and accuracy rates are mismatched as was the case for individual auditors, those who decided the inventory balance was fairly stated would take subsequent actions consistent with their high level

of confidence, whereas those who concluded that the inventory balance was not fairly stated, would continue to evaluate the data set because of their low confidence. Both actions are undesirable, the former leads to ineffective, and very often, costly judgments while the latter is inefficient. This study shows that the calibration of auditors may be improved through the use of audit groups.

Counterexplanation makes opposing arguments more salient to the auditor, and should result in higher belief revision than explanation. In the case of auditors, the need to maintain cognitive consistency is essential because inconsistency may be interpreted by superiors as incompetence (Church, 1990). However, if a counterexplanation is required as part of the formal evidence accumulation process, is documented and is known to the superior, such a need to maintain cognitive consistency is diffused. These results show that counterexplanation results in higher belief revision compared to explanation. Used on auditors who make inaccurate judgments, counterexplanation might cause auditors to switch from less accurate judgments to more accurate ones. The results show this to be the case, auditors who counterexplained inaccurate judgments were more accurate than auditors who explained inaccurate judgments. Causal explanations are generated as a by-product of decision-making, and therefore, requiring auditors to explain a judgment serves only to reinforce what is already available in memory. This is evidenced by the higher confidence of auditors who were required to explain. Counterexplanation forces auditors to consider opposing arguments that may have been overlooked in the initial judgment process, hence those auditors who counterexplained recorded lower confidence than auditors who explained.

Counterexplanation has been found to be an effective technique in social psychology for correcting over-confidence (Koriat et al., 1980) as most participants observed in social psychology are over-confident. Prior auditing literature found auditors to be under-confident (Mladenovic and Simnett, 1994; Tomassini et al., 1982). Counterexplanation is only desirable if it does not reduce the confidence of auditors further, that is, the uncertainty created by counterexplaining should not lead auditors to substitute their accurate judgment for a less accurate one. This study found significant difference in accuracy between explaining and counterexplaining accurate judgments. The former leads to significantly higher accuracy than the latter whereas the latter leads to lower accuracy. The negative effect of counterexplanation is higher for individual auditors than audit groups with the difference approaching significance. Care should be exercised in requiring a counterexplanation. While it is effective on those who make inaccurate judgments, it is counter-productive on those who make accurate judgments, and in particular, individual auditors. At the time the counterexplanation is required, the audit superior cannot tell whether the audit subordinate's judgment is accurate or inaccurate, so counterexplanation should be sparingly prescribed.

Consistent with prior studies (e.g., Hoch, 1985; Koriat et al., 1980), counterexplanation results in a decrease in confidence whereas explanation leads to increased confidence. Audit group confidence is more affected by explanation/counterexplanation than individual auditor confidence, the former decreased confidence more than the latter when asked to counterexplain (but this difference is not significant) and increased confidence significantly more than individual auditors when asked to explain. Hoch (1985) found that causal explanations cause interference, making it more difficult for participants to subsequently counterexplain. Similar results are reported in this study.

Participants who explained, generated more explanations while participants who counterexplained generated fewer counterexplanations. Similarly, the task of explaining was significantly less difficult to perform than the task of counterexplaining. Audit groups, on the other hand, could perform both tasks significantly more efficiently than individual auditors; they found the task of explaining/counterexplaining less difficult than individual auditors, and they generated more explanations and counterexplanations than individual auditors. This suggests that the negative aspects of counterexplanation may be mitigated by the group process. Additional analyses show that audit experience (except in the case of belief revision) does not have a significant effect on the results.

LIMITATIONS

The limitations from Experiments One and Two apply to this experiment as well. Due to the large numbers of participants that are required for this experiment, I was constrained to use auditors who were relatively inexperienced. While the results show that counterexplanation is an effective judgment tool on auditors, it is unclear whether the positive and negative aspects of counterexplanation could be replicated with more experienced auditors. Finally, while the results of this experiment show that audit groups are more accurate and more confident than individual auditors, it does not explain the cause(s) of these differences in performance.

CHAPTER 6 CONCLUSIONS

This chapter presents an overview of the study and a general summary of the research findings. The specific conclusions and discussions of the individual hypotheses are detailed under their respective chapters.

OVERVIEW

The objectives of this study are to examine three methods of improving audit judgment and to investigate factors that may affect their use on auditors. Experiment One examined the effect of inoculation and audit experience on the framing of internal controls. Auditors were asked to assess the internal controls in the inventory area in a between-subjects experiment which manipulated inoculation as well as audit experience. ANOVA was adopted to analyse the results. One hundred and thirty-nine participants took part in the experiment.

Experiment Two examined the differences between audit group judgments and individual auditor judgments and analysed these differences based on groups' and individuals' self-reported cue usage, cue selection and cue weighting. Auditors were asked to assess whether a hypothetical client was a going-concern. A between-subjects experiment that manipulated audit groups and individual auditors was carried out. The results were analysed using t-tests. In all, the experiment involved 172 participants.

The third experiment examined audit groups' and individual auditors' effectiveness in fraud detection. In addition, the effectiveness of counterexplanation was also examined.

Auditors were asked to assess whether a client's inventory balance was fairly stated. Two decision modes - audit groups and individual auditors, and two explanation modes - explanation and counterexplanation were manipulated. ANOVA, and t-tests were employed to analyse the results. Two hundred and seventy-one participants took part in the experiment.

This study hypothesises that inoculation would have an effect on framing and audit experience. The inoculation treatment is expected to reduce the effects of framing. Participants with different levels of audit experience are expected to react differently to the inoculation treatment. An inverted-U-shaped relation between audit experience and responsiveness to the inoculation treatment is expected. Those without audit experience and those with 19 to 60 months audit experience are not expected to be affected by the inoculation treatment, whereas those with one to 18 months audit experience are. The results showed no framing effect and consequently, no interaction effect between framing and inoculation was observed. Consistent with the hypothesis, auditors with one to 18 months audit experience were the most responsive to the inoculation treatment. The other two groups were not. The order by which auditors wrote positive and negative arguments affected their evaluation of the internal controls, and a primacy effect was observed.

It is hypothesised that audit group judgments are more superior to individual auditor judgments. Group judgments are expected to be more conservative than individual judgments. In addition, consensus, accuracy and calibration are also tested. The results showed that, not only were audit groups more accurate in their judgments, but the group process was also able to suppress outlier behaviour among individual auditors so that

consensus was higher among audit groups than individual auditors. Audit groups were also more confident than individual auditors. As confidence alone should not be taken as a measure of judgment quality, calibration was examined. Audit groups were better calibrated than individual auditors.

To account for audit groups' superior judgments, self-reported cue usage, cue selection and cue weighting are examined. The results showed that audit groups outperformed individual auditors because audit groups made use of a larger number of information cues than individual auditors. Audit groups selected more going-concern problem indicators as well as mitigating factors compared to individual auditors. However, both groups and individuals were equally proficient at weighting the selected cues.

This study hypothesises that counterexplanation will lead to higher belief revision. To test the effectiveness of counterexplanation, it is hypothesised that counterexplaining inaccurate judgments will lead to higher accuracy and counterexplaining accurate judgments will result in lower accuracy. The results showed that compared to explanation, counterexplanation led to higher belief revision. More importantly, they showed that counterexplaining inaccurate judgments led to higher accuracy and counterexplaining accurate judgments led to lower accuracy. Support for the latter hypothesis showed that the use of counterexplanation may be counter-productive. Consistent with prior psychology studies (Hoch, 1985; Koriat et al., 1980), counterexplanation decreased confidence whereas explanation increased confidence (the latter was observed among audit groups only). Causal explanation caused interference making it more difficult for individual auditors to subsequently counterexplain. Individual auditors who counterexplained wrote less counterexplanations while

individual auditors who explained wrote more explanations. Audit groups performed both these tasks with relative ease.

Mixed results are reported on the interaction effect between audit experience and responsiveness to inoculation. Experiment One reports that auditors with different levels of audit experience did not respond in the same way to inoculation. This experiment which made use of an internal control evaluation task, reports that the relation between inoculation and audit experience takes on an inverted-U shaped function. In the second experiment, a going-concern assessment task was used, and additional analyses show that audit experience does not lead to superior judgments. Because the participants were novices, most of them would probably not have performed going-concern evaluations before, and consequently, going-concern judgment quality did not improve with audit experience. In more familiar tasks such as the one used in Experiment Three (which involved the audit of inventory), accuracy and confidence increased with audit experience. In the case of explanation/counterexplanation, audit experience led to improved accuracy and confidence.

LIMITATIONS

There are limitations to the experiments reported here. First, the case materials employed in these experiments comprised relatively less information than what auditors normally work with. However, it is unclear what effect(s) more detailed case materials would have on the judgment techniques examined and what limitations the use of simplified case materials would have on the external validity of the results of this study. Second, in trying to secure responses that the participants could justify, all participants

in this study were paid. This procedure did not, however, correspond to the audit environment where auditors would suffer penalties for not providing justifiable responses. Such penalties include financial losses (in promotion and pay rises as well as litigation costs) and loss of reputation. Third, due to the large numbers of participants that were required for Experiments Two and Three, I was constrained to use auditors who were relatively inexperienced. It is unclear whether experience has an effect on the results. Finally, while the results of Experiment Three show that audit groups are more accurate and more confident than individual auditors, it does not explain the cause(s) of these differences in performance.

FUTURE RESEARCH

Future research with respect to this study may take two avenues. First, this study adds to the limited pool of judgment techniques available for the use of auditors. Obviously, there is a need to explore other ways and means of improving auditor judgments (Messier, 1995).

Second, further examination of the techniques identified in this study is necessary. As this is the first study that examines inoculation in an auditing context, our knowledge of it is very limited. Similarly, the counter-productive aspect of counterexplanation requires further testing. An interesting extension to this study would be to examine counterexplanation in relation to auditors' asymmetrical loss function. Such an examination may provide insight into the robustness of counterexplanation.

Future studies could explain why audit groups outperform individual auditors. Various theories have been suggested in prior literature, and these have been reviewed in chapter

2. For example, information load theory is untested in an auditing environment. Further studies could manipulate two information load levels (high and low) between groups and individuals. Because of the groups' ability to process more information (Chalos and Pickard, 1985), it is expected that they would outperform the individuals at the high information load level. No difference in performance at the low information load level is expected.

In a complex task, groups should outperform individuals for three reasons. First, groups could recall more information (Stasser, 1992), so that vital cues which are forgotten or overlooked by a group member may be identified by other group members. Second, individual members have different abilities and processing capabilities and may process the various facets of the task differently. This results in varying types and amounts of information being stored in memory. When they pool their different store of information, the groups' performance is improved (Stasser, 1988; Stasser and Titus, 1985). Third, the process of anchoring and adjustment reduces the effectiveness of a judgment. As described in chapter 4, groups work with a wider range of anchors because individual members bring different anchors to the judgment process (Sniezek and Henry, 1992), and the quality of the judgment is improved. An experiment that manipulates two levels of task complexity (high and low) between groups and individuals should find that audit groups outperform individual auditors in the high complexity task. Both audit groups and individuals should perform equally well in the low complexity task.

Best member theory states that groups outperform individuals by identifying their best member and using that member's judgment as a proxy for the groups' judgment

(Trotman, 1985). Iselin (1991) found that group members did not attempt to discover their best member, and consequently, did not outperform the individuals. This observation could be due to the nature of Iselin's task which comprised ten bankruptcy predictions. His experimental design had participants performed four predictions⁶⁰ (which were not used in the analyses) followed by six other predictions⁶¹. The groups' inability to identify their best member may be due to a learning effect which caused the performance difference between the experts and novices to diminish. When the nature of the task is changed to the type of task used in Experiment Three, this learning effect is reduced. This allows the best member to outperform the rest of the group, and consequently, audit groups would be more accurate than individual auditors.

⁶⁰ The first four predictions were used to identify participants as novice and experts.

⁶¹ The experimental procedures are described in chapter 2.

References:

- Abdel-Khalik, A. R. (1983). The effects of certain internal audit variables on the planning of external audit programs. The Accounting Review, 58, 215-227.
- Abdel-Khalik, A. R., Snowball, D., Wragge, J. H. (1983). The effects of certain internal audit variables on the planning of external audit programs. The Accounting Review, 58 (2), 215-227.
- Ahlawat, S. S. (1992). Explanation effects in sequential belief-revision by audit groups and individuals. An unpublished dissertation, The Pennsylvania State University.
- American Institute of Certified Public Accountants (1988). The Auditor's Consideration of an Entity's Ability to Continue as a Going Concern. Statement on Auditing Standards No. 59. New York: AICPA.
- American Institute of Certified Public Accountants. (1982). Working Papers. Statement on Auditing Standards No. 41. New York: AICPA.
- American Institute of Certified Public Accountants. (1981). The Auditor's Considerations When a Question Arises About an Entity's Continued Existence. Statement on Auditing Standards No. 34. New York: AICPA.
- Anderson, C. (1982). Inoculation and counterexplanation: Debiasing techniques in the perseverance of social theories. Social Cognition, 1, 126-139.
- Anderson, C. Lepper, M. R., Ross, L. (1980). Perseverance of social theories: The role of explanation in the persistence of discredited information. Journal of Personality and Social Psychology, 39 (6), 1037-1049.
- Anderson, C. and Sechler, E. S. (1986). Effects of explanation and counterexplanation on the development and use of social theories. Journal of Personality and Social Psychology, 50 (1), 24-34.
- Anderson, J. C., Jennings, M., Kaplan, S. E., and Reckers, P. M. J. (1995). The effect of using diagnostic decision aids for analytical procedures on judges' liability judgments. Journal of Accounting and Public Policy, 14 (1), 33-62.
- Anderson, N. H., and Zalinski, I. (1991). Functional measurement approach to self-estimation in multiattribute evaluation. In N. H. Anderson (Ed.), Contributions to information integration theory, volume 1: Cognition. Hillsdale, NJ: Erlbaum: 145-185.
- Anderson, U., and Wright, W. (1988). Expertise and the explanation effect. Organizational Behavior and Human Decision Processes, 42, 250-269.
- Arens, A. A., Loebbecke, J. K., Best, P. J., and Shailer, G. E. P. (1992). Auditing in Australia: An intergrated approach (Revised 2nd edition). Sydney, Australia: Prentice-Hall.
- Arkes, H. R. (1991). Cost and benefits of judgment errors: Implications for debiasing. Psychological Bulletin, 110 (3), 486-498.
- Asare, S. K. (1992). The auditor's going-concern decision: Interaction of task variables and the sequential processing of evidence. The Accounting Review, 67 (2), 379-393.
- Ashton, A. H. (1985). Does consensus imply accuracy in accounting studies in decision making? The Accounting Review, 13 (4), 411-454.
- Ashton, R. H. (1992). Effects of justification and a mechanical aid on judgment performance. Organizational Behavior and Human Decision Processes, 52, 292-306.
- Ashton, R. H. (1983). Research in Audit Decision Making: Rationale, evidence, and implications. Research Monograph Number 6. Vancouver: The Canadian Certified General Accountants' Research Foundation.

- Ashton, R. H. (1974). An experimental study of internal control judgments, Journal of Accounting Research, Spring, 143-157.
- Ashton, R. H., and Ashton, A. H. (1990). Evidence-responsiveness in professional judgment: Effects of positive versus negative evidence and presentation mode. Organizational Behavior and Human Decision Processes, 46 (1), 1-19.
- Ashton, R. H., and Ashton, A. H. (1988). Sequential belief revision in auditing. The Accounting Review, 63 (4), 623-641.
- Ashton, R. H., Kleinmuntz, D. N., Sullivan, J. B., and Tomassini, L. W. (1988). Audit decision making. In Abdel-Khalik, A., and Solomon, I. (Ed.), Research Opportunities in Auditing: The Second Decade. American Accounting Association: Auditing Section.
- Australian Accounting Research Foundation. (1996). AUS 208 - Documentation. Caulfield, Victoria: Prentice-Hall.
- Australian Accounting Research Foundation (1996). Australian Auditing Standards No. 708 - Going Concern. Sydney: Prentice-Hall.
- Australian Accounting Research Foundation (1996). Australian Auditing Standards No. 502 - Audit Evidence. Caulfield, Victoria: Prentice-Hall.
- Bamber, M. E., and Bylinski, H. J. (1982). The audit team and the audit review process: An organizational approach. Journal of Accounting Literature, 1, 33-58.
- Baron, R. S., Monson, T. C., and Baron, P. H. (1973). Conformity pressure as a determinant of risk taking: Replication and extension. Journal of Personality and Social Psychology, 28, 406-413.
- Barrett, M., and O'Malley, J. (1976). Choice shifts and audit opinion. Symposium on Auditing Research. University of Illinois at Urbana-Champaign publication.
- Bateson, N. (1966). Familiarization, group discussion, and risk-taking. Journal of Experimental Social Psychology, 2, 119-129.
- Bazerman, M. (1984). The relevance of Kahneman and Tversky's concept of framing to organizational behavior. Journal of Management, 10 (3), 333-343.
- Bedard, J. C., Biggs, S. F., and Maroney, J. J. (1997). Group process gains and losses in analytical procedures. American Accounting Association/Accounting, Behavior and Organization Research Conference, Pittsburgh.
- Bernardi, R. A. (1994). Fraud detection: The effect of client integrity and competence and auditor cognitive style. Auditing: A Journal of Practice and Theory, 13, Supplement, 68-84.
- Bonner, S. E. (1991). Is experience necessary in cue measurement: The case of auditing tasks. Contemporary Accounting Research, 8 (1), 253-269.
- _____. (1990). Experience effects in auditing: The role of task-specific knowledge. The Accounting Review, 65 (1), 1-20.
- Brehmer, B. (1976). Social judgment theory and the analysis of interpersonal conflict. Psychological Bulletin, 83, 985-1003.
- Brier, G. W. (1950). Verification of forecasts expressed in terms of probability. Monthly Weather Review, 78 (1), 1-3.
- Brown, R. (1986). Social Psychology: 2 edition. New York: Collier Macmillan, 200-248.
- Butterworth, J. J. (1988). Expert assessment of examination grades. Journal of Operational Research Society, 39 (7), 653-659.
- Caplan, R., Vinokur, A., Price, R., and van Ryn, J. (1989). Job seeking, reemployment, and mental health: A randomized field experiment in coping with job loss. Journal of Applied Psychology, 74 (5), 759-769.

- Carroll, J. (1978). The effect of imagining an event on expectations for the event: An interpretation in terms of the availability heuristic. Journal of Experimental Social Psychology, 14, 88-96.
- Chalos, P. (1985). Financial distress: A comparative study of individual, model, and committee assessment. Journal of Accounting Research, 23 (2), 527-543.
- Chalos, P., and Pickard, S. (1985). Information choice and cue use: An experiment in group information processing. Journal of Applied Psychology, 634-641.
- Chewning, E. G. Jr., and Harrell, A. M. (1990). The effect of information load on decision maker's cue utilization levels and decision quality in a financial distress decision task. Accounting, Organization and Society, 15 (6), 527-542.
- Choo, F., and Trotman, K. T. (1991). The relationship between knowledge structure and judgments for experienced and inexperienced auditors. The Accounting Review, 66 (3), 464-485.
- Christensen-Szalanski, J. J. J., and Bushyhead, J. B. (1981). Physicians' use of probabilistic information in a real clinical setting. Journal of Experimental Psychology: Human Perception and Performance, 7, 928-935.
- Chung, J., and Monroe, G. (1996). The effects of counterexplanation and source of hypothesis on audit judgment. A paper presented at the Accounting Association of Australia and New Zealand Annual Conference.
- Church, B. K. (1991). An examination of the effect that commitment to a hypothesis has on auditors' evaluations of confirming and disconfirming evidence. Contemporary Accounting Research, 10 (1), 333-350.
- Church, B. K. (1990). Auditors' use of confirmatory processes. Journal of Accounting Literature, 81-112.
- Church, B. K., and Schneider, A. (1993). Auditors' generation of diagnostic hypotheses in response to a superior's suggestion: Interference effects. Contemporary Accounting Research, 10 (1), 333-350.
- Cohen, C. F. (1981). Person categories and social perception: Testing some boundaries of the processing effects of prior knowledge. Journal of Personality and Social Psychology, 40, 441-452.
- Crowley, A., and Hoyer, W. (1994). An integrative framework for understanding two-sided persuasion. Journal of Consumer Research, 20, 561-574.
- Davidson, A. S. (1994). Discussion of Fraud detection: The effect of client integrity and competence and auditor cognitive style. Auditing: A Journal of Practice and Theory, 13 (Suppl.), 85-89.
- Dess, G. G. (1987). Consensus on strategy formulation and organizational performance: Competitors in a fragmented industry. Strategic Management Journal, 8, 259-277.
- Dilla, W. N., Richard, G. F., Solomon, I., and Tomassini, L. A. (1991). Predictive bankruptcy judgments by auditors: A probabilistic approach. In Poneman, L. and Gabhat, D. (Eds), Auditing: Advances in Behavioral Research. Springer-Verlag.
- Doise, W. (1969). Intergroup relations and polarization of individual and collective judgments. Journal of Personality and Social Psychology, 12, 136-143.
- Dombrowski, R. F. (1993). Auditing students: A way to get meaningful practical experience. Issues in Accounting Education, 8 (2), 320-334.
- Dubrovsky, V. J., Kiesler, S., and Sethna, B. N. (1991). The equalization phenomenon: Status effects in computer-mediated and face-to-face decision making groups. Human Computer Interaction, 6, 119-146.

- Dunegan, K. J. (1996). Fines, frames and images: Examining formulation effects on punishment decisions. Organizational Behavior and Human Decision Processes, 68 (1), 58-67.
- Einhorn, H. J. (1974). Expert judgment: Some necessary conditions and an example. Journal of Applied Psychology, 562-71.
- Einhorn, H. J., and Hogarth, R. (1986). Judging probably cause. Psychological Review, 99, 3-19.
- Einhorn, H. J., Hogarth, R., and Klempner, E. (1977). Quality of group judgment. Psychological Bulletin, 84 (1), 158-172.
- Emby, C. (1994). Framing and presentation mode effects in professional judgment: Auditors' internal control judgments and substantive testing decisions. Auditing: A Journal of Practice and Theory, 13 (Suppl.), 102-115.
- Fischhoff, B. (1982). Debiasing. In Kahneman, D., Slovic, P., and Tversky, A., (Eds), Judgment Under Uncertainty: Heuristics and Biases. New York, NY: Cambridge University Press, 422-444.
- Fischhoff, R., Slovic, P. and Lichtenstein, S. (1977). Knowing with certainty: The appropriateness of extreme confidence. Journal of Experimental Psychology: Human Perception and Performance, 3, 552-564.
- Fisher, B. A., and Ellis, D. G. (1990). Small Group Decision Making: Communication and the Group Process. New York: McGraw Hill.
- Frisch, D. (1993). Reasons for framing effects. Organizational Behavior and Human Decision Processes, 54, 399-429.
- Gaumnitz, B. R., Nunamaker, T. R. Surdick, J. J. and Thomas, M. F. (1982). Auditor consensus in internal control evaluation and audit program planning. Journal of Accounting Research, 20 (2), 745-755.
- Gibbins, M. (1984). Propositions about the psychology of professional judgment in public accounting. Journal of Accounting Research, Spring, 103-125.
- Gibbins, M., and Emby, C. (1984). Evidence on the nature of professional judgment in public accounting. In A. R. Abdel-Khalik and I. Solomon (Eds): Auditing Research Symposium, 181-212. Champaign: University of Illinois.
- Goldstein, W., and Mitzel, H. C. (1992). The relative importance of relative importance: Inferring other people's preferences from relative importance ratings and previous decisions. Organizational Behavior and Human Decision Processes, 51, 382-415.
- Hall, J., and Watson, W. H. (1970). The effects of a normative intervention on group decision-making performance. Human Relations, 23, 299-317.
- Heiman, V. (1990). Auditors' assessments of the likelihood of error explanations in analytical review. The Accounting Review, 65 (4), 875-890.
- Hemphill, J. K. (1961). Why people attempt to lead. In L. Petrullo and B. M. Bass (Eds): Leadership and Interpersonal Behavior. New York: Holt.
- Hill, G. W. (1982). Group versus individual performance: Are N + 1 heads better than one? Psychological Bulletin, 91, 517-539.
- Hirt, E. R., and Sherman, S. J. (1985). The role of prior knowledge in explaining hypothetical events. Journal of Experimental Social Psychology, 21, 519-543.
- Ho, J. L. (1994). The effect of experience on consensus of going concern Judgments. Behavioral Research in Accounting, 6, 160-177.
- Hoch, S. (1985). Counterfactual reasoning and accuracy in predicting personal events. Journal of Experimental Psychology: Learning, Memory, and Cognition, 11 (4), 719-731.

- Hoch, S. (1984). Availability and inference in predictive judgment. Journal of Experimental Psychology: Learning, Memory and Cognition, 10, 649-662.
- Hogarth, R., and Einhorn, E. (1992). Order effects in belief updating: The belief adjustment model. Cognitive Psychology, January, 55-69.
- Hollander, E. P. (1978). Leadership dynamics: A practical guide to effective relationships. New York: Free Press/Macmillan.
- Howell, W. C., Gettys, C. F., Martin, D. W., Nawrocki, L. H., and Johnston, W. A. (1970). Evaluation of diagnostic tests by individuals and small groups. Organizational Behavior and Human Performance, 5, 211-237.
- Hunt, H. K. (1973). Effects of corrective advertising. Journal of Advertising Research, 13 (5), 15-22.
- Iselin, E. R. (1991). Individual versus group decision-making performance: A further investigation of two theories in a bankruptcy prediction task. Journal of Business Finance and Accounting, 18 (2), 191-208.
- Jamieson, B. D. (1968). The "risky-shift" phenomenon with a heterogeneous sample. Psychological Reports, 23, 203-206.
- Janis, I. L. (1972). Victims of Groupthink. Boston: Houghton Mifflin Co.
- Jiambalvo, J., and Waller, W. (1984). Decomposition and assessments of audit risk. Auditing: A Journal of Practice and Theory, 3, 80-88.
- Johnson, J. R., Leitch, R. A., Néter, J. (1981). Characteristics of errors in accounts receivable and inventory audits. The Accounting Review, 56 (2), 270-293.
- Johnson, E. N. (1994). Auditor memory for audit evidence: Effects of group assistance, time delay, and memory task. Auditing: A Journal of Practice and Theory, 13 (1), 36-56.
- Johnson, P. E., Grazioli, S., and Jamal, K. (1993). Fraud detection: Intentionality and deception in cognition. Accounting Organizations and Society, 18 (5), 467-499.
- Joyce, E. J. (1976). Expert judgment in audit program planning. Journal of Accounting Research, 14 (Suppl.), 29-60.
- Kahneman, D., and Tversky, A. (1982). The simulation heuristic. In Kahneman, D., Slovic, P., and Tversky, A. (Eds.) Judgment Under Uncertainty: Heuristic and Biases. Cambridge, UK: Cambridge University Press, 201-210.
- Kahneman, D., and Tversky, A. (1979). Prospect theory: An analysis of decision under risk. Econometrica, 47, 263-291.
- Kahneman, D., Slovic, P., and Tversky, A. (Eds.). (1982). Judgment under Uncertainty: Heuristics and Biases. Cambridge, UK: Cambridge University Press.
- Kamins, M., and Assael, H. (1987). Twosided versus one-sided appeals: A cognitive perspective on argumentation, source derogation, and the effect of disconfirming trial on belief change. Journal of Marketing Research, 24, 29-39.
- Kennedy, J. (1995). Debiasing the curse of knowledge in audit judgment. The Accounting Review, 70 (2), 249-274.
- Kennedy, J. (1993). Debiasing audit judgment with accountability: A framework and experimental results. Journal of Accounting Research, 31 (2), 231-245.
- Kogan, N. and Wallach, M. A. (1967). Group risk taking as a function of members' anxiety and defensiveness levels. Journal of Personality, 35, 50-63.
- Koonce, L. (1992). Explanation and counterexplanation during audit analytical review. The Accounting Review, 67 (1), 59-76.
- Koriat, A., Lichtenstein, A., and Fischhoff, P. (1980). Reasons for confidence. Journal of Experimental Psychology: Human Learning and Memory, 6, 107-118.
- Kozielecki, J. (1981). Psychological decision theory. Dordrecht, Holland: Riedel.

- Lamm, H., and Myers, D. G. (1978). Group-induced polarization of attitudes and behavior. In L. Berkowitz (Ed): Advances in Experimental Social Psychology, 145-195. New York: Academic Press.
- Laughlin, P. R., and Branch, L. C. (1972). Individual versus tetradic performance on a complementary task as a function of initial level. Organizational Behavior and Human Performance, 8, 201-276.
- Levin, I. P., Johnson, R. D., and Chapman, D. P. (1991). Individual differences in dealing with incomplete information: Judging clinical competence. Bulletin of the Psychonomic Society, 29, 451-454.
- Libby, R. (1995). The role of knowledge and memory in audit judgment. In Ashton, R. H., and Ashton, A. H, (Eds), Judgment and Decision-Making Research in Accounting and Auditing, 176-206. New York: Cambridge University Press.
- Libby, R (1981). Accounting and Human Information Processing: Theory and Applications. Englewood Cliffs, NJ: Prentice-Hall.
- Libby, R, and Libby, P. A. (1989). Expert measurement and mechanical combination in control reliance decisions. The Accounting Review, 64, 729-747.
- Libby, R, and Luft, J. (1993). Determinants of judgment performance in accounting settings: Ability, knowledge, motivation, and environment. Accounting, Organizations and Society, 18, 425-450.
- Libby, R, and Trotman, K. T. (1993). The review process as a control for differential recall of evidence in auditor judgments. Accounting, Organizations and Society, August, 249-273.
- Libby, R, Trotman, K. T., and Zimmer, I. (1987). Member variation, recognition of expertise, and group performance. Journal of Applied Psychology, 72 (1), 81-87.
- Lichtenstein, S., Fischhoff, B., Phillips, L. D. (1982). Calibration of probabilities: The state of the art to 1980. In Kahneman, D., Slovic, P., and Tversky, A. (Eds), Judgment Under Uncertainty: Heuristics and Biases. New York: Cambridge University Press, 306-334.
- Lord, C. G., Ross, L., and Lepper, M. R. (1979). Biased assimilation and attitude polarization: The effects of prior theories on subsequently considered evidence. Journal of Personality and Social Psychology, 37, 2098-2109.
- McGarth, J. E., and Altman, I. (1966). Small group research: A synthesis and critique of the field. New York: Holt, Rinehart and Winston.
- McGuire, W. (1961). Resistance to persuasion conferred by active and passive prior refutation of the same and alternative counterarguments. Journal of Abnormal and Social Psychology, 63 (2), 326-332.
- McGuire, W., and Papageorgis, D. (1962). Effectiveness of forewarning in developing resistance to persuasion. Public Opinion Quarterly, 26, 24-34.
- McGuire, W., and Papageorgis, D. (1961). The relative efficacy of various types of prior belief-defense in producing immunity against persuasion. Journal of Abnormal and Social Psychology, 62 (2), 326-332.
- McGuire, T. W., Kiesler, S., and Siegel, J. (1987). Group and computer-mediated discussion effects in risk decision making. Journal of Personality and Social Psychology, 52 (5), 917-930.
- McMillan, J. J., and White, R. A. (1993). Auditors' belief revisions and evidence search: The effect of hypothesis frame, confirmation bias, and professional skepticism. The Accounting Review, 68 (3), 443-465.

- Messier, W. F. Jr. (1995). Research and development in audit decision aids. In Ashton, A. and Ashton, R. (Eds), Judgment and Decision Making Research in Accounting and Auditing. Cambridge, MA: Cambridge University Press.
- Miner, F. C. (1984). Group versus individual decision making: An investigation of performance measure, decision strategies and process losses/gains. Organizational Behavior and Human Performance, 112-124.
- Mladenovic, R. and Simnett, R. 1994. Examination of contextual effects and changes in task predictability on auditor calibration. Behavioral Research in Accounting, 6, 178-203.
- Moscovici, S., and Zavalloni, M. (1969). The group as a polariser of attitudes. Journal of Personality and Social Psychology, 12, 125-135.
- Myers, D. G. (1975). Discussion-induced attitude polarization. Human Relations, 28, 699-714.
- Myers, D. G., and Bishop, G. D. (1971). The enhancement of dominant attitudes in group discussion. Journal of Personality and Social Psychology, 20, 386-391.
- Myers, D. G., and Lamm, H. (1976). The group polarization phenomenon. Psychological Bulletin, 602-627.
- Nisbett, R., and Ross, L. (1980). Human Inference: Strategies and shortcomings of social judgment. Englewood Cliffs, NJ: Prentice-Hall.
- Nisbett, R., and Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. Psychological Review, 84, 231-259.
- Oskamp, S. (1962). The relationship of clinical experience and training methods to several criteria of clinical prediction. Psychological Monographs, 76 (28), 547.
- Oskamp, S. 1965. Over-confidence in case study judgments. Journal of Consulting Psychology, 29, 261-265.
- Paese, P. W., Bieser, M., and Tubbs, M. E. (1993). Framing effects and choice shifts in group decision making. Organizational Behavior and Human Decision Processes, 56 (1), 149-165.
- Pincus, K. V. (1991). Audit judgment confidence. Behavioral Research in Accounting, 3, 39-64.
- Pincus, K. V. (1990). Auditor individual differences and fairness of presentation judgments. Auditing: A Journal of Practice and Theory, 9 (Fall), 150-166.
- Pratt, J., and Jiambalvo, J. (1981). Relationships between leader behaviors and audit team performance. Accounting, Organizations and Society, 6 (2), 133-142.
- Pruitt, D. (1971). Choice shifts in group discussion: An introductory review. Journal of Personality and Social Psychology, 20 (3), 339-360.
- Reckers, P. M., and Schultz, J. J. (1993). The effects of fraud signals, evidence order, and group-assisted counsel on independent auditor judgment. Behavioral Research in Accounting, 5, 124-144.
- Reckers, P. M., and Schultz, J. J. (1982). Individual versus group assisted evaluations. Auditing: A Journal of Practice and Theory, 1, 64-74.
- Regula, C. R., and Julian, J. W. (1973). The impact of quality and frequency of task contributions on perceived ability. Journal of Social Psychology, 89, 115-122.
- Reilly, B. A., and Doherty, M. E. (1992). The assessment of self-insight in judgment policies. Organizational Behavior and Human Decision Processes, 53, 285-309.
- Reilly, B. A., and Doherty, M. E. (1989). A note on the assessment of self-insight in judgment research. Organizational Behavior and Human Decision Processes, 44, 123-131.

- Roediger, H. L. (1978). Recall as a self-limiting process. Memory & Cognition, 6, 54-63.
- Rogelberg, S. G., Barnes-Farrell, J. L., and Lowe, C. A. (1992). The stepladder technique: An alternative group structure facilitating effective group decision making. Journal of Applied Psychology, 77 (5), 730-737.
- Ross, L., Lepper, M. R., and Hubbard, M. (1975). Perseverance in self-perception and social perception: Biased attributional process in the debriefing paradigm. Journal of Personality and Social Psychology, 32 (5), 880-892.
- Ross, L., Lepper, M. R., Strack, F., and Steinmetz, J. (1977). Social explanation and social expectation: Effects of real and hypothetical explanations on subjective likelihood. Journal of Personality and Social Psychology, 35 (11), 817-829.
- Rothbart, M., Evans, M., Fulero, S. (1979). Recall for confirming events: Memory processes and the maintenance of social stereotypes. Journal of Experimental Social Psychology, 15, 343-355.
- Scheonfeld, A. H., and Hermann, D. J. (1982). Problem perception and knowledge structure in expert and novice mathematical problem solvers. Journal of Experimental Psychology: Learning, Memory, and Cognition, 8, 484-494.
- Schultz, J. J., and Reckers, P. J. (1981). The impact of group processing on selected audit disclosure decisions. Journal of Accounting Research, 19 (2), 482-501.
- Sherman, S., Skov, R. B., Hervitz, E.F., and Stock, C. B. (1981). The effects of explaining hypothetical future events: From possibility to probability to actuality and beyond. Journal of Experimental Social Psychology, 17, 142-158.
- Sherman, S., Zehner, K., Johnson, J., and Hirt, E. (1983). Social explanation: The role of timing, set and recall on subjective likelihood estimates. Journal of Personality and Social Psychology, 44, 1127-1143.
- Siegel, S., and Zajonc, R. B. (1967). Group risk-taking in professional decisions. Sociometry, 30, 339-350.
- Smith, J., and Kida, T. (1991). Heuristics and biases: Expertise and task realism in auditing. Psychological Bulletin, 109 (3): 472-489.
- Sniezek, J. A. (1992). Groups under uncertainty: An examination of confidence in group decision making. Organizational Behavior and Human Decision Processes, 52 (1), 124-155.
- Sniezek, J. A., and Henry, R. A. (1990). Revision, weighting, and commitment in consensus group judgment. Organizational Behavior and Human Decision Processes, 45, 66-84.
- Sniezek, J. A., and Henry, R. A. (1989). Accuracy and confidence in group judgment. Organizational Behavior and Human Decision Processes, 42, 1-28.
- Snyder, M., and Campbell, B. (1980). Testing hypotheses about other people: The role of hypothesis. Personality and Social Psychology Bulletin, 421-426.
- Snyder, M., and Cantor, N. (1979). Testing hypotheses about other people: The use of historical knowledge. Journal of Experimental Social Psychology, 15, 330-342.
- Snyder, M., Swann, W. B. Jr. (1978). Hypothesis-testing processes in social interaction. Journal of Personality and Social Psychology, 36, 1202-1212.
- Snyder, M., and Uranowitz, S. W. (1978). Reconstructing the past: Some cognitive consequences of person perception. Journal of Personality and Social Psychology, 36, 941-950.
- Slovic, P. and Lichtenstein, S. (1971). Comparison of Bayesian and regression approaches to the study of information processing in judgment. Organizational Behavior and Human Performance, 6, 647-744.

- Solomon, I. (1987). Multi-auditor judgment/decision making research. Journal of Accounting Literature, 6, 1-25.
- Solomon, I. (1982). Probability assessment by individuals auditors and audit teams: An empirical investigation. Journal of Accounting Research, Autumn, 689-710.
- Solomon, I., Krogstad, J. L., Romney, M. B., and Tomassini, L. A. (1982). Auditors' prior probability distributions for account balances. Accounting, Organizations and Society, 7 (1), 27-41.
- Solomon, I., and Shields, M. D. (1995). Judgment and decision-making research in auditing. In R. H. Ashton and A. H. Ashton (Eds), Judgment and decision-making research in accounting and auditing. Cambridge, MA: Cambridge University Press: 137-175.
- Stasser, G. (1992). Information salience and the discover of hidden profiles by decision-making groups: A "thought experiment". Organizational Behavior and Human Decision Processes, 52, 156-181.
- Stasser, G. (1988). Computer simulation as a research tool: The DISCUSS model of group decision-making. Journal of Experimental Social Psychology, 24, 393-422.
- Stasser, G., and Titus, W. (1985). Pooling of unshared information in group decision-making: Biased information sampling during group discussion. Journal of Personality and Social Psychology, 48, 1476-1478.
- Stein, R. T. (1975). Identifying emergent leaders from verbal and nonverbal communications. Journal of Personality and Social Psychology, 32 (1), 125-135.
- Steiner, I. D. (1966). Models for inferring relationships between group size and potential group productivity. Behavioral Science, 11, 273-283.
- Stocks, M. H., and Harrell, A. (1995). The impact of an increase in accounting information level on the judgment quality of individual and groups. Accounting, Organizations and Society, 20 (7/8), 685-700.
- Stone, D. N., and Dilla, W. N. (1994). When numbers are better than words: The joint effects of response representation and experience on inherent risk judgments. Auditing: A Journal of Practice and Theory, 13 (Suppl.), 1-18.
- Sullivan, K. (1997). Corporate managers' risky behavior: Risk taking or avoiding? Journal of Financial and Strategic Decision Making, in press.
- Surber, C. F. (1985). Measuring the importance of information in judgment: Individual differences in weighting ability and effort. Organizational Behavior and Human Decision Processes, 35, 156-178.
- Tomassini, L. A., Solomon, I., Romney, M. B., and Krogstad, J. L. 1982. Calibration of auditors' probabilistic judgments: Some empirical evidence. Organizational Behavior and Human Performance, 30 (3): 391-406.
- Trotman, K. T. (1985). A review process and the accuracy of auditor judgments. Journal of Accounting Research, 23 (2), 740-752.
- Trotman, K. T., and Sng, J. (1989). The effect of hypothesis framing, prior expectations and cue diagnosticity on auditors' information choice. Accounting, Organizations and Society, December, 565-576.
- Trotman, K. T., and R. Wood, (1991). A meta-analysis of studies on internal control judgments. Journal of Accounting Research, 29 (1), 180-192.
- Trotman, K. T., and Yetton, P. W. (1985). The effect of the review process on auditor judgments. Journal of Accounting Research, Spring, 256-267.
- Trotman, K. T., Yetton, P. W., and Zimmer, I. R. (1983). Individual and group judgments of internal control systems. Journal of Accounting Research, Spring, 286-292.

- Tuckman, J., and Lorge, I. (1962). Individual ability as a determinant of group superiority. Human Relations, 15, 45-51.
- Tutin, J. (1993). The persistence of initial beliefs in clinical judgment. Journal of Social and Clinical Psychology, 12 (3), 319-335.
- Tversky, A., and Kahneman, D. (1982). Judgment under uncertainty: Heuristics and biases. In Kahneman, D., Slovic, P., and Tversky, A. (Eds), Judgment Under Uncertainty: Heuristics and Biases. Cambridge, UK: Cambridge University Press, 3-22.
- Tversky, A., and Kahneman, D. (1981). The framing of decisions and the psychology of choice. Science, 211, 453-458.
- Tversky, A., and Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. In Kahneman, D., Slovic, P., and Tversky, A. (Eds), Judgment Under Uncertainty: Heuristics and Biases. Cambridge, UK: Cambridge University Press, 3-22.
- Tversky, A., and Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. Cognitive Psychology, 4, 207-232.
- Uecker, W. C. (1982). The quality of group performance in simplified information evaluation. Journal of Accounting Research, 20 (2), 388-402.
- Wallach, M., and Kogan, N. (1965). The roles of information, discussion, and consensus in group risk taking. Journal of Experimental Social Psychology, 1, 1-19.
- Wallach, M., Kogan, N. and Bem, D. (1962). Group influence on individual risk taking. Journal of Abnormal and Social Psychology, 65, 75-86.
- Waller, W. S., and Felix, W. L. Jr. (1984). The auditor and learning from experience: Some conjectures. Accounting, Organizations and Society, 8 (3/4), 383-406.
- Watson, D. J. (1975). The structure of project teams facing differentiated environments: An exploratory study in public accounting firms. The Accounting Review, April, 259-273.
- Weisband, S. P. (1992). Group discussion and first advocacy effects in computer-mediated and face-to-face decision making groups. Organizational Behavior and Human Decision Processes, 53 (3), 352-380.
- Whyte, G. (1993). Escalating commitment in individual and group decision making: A prospect theory approach. Organizational Behavior and Human Decision Processes, 54 (3), 430-455.
- Wright, W. F. (1988). Audit judgment consensus and experience. In Behavioral Accounting Research: A Critical Analysis, edited by Ferris, K. R., 305-328.
- Yates, J. F. (1990). Judgment and Decision Making. Englewood Cliffs: Prentice-Hall.
- Yetton, P. W., and Bottger, P. (1982). Individual versus group problem solving: An experimental test of a best-member strategy. Organizational Behavior and Human Performance, June, 307-321.
- Zadney, J., and Gerard, H. (1974). Attributed intentions and informational selectivity. Journal of Experimental Social Psychology, 10, 34-52.

APPENDICES

Appendix A

Auditors' Evaluation of Internal Controls in the Inventory Cycle

This questionnaire seeks your cooperation in an internal control evaluation task. All information provided will be treated with strict confidence.

Your cooperation in completing this questionnaire is much appreciated.

**Ms Janne Chung
Prof. Gary S. Monroe
School of Accounting
Faculty of Business
Edith Cowan University
Pearson St
Churchlands 6018
Tel: (64) 9 273 8733
Fax: (64) 9 273 8121**

Dear Participant

Thank you for agreeing to participate in our research study which should take about 10 minutes. Your participation will contribute to our understanding of how auditors formulate judgments when evaluating internal controls. Before proceeding, please read the following instructions carefully:

1. We are interested in a number of issues concerning the evaluation of internal controls. As part of our study, we would like you to evaluate the internal control system in the inventory area.
2. It is important you complete **all** the tasks in this questionnaire.
3. It should be emphasised that there are no right and wrong answers. All that you are asked is to provide your best judgments. The case is straightforward and involves no tricks.
4. If you have any questions when answering this questionnaire, please raise your hand and the instructor will attend to you. Please do not ask questions aloud. Do not consult your colleagues because each person has a different questionnaire.

It is important that you follow the instructions.

Yours faithfully,

Janne Chung
Gary S. Monroe
School of Accounting
Edith Cowan University

Task I begins on the next page.

Task

Part I. An auditor's task is to evaluate the **risk** of the internal control system in the inventory area.

Please write down in the spaces provided below, the situations you think may result in a **high** or **low** risk internal control system. Write down as many situations as you can think of. Phrases and short sentences would be sufficient.

Situations resulting in a **high risk** internal control system in the inventory area:

Situations resulting in a **low risk** internal control system in the inventory area:

Please do not turn over the page until you have completed Task I. Task II begins on the next page.

Task II

Western Australia Sporting Goods Ltd

You are in charge of the audit of Western Australia Sporting Goods Ltd (WASG) for the year ending June 30, 19A2. Your firm has been the auditor for the past 12 years and has issued unqualified audit opinions for the last five years. All adjusting entries found as a result of audits in the past have been adjusted by the client. Currently you are engaged in the final evaluation of the client's internal control system over inventory in order to finalise the audit plan for that area.

Please read through the information below describing the internal control system. **At the end of the evaluation of the internal control system, you will be asked to make a decision on the internal controls for that area.** (To keep the case brief, not all the information you might like to have for your evaluation has been included. To the extent that the information is not presented, you should assume that the situation is "normal" in that this audit does not involve any unusual or significant undisclosed circumstances.)

General History and Background

WASG is a distributor of sporting goods of all descriptions, from fishing tackle, to hockey gear, to exercise and home gym equipment. The company was founded in Perth 13 years ago. The company stocks a wide range of products and has a large customer base of wholesale and retail stores. The company purchases its products directly from various manufacturers. The products are distributed to wholesalers and retailers in Perth and Western Australia.

The company has experienced reasonable growth in the past few years. However, WASG has been affected by the current economic downturn. Many of their customers are experiencing reduced sales and profits and the industry as a whole is going through some tough times, with numerous business failures. The financial condition of WASG has declined in the first three quarters of 19A2 - sales, profits and earnings per share have not met management's expectations. Nor is this situation expected to improve throughout the remainder of 19A2. Management expects total revenues and net income to continue to decline for the remainder of 19A2.

continue next page

The Inventory System in Brief

In order to assess the client's internal control system, you have collected the following information by observation, interviewing appropriate personnel, performing preliminary tests, and other relevant procedures. (For identified controls you may assume that they operated effectively throughout the year. If a control procedure is not mentioned you are to assume that it does not exist.)

WASG has recently computerised its inventory system and now maintains perpetual inventory records. Prior to the current year the company relied on a periodic inventory system. WASG has several smaller satellite warehouses. Although all receipts of merchandise and shipments to customers are made through the main warehouse, the satellite warehouses sometimes act as a buffer to absorb extra inventory when increased demand for a particular item is anticipated. Each satellite warehouse has its own procedures based on those in place at the main WASG warehouse.

A brief description of each area involved with receipt, warehousing and shipment of merchandise is as follows:

1. **Receiving dock** - Goods are received directly from the suppliers' factories in bulk. Other items such as office supplies, repair parts and small fixed assets are also received at the receiving dock.
2. **Warehouse** - Inventory is held here for repackaging and shipment on customer order.
3. **Repackaging area** - On customer order, personnel in this area package the shipment according to customer specifications. Sometimes items can be simply re-labelled and shipped as received from the factory, other orders may require more complex packaging such as changing containers in addition to re-labelling.
4. **Shipping dock** - All materials leaving the plant facility cross this dock, and a shipping report is prepared. Inventory shipments comprise nearly all of total shipments.

continue next page

Additional Information

- Receipts of inventory are counted, inspected for damages, completeness, etc. on a test basis and a receiving report (not prenumbered) is prepared for all materials crossing the receiving dock.
- The perpetual inventory records (in units and in \$) are maintained by a person independent of the inventory stores custodian.
- The physical premises of the warehouse are in a wing of the building separate from the other functions such as administration, but there are no effective access restrictions to this wing.
- After inspection, the received items are stored in the warehouse according to general type - i.e., fishing gear, football-related, hockey-related, etc.
- A physical inventory count is done only in conjunction with the preparation of the financial report at year-end. The perpetual inventory records are not reconciled to the physical inventory on hand throughout the year.
- There is no regular review of slow-moving or potentially obsolete merchandise.
- Orders are prepared for shipment to customers from a copy of the original sales order/invoice (WASG uses multicopy sales invoices) which becomes the bill of lading (shipping docket).

Please turn over the page for the questions relating to Task II.

Task III

Your background: Please tick (✓) the appropriate spaces provided below or circle an appropriate number where applicable.

1. Your gender?
 Male Female

2. Your age? _____ yrs

3. Is English your first or native language?
 Yes No

4. What is the highest level of formal education you have completed?
 Some TAFE units
 Completed TAFE course
 Some university units
 Graduated from university

5. How many years of working experience do you have?
_____ yrs and _____ months

6. How many years of audit working experience do you have
_____ yrs and _____ mths

7. How difficult did you find the tasks given to you previously in this questionnaire?
1 2 3 4 5 6 7 8 9
Not at all Somewhat Extremely
difficult difficult difficult

8. How much effort do you think an actual internal control evaluation task will require?
1 2 3 4 5 6 7 8 9
Very little Some effort A great deal
effort of effort

9. How much effort did you expend on the tasks in this questionnaire?
1 2 3 4 5 6 7 8 9
Very little Some effort A great deal
effort of effort

Thank you for your cooperation in this study.
INCRaI

Appendix B

December 1996

Dear Professional Colleague

Thank you for agreeing to participate in this research project which should take about 20 minutes to complete. We are interested in a number of issues concerning a company's ability to continue as a going-concern. Your participation will help us develop ways for improving going-concern audits.

Below is a list of instructions you should read before proceeding:

1. You will be required to audit the going-concern status of a company.
2. The task you are required to perform will be divided into two parts. In the first part of the experiment you will be given the background information and the audit findings of your audit team. You should familiarise yourself with this information. In the second part of the experiment you will be asked some questions about yourself. Please complete both parts of the experiment.
3. If you have any questions, please raise your hand and you will be attended to individually.
4. Please complete the attached receipt. Strict confidentiality will be maintained, but details are required to assist in paying you. Your identity will be kept confidential at all times.
5. It is important that you adhere strictly to the instructions.

Thank you for your cooperation in this research project.

Yours sincerely

Janne Chung (Lecturer)
Gary S. Monroe (Professor)
School of Accounting
Edith Cowan University

Part I

You have been assigned to an audit group for the purpose of auditing the going-concern status of a company. Since this study is concerned with group judgments, it is essential that you discuss and work on the case as a group.

Assume you are an audit senior assigned to the audit of the following company: ABC Pty Ltd which is incorporated in 1967, operates mainly in the Melbourne metropolitan area. The main activity of the company is the manufacture and sale of roller shutters. Even though it is an exempt proprietary company, ABC has elected to be audited because management feels that the company would benefit from an independent audit. The following are eight information items that your audit team has noted during the 1996 financial report audit:

1. Legal proceedings have been started against the company that may result in financial judgment that could not be met.
2. The company may have idle assets which it can sell for cash.
3. The fire department may make roller shutter doors and windows illegal in residential properties because of the risk they pose in the event of a fire.
4. Management may take steps to reduce overhead and administrative expenses
5. A fire destroyed the warehouse that may be uninsured.
6. Associates of the company may take up a new issue of the company's shares.
7. Some of the company's fixed-term borrowings are approaching maturity and the company may not have realistic prospects of renewal or repayment.
8. Current shareholders may agree to contribute additional capital.

Please turn over the page for questions relating to ABC Pty Ltd.

Please ensure that Part I (i.e., pages 1 to 3) is inserted into the envelope provided before proceeding to the next task.

Part II

Based on your group's evaluation of the information presented in the main experimental part of this research project, please list below the information items that your group considered to be relevant to the audit of ABC Pty Ltd's going-concern status:

Please turn over the page when you have completed this task.

On the previous page, you have just listed those information items that were relevant to your group's audit of ABC Pty Ltd's going-concern status. Please rate how relevant each of these items was to your group's assessment of the going-concern status of ABC Pty Ltd. Please turn back to the previous page, and using the scale shown below, place the rating of each item in the left hand margin next to the item being rated.

1	2	3	4	5	6	7
Slightly relevant			Moderately relevant			Highly relevant

Please turn over the page when you are ready.

This section asks you questions about yourself and should be completed individually. Please do not discuss these questions with other members of your group. Please tick (✓) the appropriate spaces provided below or circle an appropriate number.

1. Your gender? () Male () Female
2. Your age? () yrs
3. Is English your first or native language? () Yes () No
4. I have _____ years and _____ months working experience of which _____ years and _____ months are in auditing.
5. What is your position in your firm?
() Audit Junior () Audit Senior () Supervisor
6. How difficult did you find the tasks given to you previously in the main experimental part of this research project?

1	2	3	4	5	6	7	8	9
Not at all difficult				Somewhat difficult				Extremely difficult
7. How much effort do you think an actual audit of a company's going-concern status will require?

1	2	3	4	5	6	7	8	9
Very little effort				Some effort				A great deal of effort
8. How much effort did you expend on the main experimental part of this research project?

1	2	3	4	5	6	7	8	9
Very little effort				Some effort				A great deal of effort
9. I am satisfied with my group's judgments.

Strongly disagree	1	2	3	4	5	6	7	Strongly agree
----------------------	---	---	---	---	---	---	---	-------------------
10. How would you rate the amount of interaction in your group's audit process?

1	2	3	4	5	6	7	8	9
No interaction at all				Moderate amount of interaction				A lot of interaction

Please insert all the experimental materials into the envelope provided. Thank you for your cooperation in this study.

Appendix C

December 1996

Dear Professional Colleague

Thank you for agreeing to participate in this research project which should take about 30 minutes to complete. The purpose of this research is to study judgments auditors make during the audit of the inventory balance. Your participation will help us develop ways for improving audit judgments.

Below is a list of instructions you should read before proceeding:

1. You will be required to perform the audit of the inventory balance of a real but disguised company in the experiment.
2. The task you are required to perform will be divided into two parts. Please complete both parts of the experiment. In the first part of the experiment you will be given the background information and the working papers for the audit of the inventory balance. You should familiarise yourself with this information.
3. If you have any questions, please raise your hand and you will be attended to individually.
4. Please complete the attached receipt and present it to the instructor for payment at the end of the experiment. Strict confidentiality will be maintained, but details are required to assist in paying you. Your identity will be kept confidential at all times.
5. It is important that you adhere strictly to the instructions.

Thank you for your cooperation in this research project.

Yours sincerely

Janne Chung (Lecturer)
Gary Monroe (Professor)
School of Accounting
Edith Cowan University

Part I

Instructions

1. In this part of the experiment, you have been assigned to groups. Since this study is concerned with group judgments, it is essential that you discuss and work on the case together.

Before you commence discussion as a group, you should familiarise yourself with the case materials. It should take you about 15 minutes.

2. The case materials, which begin on the next page, contain general background information about the client as well as a list of audit evidence.
3. If you have any questions during the experiment, please raise your hand and the instructor will attend to you. Please do not ask questions aloud during the experiment.

Please turn over the page when you are ready.

The Client: Tucker House Ltd

Tucker House Ltd operates a chain of fast food restaurants. The company began in 1991 and by the end of 1996 had grown to 41 units, located primarily in the eastern states. Tucker House uses a financial accounting year end of 31 October.

Restaurants

The atmosphere of the restaurants is similar to most fast food restaurants, attractive, bright and clean. All of the restaurants open 7 days a week for lunch and dinner. The busier restaurants open for breakfast as well. The menu is more up-market than most hamburger outlets serving steak sandwiches, beef burgers, selected chicken items, and the usual drinks and desserts.

Operational control has been maintained through a computerised information system. Each restaurant is equipped with specially designed computerised cash registers that have the menu items incorporated on them. This system is used to compute on a daily, weekly and monthly basis, separately for each of the restaurant units, profit and loss, sales and cost breakdown by product, labour productivity, payroll and variations from budget.

The Purchasing and Warehousing Division

The company operates its own purchasing and warehousing division (PWD) from which all the restaurants are supplied with portion-controlled units of meat, poultry and other food items, as well as paper and other supplies.

The PWD includes meat cutting rooms, refrigeration and freezer storage space. The PWD processes much of the meat required by the restaurants and makes all the sauces and dressings. The company purchases other items in large quantities in anticipation of seasonal price fluctuations.

The effort to control meat cost is the most important area of cost control at Tucker House. Meat represents 80% of food costs, or almost 33% of sales, and management feels that one of the keys to Tucker House's success is the PWD which makes it possible for Tucker House to cope with changes in the prices of meat.

The Account: Inventory at financial year end 31.10.96

<u>Type of Inventory</u>	<u>\$</u>	<u>Percentage of Total</u>
Meat	2,615,546	75.0%
Other (non-meat) food/beverages	<u>335,620</u>	<u>10.0%</u>
Total food/beverages	2,951,166	85.0%
Supplies	266,942	7.5%
Others	<u>269,286</u>	<u>7.5%</u>
Total: All inventory	<u>3,487,394</u>	<u>100.0%</u>

Instructions

In order for you to perform an audit of Tucker House's inventory balance, the following information items are available. You should only evaluate those information items you consider are necessary for forming an opinion on whether or not the inventory balance is fairly stated.

ADDITIONAL INFORMATION AVAILABLE

Information about Tucker House Directors, Management and Employees..... pg 4

1. Key management personnel
2. Members of the Board of Directors
3. Board structure and number of meetings
4. Board members comments regarding managing director and financial director

Other Information about Tucker House pg 6

5. Menu prices for main meals, 1994-96; average bill per customer
6. Main meal sales by type (percentages)
7. Number of restaurant units in operation, 1994-96
8. Expansion costs and plans
9. History of share offerings and debt financing
10. Management forecast of 1996 sales/earnings

Inventory and Inventory-related Financial Report Data/Ratios..... pg 8

11. Weighted average annual sales per restaurant unit, 1994-96
12. Cost of Sales as a percentage of Sales, 1994-96
13. Cost of Sales percentage breakdown (food/beverages vs. labour), 1994-96
14. Total purchases of food and beverages, 1994-96
15. Total inventories (\$), 1994-96
16. Inventory turnover and number of days sales in ending inventory, 1994-96
17. Inventory location (Restaurants vs PWD), 1995-96
18. Percentage breakdown of inventory by type, 1994-96
19. Percentage breakdown of meat inventory by type, 1996
20. Purchase commitments (\$) for meat at financial year end, 1994-96

Other Financial Report Information pg 11

21. Current assets, by type, as a percentage of total assets, 1994-96
22. Current liabilities, by type, as a percentage of total liabilities and shareholders equity, 1994-96
23. Gross sales, 1994-96
24. Net income as a percentage of Sales, 1994-96

Financial Ratios (Other than inventory-related)..... pg 12

25. Current ratio and Quick ratio, 1994-96
26. Receivables turnover & Number of days sales in ending accounts receivables, 1994-96
27. Earnings per share, 1994-96
28. Price-Earnings ratio, 1994-96

Results of 1996 Audit Procedurespg 13

- 29. Results of physical inventory observation at restaurants
- 30. How restaurant test units for inventory observation were chosen
- 31. Results of analytical comparisons/reviews for unobserved restaurants
- 32. Results of physical inventory observation at PWD
- 33. Results of inventory pricing tests
- 34. Results of inventory cutoff tests
- 35. Audit procedures performed related to purchase commitments
- 36. 1996 Management Letter comments re internal control weaknesses
- 37. Evaluation of effectiveness of Internal Audit function
- 38. 1996 Solicitor's Letter

Inventory Policies/Background and Selected Client Recordspg 18

- 39. PWD history (size, expansions) - 1992-96
- 40. Beef processing and purchasing policies
- 41. Accounting records/internal controls for PWD inventory
- 42. Sources (vendors) for food, beverage and supply purchases
- 43. Client records: Receiving Log, Week 52, FY 96
- 44. Client records: Receiving Log, Week 1, FY 97
- 45. Client records: Weekly Inventory Summary, Week 51, FY 96
- 46. Client records: Weekly Inventory Summary, Week 52, FY 96
- 47. Client records: Physical Inventory Counts, 3 largest items, FY 95
- 48. Client records: Weekly Inventory Summary, Week 1, FY 97

**INFORMATION ABOUT TUCKER HOUSE DIRECTORS, MANAGEMENT
AND EMPLOYEES**

1. Key Management Personnel

Jeffrey Hunerg - Founder, Managing Director, Chairman of the Board. Mr. Hunerg, an solicitor, directly owns 13.6% of the outstanding shares of Tucker House plus another 4.3% held in trust for his minor children and may be deemed to have controlling interest in the company.

Robert Dike - Deputy Managing Director, Financial Director since 1992, Treasurer, Member of the Board. Mr Dike, a CA, worked as an auditor for a Melbourne public accounting firm for 4 years prior to joining Tucker House in 1991.

Albert Nash - accountant since 1995, responsible for preparation of financial reports, reports to Dike.

Donald Chioni - Internal Auditor since 1994, previously Asst. Manager and then Manager of a restaurant unit for 2 years.

David White - Manager of Purchasing & Production since 1994, formerly Manager of PWD since 1992.

Mel Hellman - Manager of PWD since 1994, formerly Manager of the construction subsidiary for one year.

No member of management other than Hunerg owns over 2% of the outstanding Tucker House shares.

2. Members of the Board of Directors

<u>Name</u>	<u>% of O/S Shares Held</u>
Jeffrey Hunerg	13.6% directly, 4.3% through children
Robert Dike	*
Chip Hellman	*
<u>Non-executive Directors</u>	
R. J. Cramden	4.3%
John Burns	*
Edmund Lieberman	4.6%
Richard Chatte	*
Kevin Jones	*

* These directors as a group own 4.6% of shares.

All directors but Hellman were early investors and have served continuously since the firm began. Hellman is newly appointed.

3. Board Structure and Number of Meetings, 1994-96

Board meetings are called by Hunerg and are generally held when a specific matter requires Board action. Meetings tend to be fairly brief, usually 1-2 hours, and disagreements are rare. The number of meetings held per year is:

<u>1994</u>	<u>1995</u>	<u>1996</u>
6	6	4

There are only 2 subcommittees to the Board. The Executive Committee (Hunerg, Dike and Cramden) forms "the brain trust" for day-to-day operating decisions. The Audit Committee meets with the auditors semi-annually, but other than this does not meet during the year. In addition, committee members receive quarterly earnings information for their review.

4. Board Member Comments Regarding Managing Director and Financial Director

In a meeting of the audit committee with the auditors, the directors commented that Hunerg is "a brilliant promoter who runs Tucker House with an iron fist" and said that they are "impressed by both the results produced by Hunerg and by his apparently modest salary demands".

The Board feel that Dike has been instrumental in improving the company's accounting and financial operations and they appreciate the manner in which he has been able to "get along with Hunerg and accept being second in command". The first company financial director resigned in July 91 over constant disagreements with Hunerg, primarily over Hunerg's early policy of deliberately paying creditors' invoices late and "bickering with suppliers' over invoiced amounts in an effort to maximise Tucker House's cash position. The escalating disagreements between Rossmoor and Hunerg culminated in a demand by Hunerg that Rossmoor either let Hunerg run Tucker House as he wish or leave. Rossmoor stayed only long enough to help break in Dike and has been inactive in the company ever since.

OTHER INFORMATION ABOUT TUCKER HOUSE

5. Menu Prices (Main meals), 1994-96: Average Sale Per Customer

<u>Financial Year</u>	<u>Price range: dinner main meals</u>	<u>Price range: lunch main meals</u>
1994	\$3.45-\$5.95	\$1.95-\$3.95
1995	\$3.50-\$5.95	\$1.95-\$4.65

Average bill per customer:

1996: \$6.88-\$7.00
 1995: \$6.25

6. Main meal Sales by Type (Percentages)

<u>Main meal</u>	<u>Percentage of number of meals sold</u>
Steak sandwich	6%
Supreme steak sandwich	21%
Deluxe burger	15%
Doner kebabs: beef	4%
cheese	6%
chicken	3%
Tacos: beef	18%
chicken	8%
Tucker House Combination Platter	<u>19%</u>
Total	<u>100%</u>

7. Number of Restaurants in Operation, 1994-96

	<u>1994</u>	<u>1995</u>	<u>1996</u>
Number of restaurants open at FYE	28	34	41
Percentage increase over previous year	87%	21%	21%
Weighted average number of units open for year	20.42	32.25	35.42
Percentage increase over previous year	133%	58%	18%

Only 1 unit has ever been closed. It opened in 1994 and closed in 1996 when revenues failed to meet expectations.

8. Expansion Costs and Plans

All units are in leased premises. 15 new locations are planned for FY 97, including the company's first locations outside of the east. Each new location will cost between \$200-350,000 to investigate, develop and equip.

9. History of Share Offerings and Debt Financing

During 1990, Jeffrey Hunerg and R. J. Cramden, two of the three founders of the firm, raised privately about \$200,000 to finance the opening of the initial unit. During 1991, an additional \$300,000 was obtained by the sale of the 100,000 shares of ordinary shares.

In 1992, to finance the opening of the PWD, several small bank loans (\$50-100,000) were obtained during this period with Hunerg and his wife and Dike and his wife acting as guarantors.

To finance further expansion, the company made a public offering of 220,000 ordinary shares at \$12.50/share in June 1993. Expansions since that date have been financed by the proceeds of the public offering and an additional \$4,000,000 borrowed in 1995 and 1996 under a revolving credit agreement at an interest rate of 1% over the lending bank's large business prime.

The shares, after a 3:2 split in March 1996, closed at 8.80 asked, 6.60 bid on the stock exchange at 1/11/96, a trading range which has been fairly stable since the split. The company intends to offer for sale 450,000 shares not later than 120 days after FYE 31/10/96. The proceeds will be used to repay the borrowings, finance openings of additional restaurants and for working capital. The company still expects to use the revolving credit agreement from time to time to finance expansion.

10. Management Forecasts Made of 1996 Sales/Earnings

In April 1996, Hunerg and Dike appeared before the Australian Association of Financial Planners and predicted 20-25% gain in sales and a 40% increase in profit for FY 96 over FY 95.

In an April 10th press release, Tucker House disclosed an EPS prediction of \$1.86 for FY 96.

These were the first publicly announced forecasts the company had ever made.

INVENTORY AND INVENTORY-RELATED FINANCIAL REPORT
DATA/RATIOS

11. Weighted Average Annual Sales Per Restaurant, 1994-1996

<u>Year</u>	<u>Weighted average annual sales per restaurant</u>	
1994	\$813,257	5.9% increase
1995	\$805,299	1.0% decrease
1996	\$856,592	6.3% increase

The decrease from 1994 and 1995 was significantly affected by a May-July 1995 promotion offering \$2 off any main meal (which meant most meals were sold at a loss) in an effort to combat declining customer counts from late 1994 to mid-1995. The promotion was "too successful" as customer counts increased by 15%, far higher management had anticipated. Third quarter 1995 earnings were only \$0.01 versus \$0.32 the prior year.

The range of annual sales/unit in 1996 was \$516,412-\$1,972,048.

12. Cost of Sales as a Percentage of Sales, 1994-96

<u>Year</u>	<u>Cost of Sales as a % of Sales</u>
1994	59.4%
1995	59.9%
1996	59.0%

Management feels that as the size of the company has increased there has been an increase in efficiency in purchasing, production and distribution. Also, over the years there has been an increasing portion of total sales represented by relatively higher profit items, such as alcoholic beverages (which accounted for 14% of 1994 sales, 16% of 1995 sales and 26% of 1996 sales). However, the effects of these two improvements have been generally offset by increased labour costs.

13. Cost of Sales Breakdown (Food/Beverage vs. Labour) 1994-1996

Components of cost of sales:

<u>Year</u>	<u>Food & Beverage</u>	<u>Labour & Payroll Tax</u>
	65.66%	34.34%
1993	65.19%	34.81%
1994	65.96%	34.04%
1995	65.51%	34.49%
1996	61.72%	38.28%

14. Total Purchases of Food/Beverages, 1994-96*

<u>Year</u>	<u>Food/beverages purchases</u>	<u>Increase over prior year</u>
1994	\$ 7,069,325	142.7%
1995	\$11,161,728	57.9%
1996	\$12,024,682	7.7%

15. Total Inventories (\$), 1994-96*

<u>Year</u>	<u>Total Inventory</u>	<u>Increase over Prior Year</u>
1994	\$1,230,636	105%
1995	\$2,352,465	91%
1996	\$3,487,394	48%

Inventory is valued at FIFO cost, which approximates market.

16. Inventory Turnover & Number of Days Sales in Ending Inventory, 1994-96*

<u>Year</u>	<u>Inventory Turnover</u>	<u># Days Sales in Ending Inventory</u>
1994	10.8 times	45.7 days
1995	8.7 times	55.1 days
1996	6.1 times	71.1 days

PWD was expanded in mid-1995.

17. Inventory Location (Restaurant vs. PWD), 1995-96

<u>Location:</u>	<u>FY 1996</u>		<u>FY 1995</u>	
	<u>Dollars</u>	<u>%</u>	<u>Dollars</u>	<u>%</u>
Restaurants	\$ 770,001	22	\$ 480,103	20
PWD	\$2,717,393	78	\$1,872,362	80
Totals	\$3,487,394	100	\$2,352,465	100

The per unit averages at year end were:

<u>Year</u>	<u>Number of Units</u>	<u>Avg./Unit at Rest.</u>	<u>Avg./Unit at PWD</u>
1996	41	\$18,781	\$66,277
1995	34	\$14,121	\$55,069

*1993 is the first full year of operations.

18. Percentage Breakdown of Inventory by Type, 1994-96

	<u>1996</u>	<u>1995</u>	<u>1994</u>
Food and beverages	85%	84%	82%
Supplies	7.5%	7%	9%
Other supplies	<u>7.5%</u>	<u>9%</u>	<u>10%</u>
Totals	100%	100%	100%

Note: 1996 figures are unaudited.

19. Percentage Breakdown of Meat Inventory by Type, 1996

**Percentage of Dollar Value of Ending
Meat Inventory**

Prepared, portion controlled:

Steak sandwich (230-gm. serving)	23%
Supreme steak sandwich (260-gm. serving)	8%
Deluxe burger (beef, 2 per serving)	7%
Doner kebabs (beef, 3/serving)	5%
Tacos (beef, 3/serving)	7%

Bulk:

Ground beef (will be used in fillings for deluxe burger, tacos, and doner kebabs)	21%
Beef in bulk (uncut, unprocessed)	<u>29%</u>
Total	100%

20. Purchase Commitments (\$) for Meat at FYE, 1994-96

<u>Year</u>	<u>Commitment at FYE</u>	<u>Related Deposit</u>
1994	\$2,500,000	none
1995	none	--
1996	\$ 530,000	\$130,000

Purchase commitments are made only when Tucker House purchasing personnel feel that conditions are favourable.

OTHER FINANCIAL REPORT INFORMATION

21. Current Assets, by Type, as % of Total Assets, 1994-96

All percentages are percentage of total assets.

<u>Year</u>	<u>Cash</u>	<u>Trade A/c</u> <u>Rbles</u>	<u>Other A/c</u> <u>Rbles</u>	<u>Inventory</u>	<u>Prepays</u>	<u>Total CA</u>
1994	10.4%	3.0%	--	11.6%	3.7%	28.7%
1995	6.5%	3.7%	--	16.4%	4.5%	31.1%
1996	5.7%	4.6%	--	18.4%	5.6%	34.3%

22. Current Liabilities, By Type, as % of Total Liabilities + Shareholders Equity, 1994-96

All percentages are percent of total liabilities and shareholders equity.

<u>Year</u>	<u>A/c Pbles</u>	<u>Accrued</u> <u>Liab.</u>	<u>Notes Pbles</u> <u>- Current</u>	<u>Taxes</u> <u>payable</u>	<u>Total Current</u> <u>Liab.</u>
1994	9.8%	3.7%	1.4%	3.0%	17.9%
1995	13.3%	4.2%	3.8%	0.9%	22.2%
1996	11.8%	4.2%	2.4%	--	17.6%

23. Gross Sales, 1994-96

<u>Year</u>	<u>Gross Sales</u>	<u>% Increase Over Prior Year</u>
1994	\$16,603,993	147.0% increase
1995	\$25,970,904	56.4% increase
1996	\$30,337,623	16.8% increase

Management analysis: These increases are due almost exclusively to the opening of new restaurants each year. For FY 96, the aggregate sales of restaurants open in any of the three previous financial years were slightly greater than sales of those restaurants in FY 95. Sales per customer has increased in each of the years, primarily due to increased sales of beverages (a high profit margin item) and, to a lesser extent, increased menu prices. Customer counts in FY 96 were approximately 2% ahead of FY 95.

24. Net Income as a Percentage of Sales, 1994-96

<u>Year</u>	<u>Net Income as a % of Sales</u>
1994	5.1%
1995	3.6%
1996	4.5%

Management analysis: The decline in 1995 results primarily from the effects of a disastrous sales promotion. The 1995 and 1996 figures also reflect the increasingly difficult economic times.

FINANCIAL RATIOS (OTHER THAN INVENTORY-RELATED)

25. Current Ratio and Quick Ratio, 1994-96

<u>Year</u>	<u>Current Ratio</u>	<u>Quick Ratio</u>
1994	1.60:1	0.75:1
1995	1.40:1	0.46:1
1996	1.94:1	0.50:1

26. Receivables Turnover/Days Sales in Ending AR, 1994-96

<u>Year</u>	<u>Receivables Turnover</u>	<u>Days Sales in Ending A/c Rbles</u>
1994	73.6 times	7.0 days
1995	61.5 times	7.4 days
1996	43.5 times	10.4 days

Management analysis: All trade receivables represent amounts due from various national credit card companies. To date the company has had no bad debt experience with this type of receivable. Credit use by customers has been increasing over time.

27. Earnings Per Share, 1994-96

<u>Year</u>	<u>Earnings Per Share</u>		<u>Increase over Prior Year</u>	
	<u>basic</u>	<u>diluted</u>	<u>basic</u>	<u>diluted</u>
1994	\$0.68	\$0.67	112.5%	109.4%
1995	\$0.76	\$0.74	11.8%	10.4%
1996	\$1.10	\$1.05	44.7%	41.9%

Source of dilution: share options (directors, officers, and key employees)

Weighted average of outstanding shares have increased from 876,852 (primary and fully-diluted) in 1992 to 1,238,000 primary (1,333,238 fully diluted) for 1996.

28. Price-Earnings Ratio, 1994-96

<u>Year</u>	<u>Price-Earnings Ratio</u>
1994	12.13:1
1995	2.10:1
1996	7.38:1

Management analysis: 1995 drop due to poor fourth quarter 1995 earnings resulting from sales promotion's reduction of profits.

RESULTS OF THE 1996 AUDIT PROCEDURES

29. Results of Physical Inventory Observation at Restaurants

Each unit manager performs a stocktake on the final night of the financial year and forwards the counts to corporate headquarters for inclusion in the final inventory summary. In FY 96, the audit teams observed the physical inventory at 12 tested units, test-counted all the larger and more expensive items and also counted some of the smaller items. Inventory cutoff was also tested.

The conclusions of the observers were:

(1) No exceptions to the physical counts noted at any unit.

(2) All 12 units had larger than normal receipts during the last week of the FY, but no cutoff problems were noted (i.e., all restaurants receiving reports matched properly with all PWD shipping reports for the cutoff period). Several unit managers said they had been held to very "tight" inventory levels in the previous several months and welcomed the increase.

(3) One manager complained severely about a change in company policy: In previous years, unit managers retained one copy of their physical inventory counts and forwarded the other copies to corporate headquarters. This year, per a memo from D. Chioni, Internal Auditor, all copies were to be sent to headquarters. The unit manager feels this is a loss of valuable information to his unit at the start of the financial year. Chioni has said that he will send the Manager a photocopy of the count sheets and that the policy was changed to keep better control of any changes made to perpetual inventory records as a result of the physical count.

30. How Restaurant Test Units for Observation Were Chosen

In FY 96, 12 of the 41 units were selected as test units for a 29.3% coverage (FY 95: 16 out of 34 units [47% coverage]).

The reduction in scope since the previous financial year was due to five factors: (1) no significant problems had been noted in prior observations; (2) there was steady improvement over the past 4 years in internal control; (3) Nash had shown increasing sophistication as an accountant; (4) the Internal Audit function was functioning well in 1996; and (5) FY 96 tests of compliance with inventory-related controls revealed no exceptions.

The 12 units to be observed were selected in August to include the 6 largest units and a random selection of 6 other units. Details of the unit inventory observation were coordinated with Chioni well in advance of FYE. All preparations and arrangements were complete by October 1.

31. Results: Analytical Comparisons/Reviews, Unobserved Units

Audit tests of the restaurant units where the physical inventory was not observed consisted of:

- 1) All multiplication and addition on inventory count sheets tested.

Results: no arithmetical errors noted.

- 2) Compared current year quantity to prior year quantity for each item (at the same unit).

Results: At Units 10 and 16, some items quantities DOUBLED at these units. However, the dollar amounts of these items are immaterial in relation to inventory as a whole as the entire inventory at Units 10 and 16 accounts for only about 1% of total inventory.

- 3) Compared current year quantity (expressed as a percentage of inventory) for each item at the unobserved unit to the quantity for the same item at the 12 observed units.

Results: No significant differences in level of inventory at unobserved units vs. observed units.

32. Results of Physical Inventory Observation At The Purchasing and Warehousing Division (PWD)

Tucker House did a complete inventory at the PWD at the close of business Saturday October 30 (the PWD is closed on Sundays). The audit team observers performed the following procedures:

- 1) For meat: Perform a 100% check that the number of boxes in the freezers agrees with the client count sheet. Spot-check some boxes to make sure the item/box count is correct and boxes are full.
- 2) For other items: Take numerous test counts.
- 3) For all items: Tie the physical counts to the Weekly Inventory Summary (perpetual quantity records) for Week 52, FY 96.

Results: Minor differences for a few of the tested items were adjusted on the count sheets by the PWD Manager. No material differences between the counts and the Weekly Inventory Summary quantities were noted. The freezer doors were sealed and the auditors were present at the opening of business Monday November 1 to observe that the seals were unbroken.

33. Results of Inventory Pricing Tests

Results of the pricing tests generally indicated the following relationship between unit price used to value ending inventory (which was obtained as "aggregate dollar value of perpetual inventory per perpetual inventory records divided by number of units on hand per perpetual inventory records), most recent cost processed, and most recent cost purchased:

	<u>Unit Price</u>	<u>\$ -Processed</u>	<u>\$-Purchased</u>
Supreme steak sandwich	\$1.25	\$1.03	\$1.24
Steak sandwich	\$2.45	\$2.14	\$2.46
Deluxe burger	\$.98	\$.70	\$.94

Note: the unit quantity for each item is a single-serving portion.

The relationship of prices for other meat and non-meat items was similar to the above. Prices used to value inventory at the restaurants were identical to those used to value PWD inventory.

34. Results of Inventory Cutoff Tests

To test inventory cutoff, the following audit procedures were performed:

- 1) For all additions listed on the Weekly Inventory Summary for Week 52, the related receiving report and invoice were examined and the item was traced to Accounts Payable.

Result: All payables were recorded properly. No unrecorded items noted.

- 2). All invoices received in the first two weeks of FY97 were checked for any unrecorded FY 96 liabilities.

Result: No unrecorded items found.

35. Audit Procedures Performed Related to Purchase Commitments

On 2/10/96, an order for 113,600 kgs. of portion-controlled 260-gms steaks for Supreme steak sandwich at \$4.66 per kg. was placed with Canning Beef Company (CBC) for delivery from November 1996 through February 1997. Delivery, per usual, was to be by Tucker House trucks. On 16/10 and 26/10, cheques for \$65,000 each were paid to CBC as a \$130,000 advance on this order.

Per David White, Vice President of Purchasing and Production, it is the policy of CBC to request small advance payments on long-term orders at a set price. Although Tucker House has never been requested to pay such an advance before, White feels that the payment of such advances may become more common due to the general state of the economy.

Audit steps taken: The standard industry practice is to make meat purchases orally without any written confirming purchase order. However, due to the nature of this purchase, Tucker House sent a confirming purchase order to CBC on October 30 and requested the return of a signed copy. The returned P. O. signed by the Deputy General Manager of CBC, was examined and agreed in all aspects to the terms described above. The two cancelled cheques were also examined. The \$130,000 advance was originally booked by Tucker House as an Account Receivable, but the client has agreed that this should be reclassified as a prepaid for the financial report. Disclosure is complete.

36. 1996 Management Letter Comments Re Internal Control Weaknesses

Overall, audits have found the books and records to be in good condition. However, two significant internal control weaknesses are present:

- 1) Depreciation schedules are not updated on a timely basis and frequently contain errors such as failure to record transfers of items between units.
- 2) Regarding control of Accounts Payable: Most invoices are received by the Accounts Payable Department from the restaurants, the PWD or directly from the vendors. Others are received directly by department heads. Priority items such as beverages, food and utilities are recorded in the Purchases Journal immediately without further approval. However, invoices for non-priority items such as advertising and repairs and maintenance are sent to the applicable department heads for approval prior to being recorded in the Purchases Journal. No control is maintained by the A/c Pbles Department on the invoices sent out for approval. Also, vendor reconciliations are not performed on any regular basis, but only "as needed to resolve disputes".

37. Evaluation of Effectiveness of Internal Audit Function

The Internal Audit Staff was established in 1994. Review of the backgrounds of the staff and the Internal Audit reports indicates that the staff is functioning effectively. Don Chioni, the head of Internal Audit, has been very cooperative in coordinating physical inventory observation and planning for the most efficient utilisation of Tucker House personnel for the external audit.

38. 1996 Solicitor's Letter

The company is involved in no litigation other than ordinary routine litigation incidental to its business per the Solicitor's Letter dated 26/11/96 from Tucker House's legal adviser.

Tucker House has used the same legal adviser since its inception.

INVENTORY POLICIES/BACKGROUND AND SELECTED CLIENT RECORDS

39. Purchasing and Warehousing Division (PWD) History (Size, Expansions), 1992-96

The PWD was established in April 1992 in leased premises in a suburb of Melbourne. The initial term of the lease was for 10 years. In May 1995, the facilities were expanded by leasing an additional 1,100 sq. metres (bringing the total size to its current level of 2,100 sq. metres, of which 250 sq. metres is freezer space).

40. Beef Processing and Purchasing Policies

Although the company purchases some portion-controlled beef cuts from independent meat processors, almost 80% of its beef requirements are processed and supplied by the company's PWD. The PWD purchases beef in large quantities in anticipation of seasonal price fluctuations and, as a result, is better able to control costs and maintain stable menu prices. The company normally stocks approximately 12 weeks inventory of beef at the PWD although greater quantities are maintained or contracted for when it is considered economically advantageous to do so. All shipments of meat, both purchases and deliveries, are made in Tucker House's own refrigerated trucks.

41. Accounting Records/Internal Controls for PWD Inventory

Meat is inventoried on a perpetual basis with the quantity recorded in units consisting of a single-portion of the main meal. Thus, for example, a unit of Steak sandwich is a 230-gm. portion of steak and a unit of Beef Doner kebab is three doner kebabs.

Each unit of meat is sealed in vacuum packaging to preserve freshness and is stored either 40 or 80 units to a box depending on size. The boxes are plain brown boxes made specifically for meat storage and are labelled with a Health Dept Seal (the PWD is inspected and licensed by the Health Dept) and a Tucker House colour-coded labels.

The Weekly Inventory Summary lists by item, in units only, the beginning inventory, additions during the week, shipments during the week, and ending inventory. This document, along with Weekly Sales Reports (by item) from the restaurant units, is the primary source of information for purchasing and production decision making. A separate document, called the Flash, shows the dollar totals for each item currently in inventory.

Non-meat inventories are similarly accounted for.

42. Sources (Vendors) for Food, Beverages and Supply Purchases

Purchases of all inventoried items are made from suppliers unaffiliated with the company. No one supplier has accounted from year to year for a significant portion of the company's purchases, although in a particular year a single supplier may account for a significant portion.

During FY 96, one supplier of beef, Canning Beef Company (CBC), accounted for approximately 17% of the company's total food and beverage purchases and the company continues to purchase a major portion of its beef from that supplier.

Alternative sources of food and beverage products are readily available.

43. Client Records: Receiving Log for Week 52, FY 96

All receiving reports are consecutively numbered. Items received:

25/10 - 360 kgs. diced chicken meat
1,360 kgs. tomatoes
4,100 kgs. bulk beef

26/10 - 450 kgs. cheddar cheese
230 kgs. Swiss cheese

27/10 - 14,500 kgs. steak sandwich
13,100 kgs. steak sandwich

28/10 - 1,180 kgs. steak sandwich
500,000 paper napkins
900 kgs. tomatoes

29/10 - 110 kgs. pickles
360 dozen plastic cups

30/10 - 31/10 - dock closed

44. Client Records: Receiving Log For Week 1, FY97

All receiving reports are consecutively numbered. Items received:

1/11 - 1,460 kgs. tomatoes
680 kgs. corn flour

2/11 - 230 kgs. wheat flour
450 kgs. shortening

3/11 - 200 litres disinfectant

4/11 - 900 kgs. tomatoes

5/11 - no shipment received

6/11 and 7/11 - dock closed

45. Client Records: Weekly Inventory Summary, Week 51, FY 96

In portion-controlled single-serving units, three largest items:

	<u>Steak sandwich</u> (230-gm)	<u>Supreme steak</u> <u>sandwich</u> (260-gm)	<u>Deluxe burger</u> (2)
Beginning Inventory	42,800	175,000	145,000
Additions:			
Processed	-0-	-0-	-0-
Purchased	-0-	-0-	-0-
Shipments to Restaurants	(5,600)	(20,480)	(14,200)
<u>Ending Inventory</u>	<u>37,200</u>	<u>154,520</u>	<u>130,800</u>

46. Client Records: Weekly Inventory Summary, Week 52, FY 96

In portion-controlled single-serving units, three largest items:

	<u>Steak sandwich</u> (230-gm)	<u>Supreme steak sandwich</u> (260-gm)	<u>Deluxe burger</u> (2)
Beginning Inventory	197,200	154,520	130,800
Additions:			
Processed	-0-	-0-	50,000
Purchased	5,200	-0-	-0-
Shipments to Restaurants	<u>(8,400)</u>	<u>(31,520)</u>	<u>(21,800)</u>
Ending Inventory	<u>194,000</u>	<u>123,000</u>	<u>159,000</u>

47. Client Records: Physical Inventory Counts, 1/11/95

Physical inventory counts at the PWD for FY 95, 3 largest items:

Supreme steak sandwich - 204,000 units (serving)

Steak sandwich - 55,000 units

Deluxe burger - 141,000 units

48. **Client Records: Weekly Inventory Summary, Week 1, FY97**

In portion-controlled single-serving units, three largest items:

	<u>Steak sandwich</u> (230-gm)	<u>Supreme steak sandwich</u> (260-gm)	<u>Deluxe burger</u> (2)
Beginning Inventory	34,000	283,000	159,000
Additions:			
Processed	23,600	-0-	-0-
Purchased	-0-	-0-	-0-
Shipments to Restaurants	<u>(2,600)</u>	<u>(9,480)</u>	<u>(7,000)</u>
Ending Inventory	<u>55,000</u>	<u>273,520</u>	<u>152,000</u>

When you are ready, please read the information presented on page 25.

Welcome to the group section of the experiment.

For this experiment, assume that you have been assigned as a member of an audit group for the audit of the inventory balance of Tucker House Co Ltd. Please examine the enclosed materials and follow the instructions closely. As you complete the various tasks, you should insert them into the envelope provided. You are permitted to refer to the audit working papers of Tucker House Co Ltd throughout the experiment. Please place all the experimental materials into the envelope at the end of the experiment.

Please do not turn over the page until all the group members are ready.

Task 1

You should discuss and work on the case as a group to arrive at a consensus on the following questions.

Based on your group's evaluation of the case materials, indicate whether the inventory of Tucker House Co Ltd is or is **not** fairly presented by placing a tick (\checkmark) in the appropriate box.

- [] The inventory balance of Tucker House Co Ltd is fairly presented.
- [] The inventory balance of Tucker House Co Ltd is **not** fairly presented.

Indicate your group's confidence that the above answer is correct by placing a cross (X) on the line below.

0	10	20	30	40	50	60	70	80	90	100
Not confident at all										Completely confident

Based on your group's evaluation of the information, please indicate how likely does your group think that the inventory balance may be **fairly** or **not fairly** presented:

1	2	3	4	5	6	7	8	9
Fairly presented								Not fairly presented

Indicate your group's confidence that the above answer is correct by placing a cross (X) on the line below.

0	10	20	30	40	50	60	70	80	90	100
Not confident at all										Completely confident

When you have answered these questions, please insert this section into the envelope provided and proceed to Task 2.

Do not turn over the page until you have inserted Task I into the envelope provided.

Task 2

There are many ways by which people can check the appropriateness of their decision. One of these is to require your group to consider why the group's decision in Task 1 may be **incorrect**.

Please discuss among yourselves the reasons why your previous answer may be incorrect and/or why the alternative you rejected may be correct.

(Note: The reasons you will be considering for this task should be regarded as 'pre-working notes'. That is, your reasons will not go directly into the planning audit working papers but will serve as potential material for a subsequent working paper examination.)

When you have completed your discussion, please proceed to the next page. Do not turn over the page until you are ready.

Do not turn over the page until have you inserted Task 3 into the envelope provided.

Task 4

You should discuss and work as a group to arrive at a consensus on the following questions.

Based on your group's evaluation of the case materials, indicate whether the inventory of Tucker House Co Ltd is or is not fairly presented by placing a tick (✓) in the appropriate box.

- The inventory balance of Tucker House Co Ltd is fairly presented.
- The inventory balance of Tucker House Co Ltd is not fairly presented.

Indicate your group's confidence that the above answer is correct by placing a cross (X) on the line below.

-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|

0 10 20 30 40 50 60 70 80 90 100

Not confident at all Completely confident

Based on your group's evaluation of the case materials, please indicate how likely does your group think that the inventory balance may be fairly or not fairly presented:

1 2 3 4 5 6 7 8 9

Fairly presented Not fairly presented

Indicate your group's confidence that the above answer is correct by placing a cross (X) on the line below.

-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|

0 10 20 30 40 50 60 70 80 90 100

Not confident at all Completely confident

Generating the reasons to support why your group's judgment may be incorrect was

1 2 3 4 5 6 7 8 9

Very easy Very difficult

This marks the end of the group section of this research project. The next section asks you questions about yourself and should be completed individually.

Part II: Your background: Please tick (✓) the appropriate spaces provided below or circle an appropriate number.

1. Your gender?
 Male Female
2. Your age? yrs
3. Is English your first or native language?
 Yes No
4. I have _____ years and _____ months working experience of which _____ years and _____ months are in auditing.
5. What is your position in your firm?
 Audit Junior Audit Senior Supervisor

6. I am satisfied with my group's decisions

Strongly disagree	1	2	3	4	5	6	7	Strongly agree
----------------------	---	---	---	---	---	---	---	-------------------

7. How would you rate the amount of interaction in your group's audit process?

1	2	3	4	5	6	7	8	9
No interaction at all				Moderate amount of interaction				A lot of interaction

8. How difficult did you find the tasks given to you previously in the main experimental part of this research project?

1	2	3	4	5	6	7	8	9
Not at all difficult				Somewhat difficult				Extremely difficult

9. How much effort do you think an actual audit of inventory balance will require?

1	2	3	4	5	6	7	8	9
Very little effort				Some effort				A great deal of effort

10. How much effort did you expend on the main experimental section of this research project?

1	2	3	4	5	6	7	8	9
Very little effort				Some effort				A great deal of effort

**This marks the end of the experiment. Please place all research materials into the envelope provided (including materials from the tasks you performed individually).
 Thank you for your cooperation in this research project.**