AN EXPERIMENTAL STUDY OF THE EFFECTS OF THE USE OF AN EXPERT SUPPORT SYSTEM AND ITS EXPLANATION FACILITIES ON GROUP DECISION MAKING

by

Fui Hoon (Fiona) Nah

B.Sc., National University of Singapore, 1988
B.Sc. (Honours), National University of Singapore, 1989
M.Sc., National University of Singapore, 1992

A THESIS SUBMITTED IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY in
THE FACULTY OF GRADUATE STUDIES
(Business Administration – Management Information Systems)

We accept this thesis as conforming to the required standard

THE UNIVERSITY OF BRITISH COLUMBIA
October 1997
© Fui Hoon (Fiona) Nah, 1997
In presenting this thesis in partial fulfilment of the requirements for an advanced degree at the University of British Columbia, I agree that the Library shall make it freely available for reference and study. I further agree that permission for extensive copying of this thesis for scholarly purposes may be granted by the head of my department or by his or her representatives. It is understood that copying or publication of this thesis for financial gain shall not be allowed without my written permission.

Department of Commerce
The University of British Columbia
Vancouver, Canada

Date Oct 15, 1987
ABSTRACT

As information technology is increasingly used in organizations to support group work, it is important to understand how group decision making processes are moderated by the provision of computer-based decision support facilities. This research examines the effects of utilizing an Expert Support System (ESS) and its explanation facilities on group decision making. Four persuasion theories form the main theoretical foundations of this research: the elaboration likelihood model, the social judgment theory, the information processing paradigm, and the cognitive response theory.

An experimental study was carried out to examine the suitability of using an ESS to support groups in making financial analysis decisions. Three levels of decision support — no ESS support, ESS analyses without explanations support, ESS analyses plus explanations support — were examined. Two groups of subjects — experts and novices — participated in the study.

The findings are consistent with the widely-held belief that for an ESS to be useful, both ESS analyses and explanations support are necessary; they both contribute to knowledge transfer from the ESS to the novice decision makers. ESS explanations also increase users’ trust in the system. ESS support, however, decreases users’ satisfaction with the group process. The expert-novice comparison shows that novices find the ESS to be more useful than experts do. Experts are not only more capable of processing the available information, but they also tend to be more critical and ego-involved in their area of expertise. These characteristics decrease experts’ likelihood of being persuaded by the ESS and account for the lower consensus among the experts compared to novices.
This research represents one of the first studies to investigate the use of ESS technology in group settings. It integrates quantitative, statistical, and positivist methods with qualitative, case, and interpretive methods to provide a rich understanding and description of the group processes and outcomes. In terms of theoretical contributions, it integrates persuasion theories into research on the use of ESS technology for group decision making. For practitioners and managers, the findings indicate that a high quality ESS, with both its analyses and explanations components, could be used to improve the quality of group judgments.
### TABLE OF CONTENTS

Abstract .................................................................................................................. ii
Table of Contents ................................................................................................... iv
List of Tables ......................................................................................................... viii
List of Figures ........................................................................................................ xviii
Acknowledgements ............................................................................................... xix
Dedication ............................................................................................................... xx

#### Chapter 1. Introduction .................................................................................. 1
  1.1 Background ..................................................................................................... 2
  1.1.1 Expert Support Systems ........................................................................... 3
  1.1.2 Group Decision Support Systems ........................................................... 6
  1.2 Research Objectives and Motivation ............................................................. 7
  1.3 Significance of Research .............................................................................. 9
  1.4 Conduct of Research .................................................................................... 10
  1.5 Summary of Chapter ................................................................................... 11
  1.6 Organization of Dissertation ...................................................................... 11

#### Chapter 2. Literature Review ....................................................................... 13
  2.1 Expert System Technology ........................................................................... 13
  2.1.2 Components of ESS/ES ......................................................................... 15
  2.1.3 Explanation Facilities ............................................................................. 15
  2.1.4 Empirical Studies on Expert System Technology ................................... 25
  2.2 Relationship with GDSS Research ............................................................... 34
  2.2.1 Studies on Group Support Using Decision Modeling and Analysis Techniques ........................................................................................................ 35
  2.2.2 Meta-Analyses on GDSS Research ............................................................ 38
  2.3 Summary of Chapter ................................................................................... 39

#### Chapter 3. Theoretical Foundation ............................................................... 41
  3.1 Related Literature on Group Decision Making ............................................ 41
  3.1.1 Group Process Gains and Losses ............................................................ 41
  3.1.2 Informational and Normative Influence .................................................. 44
<table>
<thead>
<tr>
<th>3.1.3 Expert Power</th>
<th>45</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.2 Related Literature on Persuasion</td>
<td>47</td>
</tr>
<tr>
<td>3.2.1 Related Theories on Persuasion</td>
<td>48</td>
</tr>
<tr>
<td>3.2.2 Factors Influencing Persuasion</td>
<td>61</td>
</tr>
<tr>
<td>3.3 Expert-Novice Differences</td>
<td>67</td>
</tr>
<tr>
<td>3.3.1 Architecture of Adaptive Control of Thought (ACT)</td>
<td>69</td>
</tr>
<tr>
<td>3.3.2 Three-Stage Learning Model</td>
<td>71</td>
</tr>
<tr>
<td>3.4 Lens Model Framework</td>
<td>73</td>
</tr>
<tr>
<td>3.4.1 Mapping between ESS Support and Lens Model</td>
<td>74</td>
</tr>
<tr>
<td>3.5 Summary of Chapter</td>
<td>77</td>
</tr>
</tbody>
</table>

**Chapter 4. Research Framework, Design, and Hypotheses**

<table>
<thead>
<tr>
<th>4.1 Research Framework</th>
<th>78</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.2 Research Design</td>
<td>83</td>
</tr>
<tr>
<td>4.3 Derivation of Hypotheses on Effects of ESS Support</td>
<td>86</td>
</tr>
<tr>
<td>4.3.1 Effects of ESS Analyses Support on Judgments and Consensus</td>
<td>86</td>
</tr>
<tr>
<td>4.3.2 Effects of ESS Explanations Support on Judgments and Consensus</td>
<td>88</td>
</tr>
<tr>
<td>4.3.3 Effects of ESS Analyses and Explanations Support on Perceptions</td>
<td>89</td>
</tr>
<tr>
<td>4.4 Derivation of Hypotheses on Expert-Novice Differences</td>
<td>91</td>
</tr>
<tr>
<td>4.4.1 Effects of ESS Analyses and Explanations Support on Expert-Novice Judgments</td>
<td>91</td>
</tr>
<tr>
<td>4.4.2 Effects of ESS Analyses and Explanations Support on Expert-Novice Perceptions</td>
<td>93</td>
</tr>
<tr>
<td>4.5 Summary of Chapter</td>
<td>95</td>
</tr>
</tbody>
</table>

**Chapter 5. Research Methodology**

<table>
<thead>
<tr>
<th>5.1 Experimental Procedures</th>
<th>96</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.2 Group Effect, and ESS Analyses and Explanations Effect</td>
<td>99</td>
</tr>
<tr>
<td>5.3 Subject Characteristics</td>
<td>100</td>
</tr>
<tr>
<td>5.4 Experimental Task</td>
<td>100</td>
</tr>
<tr>
<td>5.5 Experimental ESS</td>
<td>102</td>
</tr>
<tr>
<td>5.5.1 Explanation Facilities of FINALYZER</td>
<td>102</td>
</tr>
<tr>
<td>5.5.2 Description of FINALYZER</td>
<td>103</td>
</tr>
</tbody>
</table>
5.6 Dependent Variables ................................................................. 105
  5.6.1 Quantitative Analysis ............................................................... 106
  5.6.2 Reliability and Validity of Perception Measures ........................ 113
  5.6.3 Qualitative Analysis .............................................................. 125

5.7 Summary of Chapter ................................................................. 125

Chapter 6. Results of Quantitative Analyses – Part I ................................ 127
  6.1 Statistical Analyses Employed ..................................................... 128
    6.1.1 Evaluation of Assumptions of Statistical Tests ......................... 131
  6.2 Analysis of Consistency and Consensus of Judgments .................... 132
    6.2.1 Comparison of Novice Subjects’ Performance Across Levels of ESS Support ................................................................. 133
  6.3 Comparison of Novice Subjects’ Perceptions Across Levels of ESS Support ................................................................. 163
  6.4 Summary of Chapter ................................................................. 176

Chapter 7. Results of Quantitative Analyses – Part II ................................ 177
  7.1 Comparison of Novices versus Expert Subjects’ Performance with ESS Analyses and Explanations Support ................................. 177
    7.1.1 Summary of Comparison of Novices’ versus Experts’ Judgments with ESS Support ................................................................. 177
    7.1.2 Pre-test – Experts’ versus Novices’ Initial Judgments ................. 178
    7.1.3 Group Judgments ................................................................. 182
    7.1.4 Change in Deviation Score from Individual Pre-discussion to Group Judgments ................................................................. 184
    7.1.5 Individual Post-discussion Judgments .................................... 187
    7.1.6 Change in Deviation Score from Initial to Post-discussion Individual Judgments ................................................................. 189
    7.1.7 Consensus in Judgments ....................................................... 193
    7.1.8 Summary of Results ............................................................ 196
  7.2 Comparison of Novices’ versus Experts’ Perceptions with Full ESS Support ................................................................. 197
    7.2.1 Summary of Comparison of Novices’ versus Experts’ Perceptions ................................................................. 197
    7.2.2 Perception Measure – Satisfaction with Group Process ............... 198
    7.2.3 Perception Measure – Satisfaction with Group Judgments .......... 200
7.2.4 Perception Measure – Perceived Usefulness of ESS .......................... 201
7.2.5 Perception Measure – Trust in ESS .................................................. 203
7.3 Summary of Chapter ............................................................................. 204

Chapter 8. Results of Qualitative Analysis ................................................. 206
8.1 Qualitative Analysis to Support Quantitative Analysis ....................... 207
8.1.1 Knowledge Transfer to Users Increases with Level of ESS Support .......................................................................................................................... 207
8.1.2 ESS Analyses and Explanations Support Influence Novices More than Experts ........................................................................................................... 229
8.1.3 Higher Consensus Among Novices than Experts .............................. 236
8.2 Summary of Chapter ............................................................................. 236

Chapter 9. Discussion and Conclusions ..................................................... 238
9.1 Summary and Implications of Research Findings ............................... 238
9.1.1 Effects of ESS Support on Group Decision Making by Novices .. 238
9.1.2 Effects of ESS Support on Group Decision Making by Novices versus Experts ........................................................................................................ 241
9.1.3 Supplementary Findings .................................................................. 244
9.1.4 Implications for Research ................................................................. 246
9.1.5 Implications for Practice ................................................................. 248
9.2 Contributions of Research .................................................................. 249
9.3 Limitations, Problems and Challenges ................................................. 251
9.4 Directions for Future Research ............................................................... 254

Bibliography .............................................................................................. 257

List of Appendices ...................................................................................... 276
A. Experimental Materials ........................................................................ 277
B. Sample Screens of FINALYZER ............................................................. 325
C. Materials for Recruitment of Subjects .................................................. 343
# LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2-1</td>
<td>Correspondence between Southwick's and Chandrasekaran et al.'s Categorizations of Explanation Types</td>
</tr>
<tr>
<td>4-1</td>
<td>Research Design</td>
</tr>
<tr>
<td>5-1</td>
<td>Definition of Explanation Types</td>
</tr>
<tr>
<td>5-2</td>
<td>Dependent Variables in Quantitative Analysis</td>
</tr>
<tr>
<td>5-3</td>
<td>An Example for Illustration Purposes</td>
</tr>
<tr>
<td>5-4</td>
<td>An Example to Illustrate Computation of Absolute Deviation from Original Experts' Judgments</td>
</tr>
<tr>
<td>5-5</td>
<td>Items to Measure Perceived Consensus of Group Judgments</td>
</tr>
<tr>
<td>5-6</td>
<td>Items to Measure Satisfaction with Group Process</td>
</tr>
<tr>
<td>5-7</td>
<td>Items to Measure Satisfaction with Group Judgments</td>
</tr>
<tr>
<td>5-8</td>
<td>Items to Measure User Trust in ESS</td>
</tr>
<tr>
<td>5-9</td>
<td>Items to Measure Perceived Usefulness of ESS</td>
</tr>
<tr>
<td>5-10</td>
<td>Item Reliability Statistics of “Satisfaction with Group Process” Scale</td>
</tr>
<tr>
<td>5-11</td>
<td>Item Reliability Statistics of “Satisfaction with Group Judgments” Scale</td>
</tr>
<tr>
<td>5-12</td>
<td>Item Reliability Statistics of “Perceived Consensus in Group Judgments” Scale</td>
</tr>
</tbody>
</table>
5-13 Item Reliability Statistics of “Perceived Usefulness of ESS” Scale (Novices Only) ..................................................................................................................118

5-14 Item Reliability Statistics of “Perceived Usefulness of ESS” Scale (Novices and Experts) ........................................................................................................................................118

5-15 Item Reliability Statistics of Revised “Perceived Usefulness of ESS” Scale (Novices and Experts) .........................................................................................................................................119

5-16 Item Reliability Statistics of “Trust in ESS” Scale (Novices Only) ..........................................120

5-17 Item Reliability Statistics of “Trust in ESS” Scale (Novices and Experts) ...............120

5-18 Overall Reliability of the Five Perception Scales .................................................................121

5-19 Rotated Factor Loadings ........................................................................................................122

5-20 Item Reliability Statistics of Revised “Satisfaction with Group Judgments” Scale .................................................................................................................................124

5-21 Cronbach’s Alpha, Eigenvalues, and Variances Explained by the Factors ........125

6-1 Research Design and Number of Groups in Each Treatment ........................................127

6-2 Appropriate Test Statistic for Nested Two-Factor Designs with Fixed and Random Factor Effects (B nested within A) .................................................................................................131

6-3 Summary of Results on Consistency of Novices’ Judgments ........................................134

6-4 Summary of Results on Consensus of Novices’ Judgments ........................................134

6-5 Descriptive Statistics of Consistency of Individual Pre-discussion Judgments ........136

6-6 Results of Analysis of Consistency of Individual Pre-discussion Judgments ........136
6-20 Results of Kruskal-Wallis Rank Test — Change from Individual Pre-discussion to Group Judgments Analyzed at the Group Level .......................................................... 144
6-21 Mean Ranks of Kruskal-Wallis Test — Change from Individual Pre-discussion to Group Judgments Analyzed at the Group Level ...................................................... 145
6-22 Results of Post-hoc Comparisons — Change from Individual Pre-discussion to Group Judgments Analyzed at the Group Level ......................................................... 146
6-23 Results of Multiple Pairwise Comparisons for Kruskal-Wallis Test — Change from Individual Pre-discussion to Group Judgments Analyzed at the Group Level .... 146
6-24 Results of Pairwise Comparisons using Mann-Whitney U Test — Change from Individual Pre-discussion to Group Judgments Analyzed at the Group Level .... 147
6-25 Descriptive Statistics of Consistency of Individual Post-Discussion Judgments Analyzed at the Group Level .................................................................................. 150
6-26 Results of Kruskal-Wallis Rank Test — Consistency of Individual Post-discussion Judgments .............................................................................................................. 150
6-27 Mean Ranks of Kruskal-Wallis Test — Consistency of Individual Post-discussion Judgments .............................................................................................................. 151
6-28 Results of Multiple Pairwise Comparisons for Kruskal-Wallis Test — Consistency of Individual Post-discussion Judgments ........................................................................ 151
6-29 Results of Pairwise Comparisons using Mann-Whitney U Test — Consistency of Individual Post-discussion Judgments ........................................................................ 152
6-30 Descriptive Statistics of Change in Deviation Score from Individual Pre- to Post-discussion Judgments ...................................................................................... 154
6-31 Results of Nested ANOVA — Difference in Consistency of Individual Pre- and Post-discussion Judgments ................................................................................. 154
Descriptive Statistics of Change in Deviation Score from Individual Pre- to Post-discussion Judgments Analyzed at the Group Level .................................................. 155

Results of Analysis of Difference in Consistency of Individual Pre- and Post-discussion Judgments ............................................................................................................. 155

Results of ANOVA — Change from Individual Pre- to Post-discussion Judgments Analyzed at the Group Level .................................................................................. 155

Results of Kruskal-Wallis Rank Test — Change from Individual Pre- to Post-discussion Judgments Analyzed at the Group Level ......................................................... 156

Mean Ranks of Kruskal-Wallis Test — Change from Individual Pre- to Post-discussion Judgments Analyzed at the Group Level ......................................................... 156

Results of Post-hoc Comparisons — Change from Individual Pre- to Post-discussion Judgments Analyzed at the Group Level ......................................................... 157

Results of Multiple Pairwise Comparisons for Kruskal-Wallis Test — Change from Individual Pre- to Post-discussion Judgments Analyzed at the Group Level ............................. 157

Results of Pairwise Comparisons using Mann-Whitney U Test — Change from Individual Pre- to Post-discussion Judgments Analyzed at the Group Level ................. 158

Descriptive Statistics of Total Absolute Distance between Group Members’ Post-discussion Individual Judgments Analyzed at the Group Level ................................. 160

Mean Ranks of Kruskal-Wallis Test — Total Absolute Distance between Group Members’ Post-discussion Individual Judgments ................................................................. 160

Results of Kruskal-Wallis Rank Test — Total Absolute Distance between Group Member’s Post-discussion Individual Judgments ................................................................. 160
7-6 Results of $t$ Test — Consistency of Individual Pre-discussion Judgments ........................................ 181

7-7 Results of Mann-Whitney $U$ Test — Consistency of Individual Pre-discussion Judgments ................................................................................................................................. 181

7-8 Descriptive Statistics of Deviation of Group Judgments of Expert and Novice Groups from Judgments of Original Experts ........................................................ 183

7-9 Results of Mann-Whitney $U$ Test — Consistency of Group Judgments of Experts versus Novices with Judgments of Original Experts ....................................................... 183

7-10 Results of $t$ Test — Consistency of Group Judgments of Experts versus Novices ................................................................................................................................. 184

7-11 Descriptive Statistics of Change in Deviation Score from Individual Pre-discussion to Group Judgments Analyzed at the Group Level ............................................. 185

7-12 Results of Analysis of Change in Deviation Score from Individual Pre-discussion to Group Judgments of Experts versus Novices .......................................................... 186

7-13 Results of $t$ Test — Change in Deviation Score from Individual Pre-discussion to Group Judgments ........................................................................................................ 186

7-14 Results of Mann-Whitney $U$ Test — Change in Deviation Score from Individual Pre-discussion to Group Judgments .............................................................. 186

7-15 Descriptive Statistics of Consistency of Individual Post-Discussion Judgments of Experts versus Novices Analyzed at the Group Level ......................................... 188

7-16 Results of Mann-Whitney $U$ Test — Consistency of Individual Post-discussion Judgments of Experts versus Novices .................................................................................. 189

7-17 Descriptive Statistics of Change in Deviation Score from Individual Pre- to Post-discussion Judgments of Experts versus Novices Analyzed at the Individual Level ........................................................................................................... 190
7-18 Results of Nested ANOVA – Change in Deviation Score from Individual Pre- to Post-discussion Judgments of Experts versus Novices .............................................191

7-19 Descriptive Statistics of Change in Deviation Score from Individual Pre- to Post-discussion Judgments of Experts versus Novices Analyzed at the Group Level .................................................................191

7-20 Results of Analysis of Change in Deviation Score from Individual Pre- to Post-discussion Judgments of Experts versus Novices .........................................................192

7-21 Results of $t$ Test – Change in Deviation Score from Individual Pre- to Post-discussion Judgments of Experts versus Novices Analyzed at the Group Level .192

7-22 Results of Mann-Whitney $U$ Test – Change in Deviation Score from Individual Pre- to Post-discussion Judgments of Experts versus Novices Analyzed at the Group Level .................................................................192

7-23 Descriptive Statistics of Total Absolute Distance between Group Members’ Post-discussion Individual Judgments Analyzed at the Group Level .........................194

7-24 Results of Mann-Whitney $U$ Test – Total Absolute Distance between Group Members’ Post-discussion Individual Judgments .........................................................194

7-25 Descriptive Statistics of Total Absolute Distance of Group Members’ Post-discussion Individual Judgments from Group Judgments Analyzed at the Group Level ................................................................195

7-26 Results of Mann-Whitney $U$ Test – Total Absolute Distance of Group Members’ Post-discussion Individual Judgments from Group Judgments .........................196

7-27 Summary of Results of Novices’ versus Experts’ Perceptions .........................................................197

7-28 Descriptive Statistics of Perceived Satisfaction with Group Process (Individual Level) .........................................................................................................................199
7-29 Descriptive Statistics of Perceived Satisfaction with Group Process (Group Level) .................................................................199

7-30 Results of Analysis of Novices’ versus Experts’ Satisfaction with Group Process ..............................................................................199

7-31 Results of $t$ Test – Satisfaction with Group Process .................................................200

7-32 Results of Mann-Whitney $U$ Test – Satisfaction with Group Process .................200

7-33 Descriptive Statistics of Satisfaction with Group Judgments (Individual Level) ................................................................................201

7-34 Results of Nested ANOVA – Satisfaction with Group Judgments .........................201

7-35 Descriptive Statistics of Perceived Usefulness of ESS (Individual Level) ..........202

7-36 Results of Nested ANOVA – Perceived Usefulness of ESS ..................................202

7-37 Descriptive Statistics of Trust in ESS (Individual Level) ......................................204

7-38 Results of Nested ANOVA – Trust in ESS ..............................................................204

8-1 Average Number of Feedforward and Feedback Explanations Used by Experts and Novices ................................................................234
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3-1</td>
<td>Elaboration Likelihood Model</td>
</tr>
<tr>
<td>3-2</td>
<td>The Curvilinear (Inverted-U Shaped) Relationship between Message Discrepancy and Attitude Change</td>
</tr>
<tr>
<td>3-3</td>
<td>The ACT Architecture</td>
</tr>
<tr>
<td>3-4</td>
<td>Three-Stage Learning Model</td>
</tr>
<tr>
<td>3-5</td>
<td>Brunswik's Standard Lens Model</td>
</tr>
<tr>
<td>3-6</td>
<td>Levels of ESS Support Presented using the Lens Model Framework</td>
</tr>
<tr>
<td>4-1</td>
<td>Research Framework</td>
</tr>
<tr>
<td>5-1</td>
<td>Research Procedure</td>
</tr>
<tr>
<td>5-2</td>
<td>Experimental Setup</td>
</tr>
<tr>
<td>5-3</td>
<td>Flow Chart of FINALYZER</td>
</tr>
<tr>
<td>8-1</td>
<td>Use of Explanation Types by Novices and Experts</td>
</tr>
<tr>
<td>9-1</td>
<td>Use of Explanation Types by Novice Groups</td>
</tr>
<tr>
<td>9-2</td>
<td>Use of Explanation Types by Expert Groups</td>
</tr>
</tbody>
</table>
Acknowledgements

Many people contributed to the completion of this thesis. First and most significant is my advisor, Professor Izak Benbasat, who has fulfilled his role as an advisor in every way desired. He has been supportive, helpful, and encouraging throughout the dissertation process. I appreciate his taking time from his busy Associate Dean schedule to provide me feedback on the many drafts we have gone through. The rest of my dissertation committee has also contributed to this dissertation in their own unique ways. Professor Albert Dexter has not only been very supportive and helpful but has also contributed his accounting knowledge and practical experience to increase the realism and practical value of this study. I also thank him for lending me his laptop that was brought to the financial institutions for the experiment with expert subjects. Professor Kenneth MacCrimmon has contributed his expertise in the group decision making area and prompted a number of intriguing questions that have led to improvement in the quality of this dissertation. Doctor Joy Begley, who is the Accounting expert of my committee, has taken an active role in two main aspects of this dissertation work; firstly, in ensuring that FINALYZER, the ESS that was used in this study, maintained its integrity and relevance, and secondly, in helping me with the changes that were made to the case to tailor it for the experts.

I also thank Gerry DeSanctis, Joe Valacich, Len Jessup, Jiye Mao, Wynne Chin, Keng Siau and the MIS seminar group at UBC for their comments and suggestions on the pre-experiment proposal to this research. I appreciate the help of Gary Schwartz for lending me the necessary equipment, helping me with the experimental setup, and working out a schedule to use the room space for the experiment. Gary Schwartz has also trained me to become proficient in setting up the experiment at the site of the financial institutions. I thank Andrew Gemino and Robbie Nakatsu for giving me feedback on the earlier drafts of the dissertation. Colin Ho and Andrew Gemino have also helped me with some of the statistical problems I faced. Special thanks also to all the subjects who participated in the experiment.

I thank my husband, Keng Siau, for reviewing multiple drafts and for his support in the completion of this dissertation. Lastly, I thank my “little” sister, Hwee Shan Nah, for helping me with the preparation of PowerPoint slides for my first presentation on the results of this research in Singapore, my home country, in January 1997.
Dedication

This thesis is dedicated to my husband, Keng Siau, who prompted me to pursue a Ph.D.,

and

to my parents, Tua Bah Nah and Gek Lian Toh, for their encouragement and financial support during my Ph.D. studies.
CHAPTER 1: INTRODUCTION

This dissertation examines the efficacy of using an expert support system and its explanation facilities to support group decision making. Ever since the introduction of the expert system technology, it has been targeted at individuals. With the advancement of information technology and the emphasis on teamwork, such systems are becoming popular as a group decision support tool (e.g., Swann, 1988; Sviokla, 1989). For instance, at Imperial Chemical Industries (ICI), an expert support system (so called decision assistant knowledge-based system by Swann, 1988) that supports business planning is used to assist in group decision making (Swann, 1988). The system was not originally conceived as a group support tool but arose as a group support tool due to business needs. It is increasingly being applied in group sessions and is proving very effective in that mode of use.

The development cost of Expert Support Systems (ESS) and their explanation facilities is high due to the time and resources involved in the development process. These high costs make it imperative to learn more about the effects of these systems and their explanation facilities in supporting both individual and multi-individual decision making. Several studies have examined ESS use in supporting individual decision making (Dhaliwal, 1993; Eining and Dorr, 1991; Gregor, 1996, Hsu, 1993; Lamberti and Wallace, 1990; Mao, 1995; Moffitt, 1989; Murthy, 1990; Oz, Fedorowicz, and Stapleton, 1993; Peterson, 1988; Ye and Johnson, 1995). Despite the presence of expert support systems in group meetings (Swann, 1988; Sviokla, 1989), a literature search reveals that no experimental study has investigated the effects of expert support on group decision making. Lest one assumes it is an unimportant issue, the introduction of a new “partner”, especially one with specialized knowledge, can have a profound impact on the group processes and outcomes. Benbasat, DeSanctis, and Nault (1993) point out that this phenomenon needs to be investigated.
Given that one of the main factors leading to the failure of ESS is the lack of system acceptance (Gill, 1995), we investigate if explanation facilities would change decision makers’ satisfaction with group process and judgments as well as their acceptance of the advice of ESS. The usefulness of the system for group decision making is also assessed and compared between the experts and the novices. In general, this dissertation studies 1) the effect of providing different levels of ESS support for group decision making by novices, and 2) the effect of providing the complete ESS support, i.e., both the analyses and explanations support, on group decision making by experts and novices.

1.1 Background
Expert support systems (ESS) are an extension of the expert systems (ES) concept. Expert systems were among the earliest applications of Artificial Intelligence (AI) to be commercialized. They are computer-based software tools that use artificial intelligence techniques to capture, represent, and apply expert knowledge to mimic the behavior of human experts in specific narrowly-defined problem domains. The ability to explain knowledge and reasoning, often referred to as the explanation facilities, is considered to be one of ES’s most powerful components. These explanations are machine-generated descriptions of the operations of a system — what it does, how it works, and why its actions are appropriate. Traditionally, ES technology was used to replace human decision making by transferring knowledge from human experts to the systems. However, this approach has not worked very well for a couple reasons — it is difficult to model completely the expertise of the human experts in these systems and difficult to keep the systems up-to-date due to the evolving and dynamic nature of the environment. These problems are some of the reasons that resulted in the extension of the ES concept to the ESS concept (Luconi, Malone, and Scott Morton, 1986), which is employed as a support rather than a replacement technology. In other words, an ESS provides expert advice to complement the
knowledge of decision makers, and it supports, rather than replaces, the decision makers. An ESS combines the features of two important technologies – expert and decision support technologies. In line with findings from the literature (e.g., Benbasat and Lim, 1993), which indicate that task support is effective in supporting group decision making, the analyses and explanations features of the ESS technology have a high potential in supporting group decision making (Swann, 1988).

1.1.1 Expert Support Systems

Expert support systems have gained increasingly popular use in business organizations since the mid 1980s. Despite some setbacks, many companies remain enthusiastic proponents of the ES technology and continue to develop important applications based on the technology (Gill, 1995). Some of these applications include Digital’s XCON (Kraft, 1984; Leonard-Barton, 1987; Sviokla, 1990), American Express’ Authorizer’s Assistant (Feigenbaum, McCorduck, and Nii, 1988; Rothi and Yen, 1990), Coopers and Lybrand’s ExpertTax (Shpilberg, Graham, and Schatz, 1986), Chemical Bank’s FXAA (AI Week, 1988), and Carrier’s EXPERT (Heatley, Agarwal, and Tanniru, 1995). They are used to assist in a wide range of tasks including diagnosis, prediction, planning, and design, as well as to gain competitive advantage (Feigenbaum, McCorduck, and Nii, 1988; Heatley, Agarwal, and Tanniru, 1995; Liebowitz, 1990; Sviokla, 1990). These systems also help firms to generate financial returns. Some companies have even indicated that their key businesses depend on these systems and are likely to remain so in the future (Gill, 1995).

There are two major advantages for developing expert support systems. First, they capture, preserve, and disseminate the knowledge of scarce expertise by encoding the relevant experiences of human experts and making this expertise available as a resource to less experienced persons. Second, they offer explanations to users, thus serving dual roles as justification and training devices. Four important potential outcomes of ESS usage (Liang,
1988; McKee, 1986; Shim and Rice, 1988) include (1) improved and more effective
decision making; (2) more efficient decision making; (3) higher frequency of making
correct decisions; and (4) increased job insight through the learning stimulated by the
system. On the other hand, the cost of developing and using these systems may be high.
Such cost includes that of hiring knowledge engineers, purchasing hardware and software,
taking regular and productive time away from the experts, training, undergoing
organizational changes and operational disruptions, and in some cases, hiring facilitators
and chauffeurs to operate the system which may be necessary in the group decision making
context.

As most of the ESS that have been developed are used to support individual decision
making, empirical studies to date have examined the use of ESS almost exclusively in the
context of individual decision making (Dhaliwal, 1993; Eining and Dorr, 1991; Gregor,
Oz, Fedorowicz, and Stapleton, 1993; Peterson, 1988; Ye and Johnson, 1995). Other than
the field study by Sviokla (1986; also reported in 1989 and 1990), no empirical study has
yet investigated the use of ESS to support multiple-individual decision making.
Considering that group decision making is becoming more popular and information
technology (IT) is increasingly used to support decision making in organizations, it is
important for us to understand the effects of the use of expert support technology on group
decision making processes and outcomes (Swann, 1988). This research, therefore, aims to
evaluate the appropriateness of using an expert support system and its explanation facilities
to support group decision making.

Small groups are essential units of most organizations. They are frequently formed to
specialize in performing specific tasks. They serve as project teams, committees, and
decision-making bodies for a wide range of tasks and organizational functions. In fact,
when decision makers face a genuinely important task, it is likely that a group will be assigned to the problem. Sometimes the reason is simply that one individual alone cannot be expected to handle the complexity of the task (e.g., setting the strategic direction of a company or approving a financial loan of several million dollars; both of which require a diversity of knowledge, expertise, and skills). Other times, it is because decision makers assume that the added human resources available in a group will lead to a higher quality decision — or will at least lessen the chances of making a wrong decision. Group decision making may also occur for reasons other than improving the quality of the decisions made. For instance, it may be used to enhance consensus and legitimacy of decisions, and to increase commitment to decisions. As such, it is important to employ decision support tools to increase rationality, creativity, and participation in problem-solving meetings.

Group interaction and performance are greatly influenced by a large number of factors, such as the type and difficulty of the task a group performs and the type of decision support provided to the group. The use of information technology has long been recognized as a means to support and facilitate group work. Computer-based decision aids, such as decision support systems (DSS), group decision support systems (GDSS), and expert support systems (ESS), have been designed and developed to mitigate the cognitive limitations of human decision makers. DSS are interactive computer based systems that help decision makers confront ill-structured problems through direct interaction with data and analysis models (Sprague, 1980; Sprague and Carlson, 1982). However, in order to maximize or increase the acceptance and quality of decisions, extensive consultation and discussion are usually carried out (Sprague and Watson, 1996). The concept of GDSS was therefore developed to overcome the drawbacks and limitations in group decision making (see Chapter 2).
1.1.2 Group Decision Support Systems

GDSS combine communication, computing and decision support technologies to facilitate formulation and solution of unstructured problems by a group of people (DeSanctis and Gallupe, 1987). DeSanctis and Gallupe (1987) defined three levels of GDSS. Level 1 GDSS provide technical features aimed at removing common communication barriers, such as large screens for instantaneous display of ideas, voting solicitation and compilation, anonymous input of ideas and preferences, and electronic message exchange among members. In other words, a level 1 GDSS is a communication medium only. Level 2 GDSS provide decision modeling or group decision techniques aimed at reducing uncertainty and “noise” that occur in the group’s decision process. These techniques include automated planning tools (e.g., PERT, CPM, Gantt), structured decision aids for the group process (e.g., automation of Delphi, nominal, or other idea-gathering and compilation techniques), and decision analytic aids for the task (e.g., statistical methods, social judgment models, expert system support). Level 3 GDSS are characterized by machine-induced group communication patterns and can include expert advice in the selecting and arranging of rules to be applied during a meeting. To date, very little research has been done in Level 3 GDSS because of the difficulty involved in automating the process of group decision making.

DSS is incorporated into GDSS at level 2. The meta-analysis by Benbasat and Lim (1993) indicates that level 2 GDSS generally leads to greater improvement in performance, satisfaction, and consensus than level 1 GDSS. The other two meta-analyses on GDSS (Dennis, Haley, and Vandenberg, 1996; McLeod, 1992) did not analyze the effect of level 2 GDSS independently of level 1 GDSS. The meta-analysis by Benbasat and Lim (1993) suggests that task support is more effective than process support in group decision making. The cognitive feedback (CFB) literature also points to task information rather than cognitive information as the aspect of CFB that influences performance (Doherty and
Balzer, 1988). Task information refers to relationships between cues and criterion events in the decision environment while cognitive information refers to relationships perceived by the decision maker about cues and criterion events (Balzer, Doherty, and O'Connor, 1989). Despite evidence pointing to the effectiveness of task support, no experimental research has evaluated the usefulness of expert support for group decision making. Interestingly, although expert support is a component of level 2 GDSS, none of the existing GDSS research has incorporated and evaluated the usefulness of expert support as a group decision support tool (Benbasat, DeSanctis, and Nault, 1993). The only work that has examined expert support in a group decision making context was carried out by Sviokla (1989) through the use of a qualitative case study in an organization that specializes in financial planning (refer to Chapter 2 for a review of his work). However, the study focused on the impact of the technology on the organization, rather than group. Therefore, the effectiveness of using expert support for group decision making remains to be tested.

1.2 Research Objectives and Motivation

The focus of this research is to examine if group decision making capabilities can be improved through the use of expert support technology and its explanation facilities. More specifically, it investigates

(1) the provision of expert analyses and advice, as well as the provision of the explanation facilities, increases knowledge transfer from the ESS to the decision makers, helps groups reach a higher level of consensus in their decisions, and results in a higher level of satisfaction with the group process and the group judgments,

(2) the effects of the use of expert support and its explanation facilities on group decision making processes,

(3) the effects of providing explanation facilities on perceived usefulness of and trust in ESS,
differences in the use of expert support and its explanation facilities by expert versus novice user groups, and the effects of such differences on group judgments, consensus, and satisfaction,

expert-novice differences in their perceived usefulness of ESS and trust in ESS.

A number of studies have indicated that expert support and its explanation facilities improved decisions made by individuals (Dhaliwal, 1993; Eining and Dorr, 1991; Lamberti and Wallace, 1990; Mao, 1995; Oz, Fedorowicz, and Stapleton, 1993; Peterson, 1988; Ye and Johnson, 1995). The additional knowledge provided by the ESS contributes to the improvement. In a typical group decision making context (where IT support is not available), the sharing of knowledge and information among the members as well as the multiple perspectives and approaches taken into account by the group typically lead to improvement in the decisions made (Shaw, 1981). In this case, we are interested in investigating whether or not providing expert support to decision making groups would lead to further improvement in these decisions. Although the GDSS literature, which has been mostly concerned with communication support, has indicated that group decision support typically leads to better group decisions made, the lack of focus of GDSS research in evaluating decision modeling support has prompted us to specifically investigate the effectiveness of providing expert support to decision making groups. This is important research, especially since recent evidence in GDSS research indicates that single-user tools are perceived to be more useful than group tools in supporting group work (Satzinger and Olfman, 1995).

This research also studies the usefulness of the explanation facilities in supporting group decision making. Although a number of studies (Dhaliwal, 1993; Eining and Dorr, 1991; Gregor, 1996; Mao, 1995; Murthy, 1990; Ye and Johnson, 1995) have specifically examined the usefulness and impact of the explanation facilities in supporting individual
decision making, no empirical evidence yet exists on their usefulness and impact in
supporting group decision making. As considerable effort can be expended in the design
and development of the explanation facilities, it is important to evaluate the usefulness of
the explanation facilities for supporting group decision making.

1.3 Significance of Research

As society is moving into a post-industrial or information age characterized by complexity,
diversity, and turbulence, effective exploitation of corporate knowledge and experience is
likely to be a critical determinant of an organization’s prosperity and even survival (Swann,
1988). Competitive pressures are leading to a realignment of the factors of competition
which in turn implies a need for structural change within organizations. This is likely to
lead to group decision making becoming more prevalent, and this throws greater emphasis
on the need to improve the quality of group decision making processes and outcomes.

As pointed out by Huber (1990), it is important that appropriate knowledge on decision
support be acquired by studying the effects advanced information technologies (such as
ESS and GDSS) have on decision processes and outcomes. With an increased
understanding of the use of these technologies, better and more appropriate features can be
designed and introduced into the systems to help improve decision processes and
outcomes. However, only a relatively small number of empirical studies have examined the
use of decision aiding (i.e., task support) techniques in contrast to communication support
for group decision making (Lim and Benbasat, 1993). Most MIS research assumes that
communication technology is a necessary feature for supporting group decision making
(DeSanctis and Gallupe, 1987). However, it has been noted that, for decision making tasks,
groups prefer to deal with one another face-to-face rather than through some
communication technology (Dase, Tung, and Turban, 1995; Siegel, Dubrovsky, Kiesler,
and McGuire, 1986; Watson, DeSanctis, and Scott Poole, 1988), especially with small
groups. Despite such preferences, a common face-to-face decision support scenario — where a group of decision makers, supported by a single support system (such as an ES or a DSS or a combination of both technologies), gathers together to make a decision — has been largely excluded from the stream of decision support research. This research attempts to fill this gap. Although the group of financial planners at the Financial Collaborative (TFC) has used an expert system called PlanPower to help them perform financial planning for their clients (Sviokla, 1989), the use of ESS to support group decision making is, in general, a relatively new idea that has not been explored much in either research or practice (Benbasat, DeSanctis, and Nault, 1993). To the best of our knowledge, this is the first experimental research to examine the effects of providing expert support and its explanation facilities to assist in group decision making.

1.4 Conduct of Research
This research utilizes an experimental method to examine the effects of using a financial analysis ESS and its explanation facilities to support group decision making among novice versus expert financial analysts. Multiple methods of measurement were employed to capture various aspects of the study. Expert consensus on the decisions was used to evaluate decision outcomes; questionnaire instruments were used to assess user perceptions; and process tracing was used to analyze and explain the impact of ESS and its explanation facilities on group decision making. Both quantitative and qualitative data analysis methods were used to analyze the results of the empirical investigation. The quantitative analysis comprises comparisons of performance and various system- and group-related perception measures of novices across the experimental treatments, as well as experts' versus novices' perceptions and performance in the use of an ESS and its explanation facilities. The qualitative analysis covers an analysis of the group decision making protocol to explain and justify the results obtained from the quantitative analysis.
1.5 Summary of Chapter 1
The effects of using an ESS and its explanation facilities have been examined in the context of individual decision making; however, no parallel effort exists on evaluating ESS impact on group decision making. The closest work was carried out by Sviokla (1986, 1989, 1990) which examined the organizational impact of ESS. The meta-analysis by Benbasat and Lim (1993) indicates that empirical work in group support usually deal with the collective and group communication phenomena. Cognitive and information processing aspects of group decision making have been largely neglected. Theoretical and empirical work on using the ES technology to support group decision making is also lacking. This research supplements the current stream of GDSS research by focusing on supporting the task or information processing aspects of judgment and decision making processes. It examines the effects of providing small face-to-face decision making groups with a single ESS and its explanation facilities, as well as expert-novice differences.

1.6 Organization of Dissertation
This dissertation is organized into nine chapters. Chapter 2 reviews the literature and empirical studies in two main areas: expert support and group support technologies. Chapter 3 covers the theoretical foundations of this research. It reviews relevant theories in the literature on group decision making, persuasion, and expert-novice differences. It also discusses concepts on the lens model which provides the conceptual framework for this research. Chapter 4 presents the research framework and the research design, and derives the hypotheses for addressing the research questions. Chapter 5 describes the research methodology, including the subjects' characteristics, the research task, the experimental procedures, and the dependent measures. Both qualitative and quantitative data were collected in this research. Chapters 6 and 7 present the quantitative analysis of the results whereas the qualitative analysis is discussed in Chapter 8. Chapter 9 concludes this dissertation with a summary of the research findings and their implications, the
contributions of this research, a discussion about the limitations of this research and the problems and challenges faced in carrying out this research, and future research directions.
CHAPTER 2: LITERATURE REVIEW

This chapter reviews previous research and literature on expert system technology and group support technology. Section 2.1 reviews the literature and previous empirical studies on expert system technology and its explanation facilities. Section 2.2 surveys the Group Decision Support Systems (GDSS) literature, its relationship with this research, and some related empirical studies.

2.1 Expert System Technology

The terms, expert support systems, expert systems, and knowledge-based systems, have been used in the literature to refer to systems that are developed using the expert systems technology. This section discusses both the similarities and differences between these terms, explains the components of these systems and their explanation facilities, and reviews related empirical work in this area.


Systems that are developed using the expert system technology could provide many potential benefits including cost reduction; increased output; improved quality; consistency of employee output; reduced downtime; captured scarce expertise; flexibility in providing services; easier operation of equipment; increased reliability; faster response; ability to work with incomplete and uncertain information; improved training; increased ability to solve complex problems; and better use of expert time (Fried, 1987; Stylianou, Madey, and Smith, 1992).

Organizations routinely used these systems to enhance the productivity and skill of human knowledge workers across a spectrum of business and professional domains (Durkin, 1994; Feigenbaum, McCorduck, and Nii, 1988). They are computer programs capable of
performing specialized tasks based on an understanding of how human experts perform the same tasks. They typically operate in narrowly defined task domains. Despite the name "expert systems", few of these systems are targeted at replacing their human counterparts; *most of them are designed to function as assistants or advisers to human decision makers* (Leonard-Barton and Sviokla, 1988; Luconi, Malone, and Scott Morton, 1986). Indeed, the most successful expert systems — those that actually address mission-critical business problems — are not "experts" so much as "advisors" (LaPlante, 1990). These systems eliminate the tedious, time-consuming, routine tasks that take up much of the employees' time, allowing them to concentrate on more challenging tasks that really do need a human's judgment.

A knowledge-based system (KBS) is organized in such a way that the knowledge about the problem domain is separated from the general problem solving knowledge (Waterman, 1985). The collection of domain knowledge is called the *knowledge base*, while the general problem-solving knowledge is called the *inference engine*. Both expert systems (ES) and expert support systems (ESS) are KBS, while the converse is not necessarily true (Waterman, 1985). For instance, an AI program to play tic-tac-toe would not be considered an expert system, even if the domain knowledge was separated from the rest of the program (Waterman, 1985). This is because the level of expertise required to play tic-tac-toe is too low for the program to be called an expert system. However, in practice, the terms, KBS and ES, are often used interchangeably.

ES may be used to *replace or support* decision making. On the other hand, the term, ESS, is more *specific* as it refers to *ES that are designed to provide assistance or advice to decision makers*. In other words, ESS do not replace decision making; they *support* decision making (Luconi, Malone, and Scott Morton, 1986). *ESS is the subject of interest in this research.*
2.1.2 **Components of ESS/ES**

Although ESS and ES may differ in the ways they are being used, their technical components are similar. As mentioned earlier, the two main components of an ESS/ES are:

1. the *knowledge-base*, in which the domain-specific knowledge is stored in the form of facts and rules,

2. the *inference procedure* ("inference engine") which operates on the knowledge-base, performs logical inferences and deduces new knowledge by applying rules to facts until the posed problem is solved.

Some of these systems also provide explanation facilities (Kriz, 1987). According to the pioneers of the expert system technology, Shortliffe (1976) and Buchanan (1986; Buchanan and Shortliffe, 1984), it is important for such systems to provide reasonable explanations, as well as good advice, for them to be acceptable to users. The next section describes the explanation facilities.

2.1.3 **Explanation Facilities**

The explanation facilities are an important component of the expert system technology (Southwick, 1991). In addition to analyses and advice, an ESS should also provide explanations of its behavior (on request) to be considered *usable* and *acceptable* to the user (Chandrasekaran, Tanner, Josephson, 1988). One reason often heard for favoring ESS/ES/KBS over more conventional programs is that they can explain and *clarify* their decision making (Gilbert, 1989). A system needs to explain what it has done, to assure users that its reasoning is logical and the conclusions sound. Good explanations may also *persuade* users that the system's conclusion is appropriate and relevant. If a system produces unexpected analyses or advice, a good explanation may *convince* the user of its
relevance. The ability to justify and explain the system’s advice to the user is important for a number of reasons (Chandrasekaran, Tanner, and Josephson, 1988):

1. The user may want to know if the system took into account all the knowledge that the user considers relevant,

2. The user may want to know if the strategies adopted by the system for solving the problem are satisfactory,

3. The user may wish to know if all the relevant data describing the problem state are being considered.

The explanation facilities could also be used as a training aid, by describing the system's domain knowledge and inference techniques when the objective is to teach or train users, i.e., to transfer the knowledge in the system to naive users. Finally, explanations are often useful as a debugging aid for systems designers.

While prior work has concentrated on the technology of explanation generation and presentation (e.g., Abu-Hakima and Oppacher, 1990; Chandrasekaran, Tanner, and Josephson, 1989; Lamberti and Wallace, 1995; Moffitt, 1989; Neches, Swartout, and Moore, 1985; Scott, Clancey, Davis, and Shortliffe, 1977, 1984; Swartout, 1983; Weiner, 1980), little is known about the behavioral impact of such technology on users, especially in the group context. To be considered useful and acceptable, an ESS must be able to present users with explanations of its knowledge of the task domain and the reasoning processes it employed to solve problems and make recommendations (Buchanan and Shortliffe, 1984; Ye and Johnson, 1995). At the root of every human being's understanding is the ability to seek and create explanations (Schank, 1982). For instance, the ability to explain decisions derived by a medical consultation expert system was judged to be the single most important requirement by a group of one hundred and fifty physicians, as no
user was willing to accept the system’s conclusions unless it could describe how they were derived (Teach and Shortliffe, 1981).

Decision makers in practice must deal with the real-world consequences and risks associated with their decisions. Because users of an automated decision aid are responsible for the decisions made by the systems, they are unlikely to accept decisions based on reasoning that they are unaware of or do not understand (Hollnagel, 1987). Providing appropriate explanations can increase user understanding of and confidence in machine-generated decisions.

2.1.3.1 Evolution of Explanation Facilities

Classical Explanation — Reasoning Trace Explanations

The classical explanation facilities provide explanations by generating a trace of the symbolic reasoning of the system (Shortliffe, 1976). These are termed reasoning trace explanations by Southwick (1991). The origin of reasoning trace explanations was MYCIN (Shortliffe, 1976) which provided users with the opportunity to ask “How?” and “Why?”, which are to be interpreted to mean “explain how you reached this conclusion” or “explain why you are asking me that”. A trace of the rules used during the deductive process was kept and used to explain the system’s actions.

However, there are major shortcomings with the “MYCIN paradigm” (Southwick, 1991). The major failing was described by Clancey (1983) when he attempted to use the MYCIN rule base in a tutoring system. He began with the reasonable supposition that a knowledge base used for deduction could also be used for teaching, but soon discovered that this was not the case. First, an execution trace does not provide sufficient information, because there is no justification for why a conclusion logically follows from a premise as there is no encoding of how the concepts in a rule fit together -- the rule consists of preconditions and
a conclusion; the reason for the conclusion following from the precondition has been compiled out. Second, although justifications for the inclusion of a rule were used in the design of the system, this knowledge is unlikely to appear anywhere in the rule base. In other words, the system consists of a set of rules representing the knowledge of an expert in a compiled form, with all the deep associational links removed. The problem solving approach taken by the system, therefore, cannot be fully articulated because the structure of the search space and the strategy for traversing it are implicit in the ordering of the rule concepts. In this sense the rule base is "flat", and so cannot provide a justification for the system's actions.

Helman and Bennett (1988) and Southwick (1991) stress that simply describing the behavior of a consultation process (using an execution trace) is not sufficient. It is important that a system must produce reasons for a conclusion, action, or state of affairs. The missing information is often referred to as a "deep mode" of the domain or deep explanation.

Deep Explanations

Deep explanation explicitly represents relations that are only implicitly represented in a compiled knowledge base. The task-specific goals and problem-solving knowledge of such systems are compiled from more-general domain knowledge than the knowledge in the compiled knowledge base. If the system remembers a trace of the compilation, it can justify system rules in terms of deeper knowledge. For example, a financial modeling system may give deep explanations based on economic theory. An explanation system that has access to a deep model can provide explanations that are intuitively more satisfying, since they relate to the deeper concepts that underlie the domain model (Southwick, 1991). Wick and Slagle (1989) claim that the lack of this explicit knowledge in practice systems
limits its use. Chandrasekaran and Mittal (1983) argue that “deep” knowledge can be used by the systems to provide higher quality explanations.

Since the concept of deep knowledge in reasoning systems was introduced, there has been much interest in its development (Chandrasekaran and Mittal, 1983; Swartout, 1983). Examples of its development include Ergo (King, 1986), GUIDON (Clancey, 1983), IDM (Fink, Lusth, and Duran, 1985) and XPLAIN (Swartout, 1983). Chandrasekaran and Mittal (1983) also found that diagnostic problem solving (and explaining such problem solving) may require different levels of knowledge. Such multi-level knowledge was implemented in FINALYZER (Mao, 1995) using the hypertext concept (see Chapter 5 for a description of FINALYZER).

Strategic Explanations

The systems designer uses a particular problem solving strategy when constructing a knowledge base. The order in which the rules are written, for example, affects the behavior of a rule-based system, and very often system behavior is controlled through rule or task ordering. An explicit representation of the problem-solving process is required to explain the strategy employed in solving a problem. Such a representation is encoded in a body of meta-knowledge, and is structured into tasks consisting of meta-level goals and subgoals, and meta-rules. These goals and rules are the methods used for performing tasks. Such additional knowledge allows the system to explain its strategy during a consultation. For example, NEOMYCIN is an extension of MYCIN to include strategic explanations (Hasling, Clancey, and Rennels, 1984). That is, NEOMYCIN explicitly outlines system strategies in its knowledge base, which renders them available for explanation. Thus, NEOMYCIN gives explanations of the overall problem-solving strategy.
2.1.3.2 Explanation Types

The three main approaches to explanations that have been discussed so far are (Southwick, 1991):

(1) **Reasoning trace explanations**: these are explanations at the system level, able to give information about the contents and structure of the knowledge base. They explain why a conclusion was reached, or a decision made, by describing the reasoning steps that led to the conclusion.

(2) **Deep explanations**: deep, or model-based explanations justify system results by linking them to a deep, causal model. Thus deep explanations attempt to give the underlying reasons for an action or state.

(3) **Strategic explanations**: rather than explaining a result by listing rules used, these explanations describe the strategy employed by the problem solver. Strategic explanations give the user an insight into the problem-solving methodology.

While reasoning trace explanations rely only on the formulation of knowledge that comprises the knowledge base, the other two explanation types require additional supplementary knowledge. A system that can give strategic explanations must be able to reason about its own activity, which may require knowledge about the ordering of problem solving tasks, for example. Deep explanations obviously require a great deal of information in the form of a causal model of the domain. Several systems have been developed using the above concepts. For instance, XPLAIN (Swartout, 1983) uses deep knowledge ("the domain model") and a representation of problem-solving control strategies ("domain principles") to compile a knowledge-based system. Thus, the system can examine control strategy to analyze system behavior and can use the deep model to justify system rules.

To see the roles that these explanation types play, a diagnostic system for car maintenance was used by Southwick (1991) to illustrate these explanation types: Suppose that the
system is explaining its conclusion that a clogged fuel filter caused an engine to die. Such a system might give the following explanations, corresponding to the three categories:

(1) **Reasoning trace**: “You told me that the engine spluttered, and I know that if the engine coughs, then the filter may be at fault.”

(2) **Deep model**: “A clogged fuel filter prevents petrol from reaching the carburettor, thus causing engine failure.”

(3) **Strategic**: “There are three engine subsystems to check. I checked the fuel system first, because the symptoms indicated the likelihood of a fuel system problem.”

Reasoning trace explanations are referred to as *shallow knowledge* by Hollnagel (1987), while deep explanations concern *deep knowledge*. Shallow knowledge (and reasoning) refers to the phenomenological level while deep knowledge (and reasoning) refers to the conceptual or morphological level of system description. Shallow knowledge is about those input-output relations that are perceived by the user whereas deep knowledge is about the factual input-output relations (Hollnagel, 1987). The difference is also one of the degree of elaboration of knowledge—shallow knowledge being less elaborated than deep knowledge.

Hollnagel (1987) also identifies the limitations of deep and shallow knowledge for diagnosis and argues for the combination of both as the solution. Deep knowledge is difficult to apply for abnormal situations (accidents, etc.) because it describes the mechanisms for normal system behavior in a closed world. For normal situations, it is inefficient to use deep knowledge because the number of possible paths and side-connections would quickly become astronomical. In short, it is both inefficient and insufficient to use only deep knowledge for decision making. On the other hand, shallow knowledge, by virtue of being derived from experience, will be incomplete. Furthermore, if a system is unable to explain the reasons for its advice or recommendations to the user, it is of very limited use. Consequently, taken by themselves both shallow and deep knowledge
are insufficient as basis for a diagnosis (Hollnagel, 1987). Thus, the most obvious solution is to combine shallow and deep knowledge, using the shallow knowledge as a way of controlling the deep knowledge.

The classification of explanations into reasoning trace, deep, and strategic types is consistent with the categorization proposed by Chandrasekaran, Tanner, and Josephson, (1988, 1989). The three generic categories of explanations proposed by Chandrasekaran, Tanner, and Josephson (1988, 1989) are:

1. **Trace (Type 1):** explaining why certain decisions were or were not made; showing what pieces of knowledge have been invoked by the expert system to produce a solution. (A piece of knowledge is typically structured as an association between specific data and a conclusion/hypothesis.)

2. **Justification (Type 2):** explaining knowledge base elements; concerning explicit representation of the causal argument underlying individual pieces of knowledge.

3. **Control (Type 3):** explaining the control behavior and problem-solving strategy; describing the partial goal structure of the problem-solving task, typically represented as a plan according to which individual pieces of knowledge are invoked.

Table 2-1 shows the correspondence between Southwick’s (1991) and Chandrasekaran, Tanner, and Josephson’s (1988, 1989) categorizations of explanation types.
Another classification of explanations into feedforward and feedback was proposed by Dhaliwal and Benbasat (1996). Feedforward and feedback explanations were designed based on two learning operators from the cognitive learning perspective, cognitive feedback and feedforward (Bjorkman, 1972). The cognitive feedforward and feedback paradigm, which is applied in the context of problem solving, emphasizes a particular order among events (Bjorkman, 1972). Feedforward knowledge was presented as cognitive feedforward prior to analysis, and feedback explanations were accessible as cognitive feedback after the system has presented its analyses and advice. The advantages of providing domain knowledge as cognitive feedforward include promoting more accurate and consistent knowledge acquisition, relieving the learner from certain cognitive strain, and favoring an analytical rather than intuitive mode of thought (Bjorkman, 1972).

Cognitive feedback constitutes case-specific information provided to users at the end of an analysis, while feedforward constitutes non case-specific, generalized information pertaining to the input cues of an analysis that is provided to users prior to the performance of an analysis.

There are three major distinctions between feedback and feedforward (Dhaliwal and Benbasat, 1996):

1. **Case Specificity.** Cognitive feedback provides information that clarifies the results of the analyses. It uses the results of the analyses as the starting reference point for

<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Reasoning trace</td>
<td>Trace (Type 1)</td>
</tr>
<tr>
<td>Deep</td>
<td>Justification (Type 2)</td>
</tr>
<tr>
<td>Strategic</td>
<td>Control (Type 3)</td>
</tr>
</tbody>
</table>

*Table 2-1: Correspondence between Southwick’s and Chandrasekaran et al. ’s Categorizations of Explanation Types*
improving the decision maker’s understanding of the task. Feedforward, on the other hand, is not related to the results of analyses of the specific case being considered but focuses rather on the input cues of the task.

(2) Temporal Order. Feedforward is always provided prior to task performance, while feedback is presented subsequent to completion of the task and provision of the results of analyses.

(3) Types of Cues Focused Upon. Feedforward relates to the cues which serve as input variables, while feedback is information relating to the results of analyses. Cognitive feedback uses the clarification of case-specific outcomes as a starting point and provides information that traces the reasoning backward to the input cues. Feedforward focuses on the clarification of the input information and traces the reasoning forward to results of analyses.

In this research, the feedforward and feedback classification of explanations is adopted. The relationship between the two classifications—1) feedforward and feedback, and 2) deep, reasoning trace, and strategic—is as follows: Basically, deep explanations are feedforward explanations as they are not case-specific and are available prior to task performance. Reasoning trace explanations are feedback explanations because they are case-specific and are only relevant subsequent to analysis or task execution. Strategic explanations are available under both feedforward or feedback. Feedforward strategic explanations clarify the overall manner in which input information to be used is organized or structured, and specify the manner in which each input cue to be used fits into the overall plan of assessment that is to be performed. Feedback strategic explanations clarify the overall goal structure used by the system to reach a particular conclusion, and specify the manner in which each particular assessment leading to the conclusion fits into the overall plan of assessments that were performed.
Given that the explanation types vary in their structures and information content, user requirements for each type of explanation are expected to be contingent upon 1) user characteristics, such as prior knowledge and experience of the task domain (expertise) and prior experience with similar systems, and 2) task characteristics, such as the environmental context of the task, and the type of knowledge and inference processes required to accomplish the task (Ye, 1990).

2.1.4 Empirical Studies on Expert System Technology

A number of empirical studies have evaluated various aspects of expert system technology (Berry and Broadbent, 1984, 1987; Dhaliwal, 1993; Eining and Dorr, 1991; Gregor, 1996; Hsu, 1993; Lamberti and Wallace, 1990; Mao, 1995; Moffitt, 1989; Murthy, 1990; Oz, Fedorowicz, and Stapleton, 1993; Peterson, 1988; Sviokla, 1986, 1989, 1990; Ye, 1990; Ye and Johnson, 1995). A review of these studies indicates that, with the exception of the work by Sviokla (1989), no empirical studies have examined the use of expert system technology and its explanation facilities in a multiple-individual decision making context; the others examined its use by single individuals. This dissertation therefore represents one of the pioneering effort to investigate the effects of the use of expert system technology and its explanation facilities on group decision making processes and outcomes.

2.1.4.1 Review of Empirical Studies on Expert system technology

Berry and Broadbent (1984) show that practice improved procedural ability but did not improve the ability to answer related questions (declarative knowledge). Verbal instructions given before the tasks significantly improved ability to answer questions but had no effect on procedural performance. The insight gained from this research is that to both perform and explain a task, both practice and verbal instructions are important. In another study, Berry and Broadbent (1987) compare two forms of explanation on a complex search task. Subjects who were allowed to ask "why" each computer
recommendation was made performed significantly better than those who were provided with a block text of explanation at the start of each trial. It is concluded that success in learning depends upon the amount, level of specificity, and timing of explanations.

Dhaliwal (1993) find that explanations provided by an ESS increase both the accuracy of individual judgmental decision-making and the user perceptions of system usefulness, although different types of explanations are responsible for each. Two types of explanations, feedforward and feedback, were studied. Feedback explanations constitute *case-specific* information provided to users at the end of an analysis. On the other hand, feedback explanations constitute *non case-specific*, generalized information pertaining to the input cues of an analysis provided to users prior to the performance of an analysis. Feedback explanations were found to improve the accuracy of judgmental decision making but had no effect on user perceptions of usefulness. Feedforward explanations were found to increase user perceptions of usefulness but had no effect on the accuracy of judgmental decision making. The use of the Why explanation as feedback improved the accuracy of judgmental decision making. This finding is consistent with the finding by Berry and Broadbent (1987). Other findings in the study include: 1) user expertise is not a determinant of the proportion of explanations used, but does influence the types of explanations that are used – novices used significantly more Why explanations than the Strategic explanations, while experts used significantly more How explanations as compared to the Strategic explanations; 2) the Why and How explanations are used significantly more than the Strategic explanations.

Eining and Dorr (1991) study the experiential learning of novice auditors using an ESS as a decision aid, both with and without explanatory capability. The study took place over five one-hour sessions during a five-week period. Participants using the expert system (with explanatory capability and without explanatory capability) performed better with respect to
both time and accuracy than did participants in the group not provided with decision aid and in the group using only a questionnaire as an aid (i.e., conventional decision aid). No difference in performance resulted from the use of the expert system with explanatory capability versus one without this capability, which is inconsistent with the findings by Berry and Broadbent (1987) and Dhaliwal (1993). However, there was no measure as to whether the explanation facilities were used when available.

Gregor (1996) investigates three determinants of the frequency of explanation use in the accounting context: 1) the level of user expertise, 2) the goal of the user (learning versus problem solving), and 3) the nature of the problem solving situation (collaborative versus non-collaborative). A higher level of expertise of the user was related to higher problem solving performance, greater use of explanations, and greater confidence in the system. Both a goal of learning rather than problem solving and a requirement for collaborative problem solving led to higher frequency of explanation use. There was no significant difference in performance between groups with and without explanations. However, there was support for a positive relationship between frequency of use of explanations and help, and problem solving performance. This relationship was observed only in groups where explanation facilities were made available.

Hsu (1993) investigates the effects of using a financial statements analysis expert system on knowledge transfer. The study indicated that both cognitive styles and interface designs were important factors that influenced knowledge transfer. Field-independents were affected more by different interface designs than were field-dependents. In addition, justification explanations resulted in a greater degree of knowledge transferred than using rule-trace explanations alone. The research concludes that it is important to consider individual differences, explanation presentation formats, and multiple trials in knowledge transfer.
Lamberti and Wallace (1995) evaluate intelligent interface requirements for knowledge presentation in an expert system used for diagnostic problem solving. They examined interactions between user expertise, knowledge presentation format (procedural versus declarative), question type (requiring abstract versus concrete answers), and task uncertainty, in terms of speed and accuracy of decision making performance. The expert system has a greater impact on improving performance for low-skill users than for high-skill users. A relationship was found between skill level and task uncertainty indicating that different skill-level users require different presentation formats paralleling their conceptual representations of the problem. For higher uncertainty tasks and high-skill users, response time and accuracy improved when explanations were declarative rather than procedural. For low uncertainty tasks, low-skill users performed better than high-skill users when given declarative explanations. Both high- and low-skill users with low uncertainty tasks were more confident with procedural explanations. In relation to concrete versus abstract knowledge organization, low-skill users performed faster and more accurately when answering questions requiring concrete knowledge organization but high-skill users performed faster, although not necessarily more accurately, when responding to questions requiring abstract knowledge organization.

Mao (1995) investigates the behavioral and cognitive basis of the use of hypertext in providing explanations in an ESS. The use of hypertext for providing explanations significantly improved decision accuracy, and influenced users’ preference for explanation types, and the number and context of explanation requests. Enhanced accessibility to deep explanations via the use of hypertext significantly increased the number of deep explanations requested by both novices and experts. Increased use of deep explanations led to higher knowledge transfer from the system to the users, especially the novices. Verbal protocol analysis shows that the lack of knowledge and the means of accessing deep
explanations could make it difficult to understand system recommendations, and that deep explanations could improve the understandability of system advice, especially in cases where unfamiliar domain concepts were involved. Experts and novices also had different preferences for explanation types. Experts requested a much higher percentage of How, and lower percentages of Why and Strategic explanations, than did novices. Verbal protocol analysis illustrates that experts and novices used explanations for different purposes, the novices mainly for learning and the experts mainly for verifying their knowledge against that of the system.

Moffitt (1989) assesses incidental learning of declarative knowledge and decision processes in which four forms of explanation provision (no formal explanation, user-invoked rule-trace explanation, user-invoked canned-text explanation, embedded-text explanation) were manipulated. Rule-trace explanation facilities allow users to access currently active rules and a post-session trace of the rules which fired during the session when the facilities are invoked by users. Canned-text explanation facilities provide English-language explanations when invoked. Embedded-text explanation facilities are the same as canned-text explanation facilities except that the explanations were automatically provided to the user instead of user-invoked. It was found that users of the expert system did learn declarative knowledge and were able to apply the heuristics used by the system in an unaided decision context. Embedded-text explanation facilities provide a higher level of incidental learning than the other explanation facilities. Furthermore, text-based treatment groups (embedded, canned) rated the system significantly higher than the no formal explanation treatment group when evaluating the system on its usefulness as a learning device and on the amount of scheduling information they had learned while using the system. In addition, the embedded-text treatment group rated the system significantly higher than did the rule-based treatment group on these two measures. Nevertheless, the
users' ratings on the usefulness of the system as a decision aid were independent of the explanation provision techniques.

Murthy (1990) finds the performance of participants using an expert system without explanations to be better than that of participants using the expert system with explanations, *though the difference was not significant*. He concludes that the availability of explanation facilities is counterproductive. However, the experiment appears to have been conducted over a short period of time, which may not have given the subjects enough time for experiential learning to occur.

Oz, Fedorowicz, and Stapleton (1993) examine the improvement in decision making skills when an expert system is provided to novices to support risk evaluation. The users improved their decision quality more than non-users. However, no difference was found between the users and non-users with respect to decision making time. The users' confidence in the decisions increased over time more than non-users' confidence. Furthermore, no relationship was found between the users' attitude toward computers and the improvement of their decision quality and increased confidence in their decisions.

Peterson (1988) examines the usability and usefulness of an expert system in providing managers with their managerial knowledge skills such as giving performance feedback. He conducted a laboratory study with actual managers and found that the performance of both inexperienced and experienced managers improved with the use of the system. Inexperienced managers not only demonstrated greater improvement in accuracy from the use of the system but also found the system more useful than experienced managers. More interestingly, inexperienced managers outperformed experienced managers when the system was used.
Sviokla (1986, 1989, 1990) examined three field sites with expert systems in active use to generate knowledge about the effects of the use of expert systems on the organizations which use them. The study used a comparative, three-site, pre-post exploratory design to describe and compare the effects of expert systems use on three organizations: The Financial Collaborative (using the PlanPower system for financial planning), Digital (XCON for computer configurations) and Baroid (MUDMAN for drilling decisions). In all three cases, the expert systems seemed to increase the effectiveness and efficiency of the user firm at the expense of an increased rigidity in the task. Progressive structuring of the tasks was observed in all the three sites surveyed. As the systems were being used, maintained and improved upon, problem-solving knowledge improved and problem structure increased. They also helped in managing complexity, improving representation, improving standards, and providing a rigorous understanding of some of the previously uncertain parts of the task. Furthermore, the task process changed, resulting in shifts in roles and responsibilities.

Ye (1990; Ye and Johnson, 1995) investigates the impact of expert system explanations on changes in user beliefs toward the conclusions generated by the system. Three alternative types of explanations — trace, justification, and strategy — were provided in a simulated diagnostic expert system performing auditing tasks. The results indicate that the explanation facilities can make advice generated by the system more acceptable to users and that justification (Why explanation) is the most effective type of explanation to bring about changes in user attitudes toward the system. Although the Why explanation was the most preferred explanation across all levels of user expertise, experts perceived the How explanation as being most useful, and novices perceived the Why explanation to be most useful.
2.1.4.2 Summary of Empirical Studies on Expert system technology

Usefulness of ESS Analyses and Explanations

Users of ESS achieve a higher degree of decision accuracy than non-users (Eining and Dorr, 1991; Oz, Fedorowicz, and Stapleton, 1993). On the other hand, the explanation facilities increase user acceptance of ESS generated advice (Ye, 1990; Ye and Johnson, 1995). Feedforward explanations increase users' understanding of the tasks (Berry and Broadbent, 1984; Hsu, 1993) and perceptions of usefulness of the system (Dhaliwal, 1993), but they do not lead to better performance (Berry and Broadbent, 1987; Dhaliwal, 1993). Feedback explanations, however, are more effective than feedforward explanations in improving performance (Berry and Broadbent, 1987; Dhaliwal, 1993), but they have no effect on users' perceptions of usefulness of the system (Dhaliwal, 1993). There are some situations where feedback explanations do not lead to better performance (Eining and Dorr, 1991; Gregor, 1996; Murthy, 1990). In addition, evidence indicates a positive relationship between frequency of use of explanations and problem solving performance (Gregor, 1996).

Factors Influencing Use of Explanations

A requirement for collaborative problem solving and a goal of learning rather than problem solving lead to higher frequency of explanation use (Gregor, 1996). Using hypertext explanations also increases the frequency of access of explanations (Mao, 1995). In general, the Why and How explanations tend to be used more than the Strategic explanations (Dhaliwal, 1993).

User Expertise and Use of Explanations

A higher level of user expertise also relates to higher problem solving performance (Gregor, 1996; Lamberti and Wallace, 1995) and greater confidence in the system (Gregor, 1996). Expert systems, however, improve performance for less experienced users more than for experienced users (Lamberti and Wallace, 1995; Peterson, 1988). Users with less experience also find the system more useful than experienced users do. Most interestingly, inexperienced users outperformed experienced users with the system.

For highly uncertain tasks, high-skill users perform better in both response time and accuracy when explanations are declarative rather than procedural (Lamberti and Wallace, 1995). For low uncertainty tasks, low-skill users perform better than high-skill users when given declarative explanations. Both high- and low-skill users with low uncertainty tasks are more confident with procedural explanations.

Experts and novices also have different preferences for explanation types. Experts tend to use a much higher percentage of How and a lower percentage of Why than novices (Dhaliwal, 1993; Mao, 1995). In contrast, Ye (1990) finds Why explanations to be the most preferred explanation across all levels of user expertise. However, experts value the How explanation most while novices find the Why explanation to be most useful (Ye, 1990).

Use of ESS in Organizations
Sviokla (1986, 1989, 1990) observed six main changes in organizations that adopted expert systems: 1) Expert systems increase the effectiveness and efficiency of organizations. The information processing capacity of the organizations improves and the cycle time decreases. However, this is at the expense of an increased rigidity in the task. 2) The use of ES also increases the structure of the problem. As an ES was being used, maintained, and improved upon at the sites, problem-solving knowledge improved and problem structure increased. 3) Using an ES changes the task procedures. For example, in the PlanPower
case, the focus of the financial planners' regular Wednesday meetings shifted from planning design to a critique and study of the system. The task may also become more diverse and complicated in terms of its process flow. In some cases, tasks became differentiated into categories: those that were solved by ES; those that would be modified from the ES solution; and those that would still need to be solved by hand. 4) The change in task procedures resulted in shifts in roles and responsibilities. For instance, PlanPower took the place of some work roles such as report writer and as a result of Planpower, a new role was created to facilitate the use of the system. As another example, XCON changed the job characteristics of Digital's technical report editor and caused it to become less challenging after using XCON. 5) Using an ES makes task execution less flexible. The PlanPower and XCON cases show the tedious work required to make changes. To add new features to an ES, the help of the ES developers is needed, which increases the dependence upon the ES vendor. 6) An ES also improves work quality, increases the scope and formality of data used, and provides better data and decision support. For example, the simulation program in MUDMAN gives people more confidence in their decisions.

2.2 Relationship with GDSS Research

As defined earlier in Chapter 1, a GDSS is "a system that combines communication, computing, and decision support technologies to facilitate formulation and solution of unstructured problems by a group of people" (DeSanctis and Gallupe, 1987). The definitions of the three levels of support in GDSS proposed by DeSanctis and Gallupe (1987) were also discussed in Chapter 1. Level 1 GDSS is a communication medium only. Level 2 GDSS is the addition of decision modeling or group decision techniques to level 1 GDSS. These include tools or aids commonly found in individual decision support systems as well as group structuring techniques. This research addresses the effectiveness and usefulness of a decision modeling and analysis technique — the expert system technology and its explanation facilities — for supporting group decision making. It examines one
specific element of level 2 GDSS independently of level 1 GDSS by evaluating the impact of providing different levels of expert support (including no support) on group judgments. Communication support is not studied in this research because message exchange is rated by users to be the least important in supporting face-to-face meetings (Satzinger and Olfman, 1995). In contrast, traditional single-user tools such as spreadsheets/DSS, database retrieval, and presentation support are perceived to be more useful than group tools in supporting group work (Satzinger and Olfman, 1995). Satzinger and Olfman (1995) therefore conclude that group support packages would become more receptive to users if they integrated traditional single-user tools.

### 2.2.1 Studies on Group Support Using Decision Modeling and Analysis Techniques

To date, empirical research in GDSS has focused largely on communication and process support (Jessup and Valacich, 1993). Although researchers find such group support tools useful (e.g., Benbasat and Lim, 1993; Connolly, Jessup, and Valacich, 1990; Dennis, Haley, and Vandenberg, 1996; Gallupe, DeSanctis, and Dickson, 1988; Jarvenpaa, Rao, and Huber, 1988; Jessup, Connolly, and Galegher, 1990; Jessup and Tansik, 1991; McLeod, 1992; Watson, DeSanctis, and Poole, 1988), Satzinger and Olfman (1995) believe developers should integrate traditional single-user tools into face-to-face meeting systems as they are perceived by users to be more useful than group tools. Despite such a finding, little research has investigated the impact of using traditional single-user tools for group support. To more clearly understand the impact of ESS and its explanation facilities for group support, we evaluated their impact independently of other forms of (individual and group) support. For instance, research on GDSS has been criticized for its heterogeneity (McLeod, 1992), making it difficult to pinpoint exactly which feature or combination of features is responsible for a specific effect.
We have identified four empirical studies that have specifically examined the use of computer-based decision modeling and analysis techniques independently of communication support for group decision making (Joyner and Tunstall, 1970; McCartt and Rohrbaugh, 1989; Scott Morton, 1971; Sharda, Barr, and McDonnell, 1988).

Joyner and Tunstall (1970) test a program called CONCORD (CONference COoRDinator) to help groups of decision makers apply a satisficing model. Users of CONCORD performed no better than nonusers in solving human relations problems.

McCartt and Rohrbaugh (1989) assess the perceived effectiveness of group decision making process with decision conferencing using a self-administered survey of participants. Decision conferencing is a level 2 GDSS that employs a portable, single-user computer system to support groups of managers and executive teams working face-to-face on a wide variety of organizational problems. Verbal and nonverbal communication in decision conferences is not restricted by electronic networking but, rather, takes a completely connected, “each to all” pattern enhanced by the presence of a group facilitator. A distinguishing feature of decision conferencing is the on-the-spot development of a computer-based model that incorporates the differing perspectives of participants. The group can examine the implications of the decision model, modify it, and test the effects of different assumptions, thereby ruling out ineffective strategies and focusing quickly on primary issues of major impact. During a decision conference, the group is assisted by at least three people from outside the organization: a facilitator of the meeting, an analyst who records explicit information and participants’ assessments of expressed priorities on a computer, and a correspondent who maintains a record of the meeting. Of the 14 decision conferences surveyed, decision conferencing had been a very effective intervention for 5 organizations, judged consistently and positively both in terms of process and outcomes. For 5 other organizations, the intervention clearly had been problematic, evaluated with
sharply lower ratings on process and outcome items. Participants from the remaining 4 organizations provided responses that were intermediate. It is interesting and instructive to find that decision conferencing, which is only a single type of GDSS intervention, can produce such dramatically different profiles of decision process effectiveness. All the participants in decision conferences, however, clearly view the intervention (no matter how successful) as providing a fully participatory and goal-centered process. Differences in perceived conference success were related to: 1) the proportion of participants who believed the conference resulted in a decision, 2) the level of benefits derived from full support of the structure or preference technology, 3) the opportunity for full, extended discussion, 4) development of an action plan, and 5) expected resolution of the problem by the conference end. Thus, the greatest distinguishing factor for perceived conference success could be made with respect to the perceived usefulness of their collective work supported by the analyst's use of a personal computer, that is, the group's examination, refinement, and testing of the decision model projected on a large viewing screen.

In Scott Morton's (1971) study, a DSS was used in an organization to help the marketing manager, the production manager, and the market planning manager negotiate and come to an agreement for the manufacture, sales, and distribution of their laundry products. One of the drawbacks of the study was that there was no sure way to measure the quality of the decisions made with the support of the DSS. However, the amount of time spent by the manager actually working on the problem was sharply reduced by a factor of 12 to 1, which not only released managerial talent for other problems, but it also induced more vigorous, logical, problem solving. In addition to the time effect, there was a change in the problem-finding and problem-solving process. The impact of the DSS on the change in decision structure was analyzed based on Simon's (1958) "Intelligence, Design, Choice" stages. In addition, communication between the managers changed considerably with the use of the
system. Less effort was required to make one's point clear and less time was spent on discussing misunderstood issues.

Sharda, Barr, and McDonnell (1988) assess the effectiveness of a DSS for supporting a business simulation game. The DSS was developed using Interactive Financial Planning System (IFPS), a modeling language which allows for "what-if" analysis. They found that groups with access to the DSS made significantly more effective decisions in the business simulation game than their non-DSS counterparts. The DSS groups took more time to make their decisions than the non-DSS groups at the beginning of the experiment but the decision times converged in a later period. The DSS teams also exhibited a higher confidence level in their decisions than the non-DSS groups.

2.2.2 Meta-Analyses on GDSS Research

One implicit and secondary objective of this research is to find out if the effects of using ESS support for group decision making are congruent or similar with findings in the GDSS literature. The similarities and differences between this research and GDSS research has been discussed earlier. As such, this section summarizes the findings of the GDSS literature while Chapter 9 reports on the results of the comparison. To date, three meta-analyses on GDSS research have been published (Benbasat and Lim, 1993; Dennis, Haley, and Vandenberg, 1996; McLeod, 1992). The results of the meta-analysis by Dennis, Haley, and Vandenberg (1996) suggest that, in general, GDSS use improves decision quality, increases time to make decisions, and has no effect on participant satisfaction. Consistent with the findings from Benbasat and Lim (1993), they also found larger groups to be more satisfied and to perform better with the use of GDSS than smaller groups. Consensus was not investigated in the study. The findings from Benbasat and Lim's (1993) and McLeod's (1992) meta-analyses are consistent in that GDSS use increases decision quality, time to reach decisions, and equality of participation, but decreases consensus and satisfaction.
Benbasat and Lim (1993) also investigate the moderating effect of level 2 versus level 1 GDSS. The level of GDSS support was found to moderate the relationships between GDSS use and decision quality, the number of alternatives, time to reach decision, satisfaction with process, satisfaction with outcome, and consensus. Level 2 GDSS generally lead to greater improvement in performance, satisfaction, and consensus than level 1 GDSS. These results highlight the importance of the modeling and structuring capabilities of GDSS, which also signify the importance of this research. Level 1 GDSS, though it increases meeting effectiveness, lowers the satisfaction of group members. Putting together communication aid (level 1 GDSS) and modeling support (level 2 GDSS) not only leads to greater member satisfaction, but also enhances group performance.

2.3 Summary of Chapter 2

This chapter reviewed the literature on expert support and its explanation facilities, which are the focus and subject of this research. The literature on expert system technology and its explanation facilities provides the foundation that drives the entire research. More specifically, it drives the research question and hypotheses (Chapters 1 and 4), the development of FINALYZER (Chapter 5), and the experimental design and task (Chapters 4 and 5).

This chapter also explained the relationship of this research with GDSS research and reviewed related empirical studies. The review fails to find existing literature on the evaluation of expert system technology and its explanation facilities in the group context. Past empirical studies either evaluated the use of expert system technology by single individuals or examined the use of more conventional decision support facilities by groups. Despite the popular use of expert system technology for supporting decision making and the increasing emphasis on teamwork and group decision making, literature on the
integration of these two areas is lacking. This research, therefore, represents a new branch of group research with its focus on expert system technology and its explanation facilities.
CHAPTER 3: THEORETICAL FOUNDATIONS

This chapter reviews previous research and literature on group decision making, persuasion, expert-novice differences, and the lens model framework which form the theoretical foundations of this research. Section 3.1 reviews the concepts of group process gains and losses, informational and normative influence, and expert power. Section 3.2 reviews theories of persuasion as well as factors influencing it that are related to this research. Section 3.3 reviews the literature on expert-novice differences, the ACT architecture of human information processing model, and the three-stage learning model. Section 3.4 covers the lens model framework which is the conceptual framework for this research. It explains the concept of ESS analyses and explanations support using the lens model framework. The hypotheses of this research, which are presented in Chapter 4, were derived based on the literature covered in this review.

3.1 Related Literature on Group Decision Making

The use of expert support systems and the explanations facilities for group decision making is expected to influence not only the decision outcomes, but also the group process. According to Hackman and Morris (1975), group performance in decision making can be explained by the group interaction process. As such, the literature reviewed in this section concentrates on the process of group decision making and its implications on decision outcomes.

3.1.1 Group Process Gains and Losses

There are advantages and disadvantages for making decisions in groups. These positive and negative factors were identified by Maier (1967) and termed group assets and liabilities, and by Nunamaker, Dennis, Valacich, Vogel, and George (1991) who used the terms group process gains and losses, adopted from Steiner (1972).
According to Steiner (1972):

\[ \text{Actual Group Performance} = \text{Potential Group Performance} - \text{Group Process Losses} \]

where Potential Group Performance includes the process gains from working in a group.

**Process Gains**

When decisions are made in groups, more information and a broader range of knowledge are available for decision making (Shaw, 1981; Steiner, 1972), multiple perspectives and approaches are taken into account (Maier, 1967), synergetic effects can be obtained when members build on one another's ideas (Osborn, 1957; Vangundy, 1984), errors are often checked and corrected (Barnlund, 1959; Hill, 1982; Shaw, 1981), stimulation from the group may encourage individual members to perform better (Shaw, 1981), members may learn and imitate more skilled members to improve performance (Hill, 1982), and acceptance of the decision made increases due to participation in problem solving (Maier, 1967). As a result, the group, as a whole, may gain an increased understanding of the problem, work out solutions more thoroughly, and make better judgments and decisions.

**Process Losses**

However, some negative aspects of group decision making are also observed. When individuals work in groups, more time is consumed in decision making (Husband, 1940), unreasonable social pressure may be used to push for conformity rather than for high-quality solutions (Maier, 1963), engagement in conflict can delay decision making and create ill will among group members (Huber, 1980), and excessively conservative or risky decisions may be made (Myers and Lamm, 1976; Wallach, Kogan, and Bern, 1962). The negative aspects of group decision making are well documented by Janis (1982) and Steiner (1972).
Another problem with many group discussions is that they tend to be disorganized or exhibit a lack of focus. In other words, the problem at hand is often not properly defined or broken down into its component parts for analysis, resulting in incomplete analysis and understanding of the task. This often leads to superficial discussions as well as incomplete access to and use of information necessary for successful task completion (Hirokawa and Pace, 1983). Groups may also face coordination problems. When an appropriate decision making strategy is not used, integration of members' contributions can be difficult, which can lead to dysfunctional cycling or incomplete discussions, resulting in premature decisions.

Cognitive inertia can also be a problem in group decision making (Nunamaker, Dennis, Valacich, Vogel, and George, 1991). This occurs when discussion moves along one train of thought without deviating because group members refrain from contributing comments that are not directly related to the current discussion. In some groups, individual domination occurs when some group member(s) exercise undue influence or monopolize the group's time in an unproductive manner (Maier, 1967). On the other hand, free riding may occur when members rely on others to accomplish goals. This may occur due to cognitive loafing, the need to compete for air time, or when members perceive their input to be unneeded (DeSanctis and Gallupe, 1987; Nunamaker, Dennis, Valacich, Vogel, and George, 1991). Furthermore, some members may withhold ideas or comments during discussion for fear that they may be negatively evaluated.

An overemphasis on social-emotional aspects may also reduce task performance. Members may be reluctant to criticize the comments of others due to politeness or fear of reprisals (Shaw, 1981). The desire to be accepted and to be a good group member tends to silence disagreements, favor consensus, and produce unreasonable social pressure for conformity. Inferior decisions may be made as a result.
Relationship to Current Research

In this research, the process gains and losses that the ESS has effected will be identified in the qualitative analysis presented in Chapter 8 to explain the quantitative analysis presented in Chapters 6 and 7.

3.1.2 Informational versus Normative Influence

Group discussion often has the effect of inducing shifts in both individual opinions and group decisions. Two alternative influence modes by which groups exert influence on members during discussion are informational influence and normative influence (Deutsch and Gerard, 1955; Kaplan and Miller, 1987). Informational influence is based on sharing of facts and persuasive arguments about the issue. It arises when individuals accept information from others as evidence about reality. Normative influence refers to conformity to others' preferences. It is based on the desire to conform to the expectations of others. Thus, informational influence is derived from the desire to be correct, and normative influence stems from concerns with others' reactions when one's opinion differs from others.

Kaplan and Miller (1987) find that intellectual issues (attempting to discover the true or correct answer) elicited more informational influence than normative influence, and judgmental issues (deciding on the moral, valued, or appropriate position) provoked more normative than informational influence. Furthermore, the relatively greater use of normative influence, when the issue is judgmental, and of informational influence, when the issue is intellectual, tend to increase when the decision rule is unanimity rather than majority.
Two concepts that are related to these two modes of influence are persuasive arguments theory (Burnstein and Vinokur, 1977) and social comparison theory (Festinger, 1954). Persuasive arguments theory states that groups will shift toward the point of view that is supported by the largest number of convincing arguments during discussion. It focuses on the cognitive learning that results from exposure to persuasive arguments during the course of the group discussion. This is normally the conjunctive influence of the actual content of the group discussion and the degree to which members are persuaded by the new information presented. Hence, this persuasion may involve mutual reinforcement and is attributed to the sharing of relevant arguments and factual information about the judged issue. On the other hand, social comparison theory proposes that group influence is embedded in members' desire to reevaluate their own preferences in light of thinking about others' choices, and in members' feelings of either external pressures or internal pressures to conform (Seibold and Meyers, 1986).

**Relationship to Current Research**

The way in which ESS has influenced the group decision making will be identified from an analysis of the conversations and group interactions. These will be discussed in Chapter 8 under qualitative analysis.

**3.1.3 Expert Power**

*Expert power* refers to the influence that results when group members perceive another member to have specialized knowledge, information, or skills (French and Raven, 1959). For instance, most people will readily seek advice from professors, physicians, electricians, or attorneys. Because experts are perceived to have specialized knowledge and experience, power is given over to them. Such individuals may exert expert power over the decision makers. Similarly, an ESS and its explanation facilities may also exhibit such power.
The persuasion literature provides support for expert power. The expertise of the source of a message is one of the most important features of the persuasion situation and one of the earliest variables to be investigated. *Expertise* is important in inducing attitude change especially when the advocated position was quite different from the recipients' initial attitude (Petty and Cacioppo, 1981a, p. 64). A communication represented as coming from a high credibility source is more persuasive than the same communication represented as coming from a low credibility source (Hovland and Weiss, 1951). *Provision of supporting arguments by an expert* further increases the persuasive impact of a message (Petty and Cacioppo, 1981a).

Incentives play an important explanatory role in Hovland, Janis and Kelley's (1953) analysis of communicator variables in persuasion. Their prediction that expert sources confer greater persuasion than nonexpert sources was based on the logic that experts' statements are usually regarded as veridical (whereas nonexperts' statements are not), and that holding veridical beliefs and attitudes is inherently reinforcing. Thus, attending to and comprehending expert communicators' arguments, or accepting their recommendations, was assumed to have greater incentive value than learning the arguments or accepting the recommendations presented by nonexperts. Similar reasoning can also be used to explain message recipients' tendency to exhibit greater agreement with the beliefs and attitudes recommended in persuasive messages when the sources of these messages were portrayed as higher in trustworthiness, status, likability, or attractiveness (see the reviews by McGuire, 1969, 1985).

Two information-processing explanations have been considered in explaining the persuasive impact of communicator variables discussed above. One explanation holds that positive source cues such as high expertise increased recipients' motivation to *attend to* and *comprehend* persuasive message content. The other holds that such cues increased
recipients' motivation to accept communicators' recommendations. Hovland and his colleagues (Hovland, Janis, and Kelley, 1953; Hovland and Weiss, 1951) conclude that the persuasive impact of communicator variables was mediated most typically by differential acceptance of messages' recommendations rather than by differential attention to and comprehension of persuasive message content.

Expert power is to be distinguished from informational power or influence. When the acceptance of the truth of fact is based on the credibility of the "expert", it is an actualization of expert power. Informational power is based on characteristics of the stimulus, such as the logic of the argument or the "self-evident facts".

**Relationship to Current Research**

Expert power is a useful and relevant concept for understanding the effects of ESS in this research. This is particularly true for novices who have limited experience and knowledge of a specialized domain. Novices may accept the recommendations given by an ESS because of its expert power. This concept is particularly useful in explaining the hypothesized differences between domain expert and novice users in their attitude and perceptions toward the ESS.

### 3.2 Related Literature on Persuasion

This section covers a review of process theories in persuasion that are relevant to this research. These theories are used for hypotheses generation in Chapter 4, and to explain the phenomena and results in this study. Following the review, factors influencing persuasion that are related to this research will be identified and discussed.
3.2.1 Related Theories on Persuasion

Four process theories of attitude formation and change are covered in this review. They are:

2. Greenwald’s (1968) cognitive response theory (Petty, Ostrom, and Brock, 1981);
3. Petty and Cacioppo’s (1981a, 1986a, 1986b) elaboration likelihood theory;
4. Sherif and Hovland’s (1953, 1961; Hovland and Sherif, 1952) social judgment theory.

The first two theories, which cover McGuire’s information processing paradigm and Greenwald’s cognitive response model, emphasize the importance of message recipients’ detailed processing of persuasive message content in producing new or changed attitudes. The other two theories, which cover Petty and Cacioppo’s elaboration likelihood model and Sherif and Hovland’s social judgment theory, focus on theoretical perspectives that either feature or incorporate mechanisms of attitude formation and change that do not implicate message recipients’ comprehension or elaboration of persuasive message content. These four theories cover different aspects of persuasion, with the idea of central and peripheral routes to persuasion introduced in Petty and Cacioppo’s elaboration likelihood model as the model that attempts to integrate these different ideas and perspectives.

3.2.1.1 McGuire’s Information-Processing Paradigm

McGuire’s information-processing paradigm stemmed directly from Hovland, Janis, and Kelley’s (1953) suggestion that the impact of persuasive communications could be understood in terms of three information-processing phases: 1) attention to the message, 2) comprehension of its context, and 3) acceptance of its conclusions. According to them, independent variables that influence persuasion act not only directly on people’s tendencies
to accept messages' conclusions, but also indirectly through their impact on two causally prior processes, attention and comprehension.

The role that cognitive processes play in persuasion was developed more systematically in the late 1960s by McGuire (1968). He proposed that the persuasive impact of messages could be viewed as the multiplicative product of six information-processing steps: 1) presentation, 2) attention, 3) comprehension, 4) yielding, 5) retention, and 6) behavior. According to this information-processing paradigm, the message recipient must first be presented with the persuasive message. Given that exposure occurs, the recipient must pay attention to the message in order for it to produce attitude change. If the message attracts the recipient's attention, the overall position it advocates and the arguments provided to support this position must be comprehended. It is also necessary that the recipient yield to, or agree with, the message content he has comprehended if any attitude change is to be detectable. And if this change is to persist over a period of time, the message recipient must retain, or store in memory, his changed attitude. Finally, the recipient must behave on the basis of his changed attitude.

Relationship to Current Research
The McGuire's information-processing paradigm is relevant to this research because it explains the role of ESS explanations and the reasons for their provision. The ESS explanation facilities provide users the opportunity to comprehend the reasoning behind the conclusions. In this way, they convince users of the validity of ESS advice, thus increasing their acceptability of the conclusions.

3.2.1.2 Greenwald's Cognitive Response Model
The cognitive response approach shares with the Hovland, Janis, and Kelley (1953) and McGuire (e.g., 1972) frameworks the assumption that some kind of learning plays a role in
determining attitude change and its temporal persistence. However, whereas the Hovland group, and especially, McGuire emphasized the mediational role of reception processes, the cognitive response approach emphasizes the mediating role of idiosyncratic thoughts or "cognitive responses" that recipients generate — and, thus, rehearse and learn — as they receive and reflect upon persuasive communications.

Thoughts that people generate on their own can be as effective in producing attitude changes as messages that originate externally (sometimes even more effective!) (Petty and Cacioppo, 1981). The cognitive response approach takes this reasoning one step further: it contends that even the persuasion that results from exposure to externally originated messages is due to the thoughts that the message recipient generates in response to the communication. These thoughts generated in response to the communication are called cognitive responses and are the end result of information processing activity.

According to the cognitive response model, people actively relate information contained in persuasive messages to their existing feelings and beliefs about the message topic. Cognitive responses represent the content of this internal communication on the part of the message recipients and are assumed to reflect recipient-generated thoughts that are not merely repetitions of message content. Most importantly, the model assumes that cognitive responses mediate the effect of persuasive messages on attitude change. Messages that evoke predominantly favorable recipient-generated thoughts should be persuasive, whereas those that evoke mostly unfavorable thoughts should be unpersuasive (and may even result in attitudes that are less favorable to the advocacy than the recipients' prior attitudes). In essence, then, the cognitive response model asserts that the cognitions generated in response to persuasive messages determine both the direction and magnitude of attitude change (Greenwald, 1968; Petty, Ostrom, and Brock, 1981).
**Relationship to Current Research**

The provision of ESS conclusions and explanations increases users' chances of rehearsing and learning from these messages, which in turn leads to persuasion. The more "cognitive responses" or thoughts recipients generate from these messages, the greater the persuasion.

### 3.2.1.3 Petty and Cacioppo’s Elaboration Likelihood Model

The McGuire's information processing theory and Greenwald's cognitive response theory are *systematic* process theories that emphasize the role of people's reception and cognitive elaboration of persuasive argumentation in producing new or changed attitudes. On the other hand, Petty and Cacioppo's elaboration likelihood model and Sherif and Hovland's social judgment theory incorporate or feature the idea that people may adopt attitudes on bases *other* than their understanding and evaluation of persuasive argumentation. The elaboration likelihood model (ELM) (Petty and Cacioppo, 1981a, 1986a, 1986b) incorporates this viewpoint by positing a *peripheral route* to persuasion — persuasion that occurs in the *absence* of argument scrutiny.

Although similar to the cognitive response model in many respects, the ELM offers an extended view of persuasion insofar as it 1) specifies the conditions under which persuasion should be mediated by message-related thinking, and 2) postulates that alternative *peripheral* mechanisms account for persuasion when these conditions are not met. The model also attempts to place existing persuasion theory and research under one conceptual umbrella. Petty and Cacioppo specify two *qualitatively* different routes to persuasion in the ELM and assert that most attitude theories can be viewed as exemplifying one or the other. Theories emphasizing the mediational importance of argument-based thinking — theories termed *systematic* — are labeled *central route* perspectives. Two examples of systematic theories are McGuire's information processing theory and Greenwald's cognitive response theory. In contrast, theories that specify psychological
mechanisms that do not implicate argument processing — theories termed heuristic — are labeled peripheral perspectives. The ELM is therefore a heuristic-systematic model.

Petty and Cacioppo suggest that source, message, and other variables that influence persuasion may do so in one or more of three distinct ways: 1) serving as persuasive arguments — "bits of information contained in a communication that are relevant to a person's subjective determination of the true merits of an advocated position", 2) serving as peripheral cues, thereby affecting persuasion via one or another peripheral mechanism, 3) affecting either motivation or ability for elaboration, thereby moderating the route to persuasion. For example, when people receive a personally relevant message under non-distracting conditions (establishing high motivation and high ability), the model predicts that central route persuasion will take place. However, when this same message is received under highly distracting conditions, or when a personally irrelevant message is received, the model predicts that peripheral route persuasion will occur. Importantly, Petty and Cacioppo also propose that variables influencing the elaboration likelihood may lead to either objective message processing or biased message processing.

Figure 3-1 illustrates the antecedents and consequences of the elaboration likelihood model's central and peripheral routes to persuasion. The model assumes that people desire to attain correct attitudes, but that the extent and nature of their processing of persuasive arguments depends upon motivation and ability, as addressed in the cognitive response model. The term elaboration in the model refers to the extent to which people think about issue-relevant arguments contained in persuasive messages. When situational and individual difference variables ensure high motivation and ability for issue-relevant thinking, the elaboration likelihood is said to be high. As a consequence, the probability that recipients follow the central route is high. Attitudes formed or changed via the central route are hypothesized to be relatively persistent, predictive of behavior, and resistant to
change until challenged by convincing counterarguments. For instance, since experts tend to possess a higher level of ability to process messages on an issue or topic, they are more likely to employ the central route and hence, they are more likely to persist and resist changing their attitudes.

However, when motivation or ability for elaboration is low, attitudes are still formed and changed under such conditions, but not via the central route. Instead, when the elaboration likelihood is low, the probability is high that the recipients will follow the peripheral route to persuasion. This type of persuasion is regarded as more ephemeral than central route persuasion. These peripheral mechanisms include cognitive, affective, and social role mechanisms. Like the central route, the peripheral route refers to a family of attitude theories. Because the distinction between the two families is that central route perspectives emphasize argument-based processing whereas peripheral route perspectives do not, the model’s two routes are complements of one another.

In summary, the elaboration likelihood model not only integrates the existing persuasion theories into one framework, but it also provides a comprehensive overview of the persuasion process and the factors influencing it.
Figure 3-1: Elaboration Likelihood Model

Relationship to Current Research

The elaboration likelihood model predicts that novices are more likely to take the peripheral route to persuasion than experts. By taking the peripheral route, novices accept the messages
based on peripheral cues, such as the perceived expertise of the ESS. Novices would trust the conclusions of the ESS due to its source credibility, i.e., based on the fact that the conclusions were derived from the knowledge of experts. This is similar in concept to expert power. On the other hand, experts are more likely to take the central route to persuasion. Therefore, ESS explanation facilities are more important to experts than novices. The explanation facilities increase the influence of ESS on users by providing a means for users to comprehend its conclusions. According to the information processing paradigm discussed earlier, increased comprehension leads to increased persuasion.

3.2.1.4 Sherif and Hovland's Social Judgment Theory

Social judgment theory features mechanisms of attitude formation and change that can occur in the absence of argument-based processing. It emphasizes how people’s prior attitudes affect their perceptions of the attitude positions that communicators express and how these perceptions, in turn, influence agreement with persuasive communications.

The social judgment theory has also been called assimilation-contrast theory (e.g., Insko, 1967) and the social judgment-involvement approach (e.g., Sherif, Sherif, and Nebergall 1965). These names highlight the theory’s key constructs: 1) assimilation and contrast effects in perception, and 2) ego-involvement, the extent to which an attitude is part of one’s self-concept and thus “intimately felt and cherished”.

Assimilation and Contrast

In general terms, the social judgment theory assumes that a recipient’s own attitudinal position serves as a judgmental standard or anchor that influences where along an evaluative continuum a communicator’s advocated position is perceived to lie. Attitudes that are relatively close to one’s own are assimilated (perceived to be closer than they
actually are), but attitudes that are very discrepant from one’s own are *contrasted* (perceived to be further than they actually are).

Petty and Cacioppo (1981a, 1986b) classify social judgment theory as a peripheral theory of attitude change because the *perceptual* phenomena of assimilation and contrast occur prior to the argument-based processing that epitomizes central route persuasion. Social judgment theory recognizes the influence of motivation by its ego-involvement construct and the influence of cognition by its perceptual mechanisms of assimilation and contrast. It assumes that perceptual processes mediate the persuasive effects of ego-involvement and other distal variables.

*Latitudes of Acceptance, Rejection, and Noncommitment*

Social judgment theory explains the impact of people’s prior attitudes on the encoding of attitude-relevant information. It draws upon the basic assumption that attitude is best conceptualized as a range of acceptable attitudinal positions rather than a single point along an evaluative continuum. The theory posits a tripartite division, i.e., three categories, of the evaluative continuum. The *latitude of acceptance* refers to that region of the continuum that contains beliefs about the attitude object that the individual considers acceptable — and within which his attitude lies. On the other hand, the *latitude of rejection* contains beliefs that are viewed as unacceptable, and the *latitude of noncommitment* contains beliefs that are considered neither acceptable nor unacceptable. Whether or not the advocated position of a communication falls within a person’s latitude of acceptance or rejection will thus be a primary determinant of whether or not persuasion will occur.

Specifically, the widths of people’s latitudes and the location of their preferred attitude position within the latitude of acceptance determine how they judge attitude statements. When a single statement or position advocated in a message falls within the latitude of
acceptance or nearby in the latitude of noncommitment, assimilation occurs — the stimulus statement or position is seen as closer to the person’s own attitude “anchor” than it truly is. Within this assimilation range true discrepancy is underestimated, and the magnitude of underestimation grows larger as true discrepancy increases. When an attitude statement or advocated position falls within the latitude of rejection or just outside this range in the latitude of noncommitment, contrast occurs — the statement or position is perceived to be farther from the person’s own attitude than it truly is. Within this contrast range, true discrepancy is overestimated and the magnitude of overestimation grows larger as true discrepancy increases.

*Explaining Attitude Change*

So far, only the first part of social judgment theory’s thesis for persuasion — people’s existing attitudes distort their perception of the positions advocated in communicators’ messages — has been explained. The second part of the thesis is that these perceptual displacements mediate persuasion. When a communication falls within the latitude of acceptance, its position is assimilated, its content is positively evaluated, and attitude change occurs. Moreover, in this latitude (or nearby), greater levels of true discrepancy between the message’s position and the recipient’s own attitude lead to increasing degrees of assimilation and positive evaluation and thus, to greater amounts of attitude change. In a parallel fashion, when a message falls within the latitude of rejection, its position is contrasted, its content is negatively evaluated, and attitude change is inhibited. In this latitude, higher levels of true discrepancy lead to increasing degrees of contrast and negative evaluation and, thus, to increasingly lesser amounts of attitude change. (In fact, social judgment theorists further suggested that for recipients who were higher ego-involved in their attitudes, extreme levels of discrepancy might sometimes result in negative or *boomerang* attitude change, wherein recipients shift their attitudes in a direction opposite to that advocated in the message.)
In general, social judgment theory posits an inverted U-shaped relation between message discrepancy and attitude change (see Figure 3-2). As discrepancy increases from minuscule to moderate, increasing degrees of attitude change should be observed since the message is likely to fall within the recipient's latitude of acceptance or just beyond it in the latitude of noncommitment. After this point, however, further increases in discrepancy should produce decreasing amounts of attitude change because the message is increasingly likely to fall within the latitude of rejection. Importantly, though, the location of the point along the discrepancy continuum where attitude change ceases to increase and starts to decline depends primarily on latitude width. This inflection point should occur at lower levels of true discrepancy for recipients whose latitudes of rejection are wider rather than narrower and, similarly, for recipients whose latitudes of acceptance are narrower rather than wider.

Figure 3-2: The Curvilinear (Inverted-U Shaped) Relationship between Message Discrepancy and Attitude Change
According to social judgment theory, any factor that influences latitude widths should exert a corresponding influence on the shape of the relation between message discrepancy and persuasion. Heightened source credibility, such as status, expertise, or trustworthiness, might function to extend recipients' latitudes of acceptance (Sherif and Sherif, 1967) and raise the inflection point in the discrepancy-persuasion relation. To cite another example, Sherif and Hovland (1961) suggest that recipients who lack established (i.e., strong) attitudes should have very broad latitudes of acceptance. If so, the discrepancy-persuasion relationship ought to prove predominately positive for such persons, even at relatively high discrepancy levels. Although other width-affecting factors may exist, social judgment theory accords a special role to one, the recipient's degree of ego-involvement in his/her attitude.

**Role of Ego-Involvement**

Latitude width is assumed to vary as a function of *ego-involvement*. Sherif and Cantril (1947) define ego-involved attitudes as those that are part of the person's self-concept or "ego", attitudes that "have the characteristic of belonging to *me*, as being part of *me*" (p. 93). Sherif and Cantril (1947) view ego-involvement as having important motivational and affective consequences:

> This degree of ego-involvement, this intensity of attitudes, will determine in large part which attitudes he will cling to, how annoyed or frustrated he will feel when his attitudes are opposed, what action (within the range of his individual temperament and ability) he will take to further his point of view (p. 131).

The concept of "ego-involvement" is important in this research as the majority of experts possess such a characteristic. Experts are known to possess special skills, knowledge, or experience in a particular domain and would feel frustrated, annoyed, or even "threatened" if their status or ideas are challenged. Social judgment theory assumes that exposure to
discrepant attitude positions creates little "tension" or "incongruity" for the uninvolved
person, but a great deal of psychological discomfort for the ego-involved person (Sherif
and Sherif, 1967, p. 130). The ego-involved person "perceives his stands as part of what he
is and what he claims to be. His personal identity and the stability of his conception of
himself depend in no small part on the stability and perpetuation of his stands" (Sherif and
Hovland, 1961). Because of this need to maintain and protect the self-concept, the ego-
involved person is presumed by the theory to become highly engaged in attitude-relevant
tasks and to encode attitudinal information in a highly personalized, self-protective
manner. By contrast, the uninvolved person is presumed to be less personally engaged in
such tasks and to encode attitudinal information in a relatively detached, objective, and
factual manner (Sherif and Hovland, 1961).

These motivational assumptions are integrated with cognitive aspects of the theory in two
main ways: 1) Ego-involvement is assumed to strengthen the anchoring effects of prior
attitudes – the more involved the individual is, the more likely his or her attitude will serve
as an internal reference point in judging attitudinal stimuli. Thus, the magnitude of
assimilation-contrast tendencies in attitudinal perception should be a positive function of
ego-involvement (Sherif and Hovland, 1961; Sherif and Sherif, 1967). In other words, the
tendency to assimilate a message that falls within the latitude of acceptance should be
greater for the highly involved recipient than for the uninvolved recipient; similarly, the
involved person should manifest a greater tendency to contrast a message that falls within
the latitude of rejection; 2) Ego-involvement affects latitude width and exerts a marked
influence on people's tolerance for beliefs different from their own. Initially, it was
hypothesized that ego-involved attitudes would be characterized by broader latitudes of
rejection and narrower latitudes of acceptance. However, results of studies indicate that
heightened ego-involvement was typically associated with broader latitudes of rejection
and small-to-nonexistent latitudes of noncommitment. Although ego-involvement
broadens the latitude of rejection and narrows the latitude of noncommitment, ego-involvement bore little relation to the width of latitudes of acceptance.

In short, *since highly involved persons have larger latitudes of rejection and little, if any, latitude of noncommitment, they should generally be more resistant to persuasion than less involved persons* because any given message has a greater probability of falling in the rejection region for the highly involved. As Sherif and Sherif (1967, p. 133) have strongly stated: “Regardless of the discrepancy of the position presented, we predict that the more the person is involved in the issue, the less susceptible he will be to short-term attempts to change his attitude.”

**Relationship to Current Research**

According to the social judgment theory, experts are less likely than novices to accept the conclusions provided by the ESS. As experts tend to be more ego-involved, they have a larger latitude of rejection, indicating that they are more likely to strongly reject arguments that are different from theirs. The social judgment theory also suggests that experts tend to be critical not only with the ESS but also among themselves, thus making it more difficult for experts to achieve true consensus than novices.

### 3.2.2 Factors Influencing Persuasion

From the four process theories above, the relevant factors in this research that influence attitude change were identified. These factors are listed below under three main groupings: 1) *Message* factors, 2) *Recipient* factors, and 3) *Source* factors.
Message Factors

(1) Message Comprehensibility, Argument Quantity, and Argument Quality

Message comprehensibility, and argument quantity and quality may impact persuasion primarily when recipients are more concerned with maximizing the validity of their attitudes (i.e., taking the central rather than the peripheral route) than with achieving other, more interpersonal goals (Eagly and Chaiken, 1993, Norman, 1976).

As suggested by McGuire’s (1968, 1989, 1972, 1985) information processing paradigm and Hovland, Janis, and Kelley’s (1953) message learning approach, message comprehension is necessary in order for persuasion to take place. Lowering message comprehensibility lowers recipients’ retention of persuasive arguments, and more important, significantly lessens their agreement with the message’s recommendations (Eagly, 1974). The decreased persuasiveness of low comprehensibility messages may also take place due to the negative affect that message recipients experienced as they tried to comprehend the message (Eagly, 1974). Lower message comprehensibility presumably decreases the persuasiveness of (high quality) messages by lessening the amount of supportive argumentation received (Eagly and Chaiken, 1993). The tendency for persuasion to decrease when fewer (strong) arguments are presented has been well documented in the persuasion literature (Calder, Insko, and Yandell, 1974; Insko, Lind, and LaTour, 1976).

Argument quantity may sometimes affect message acceptance directly, by influencing people’s global judgments of message validity (Petty and Cacioppo, 1984a). Although it is generally true that increasing the number of (strong) supporting arguments increases the effectiveness of a message, persuasion does not increase invariably with quantity of persuasive argumentation. The ELM predicts that the quality of persuasive argumentation influences attitude judgments more when recipients are highly motivated and/or highly able
to engage in elaborative processing. In the low relevance (i.e., low ability and low motivation) condition, persuasion tends to increase with the number of arguments (until the saturation point is reached) regardless of argument quality. In the high relevance (i.e., high ability and high motivation) condition, recipients agree significantly more with strong than with weak messages.

Consistent with the cognitive response model, increasing the number of high quality arguments can increase persuasion while increasing the number of low quality arguments can reduce it (Petty and Cacioppo, 1984). Presenting a large number of low quality arguments may reduce persuasion by lowering recipient's global judgments of message validity or perceived credibility of the source.

(2) Message Repetition
According to the message-learning approach and the information-processing paradigm, repetition should enhance the total attention to, comprehension of, and retention of a message, which in turn enhances recipients’ abilities to engage in cognitive responding (Cacioppo and Petty, 1985). Agreement with high quality messages (i.e., those supported by strong arguments) increases with the amount of exposure (Cacioppo and Petty, 1979), while agreement with low quality messages decreases with increased exposure (Cacioppo and Petty, 1985). Although repeated exposure to a high quality persuasive message increases attitude change toward the position advocated by the message, the level of agreement does not increase further when repetition reaches a “tedious” level (about three exposures) (Cacioppo and Petty, 1979). Cacioppo and Petty interpreted their curvilinear persuasion data as reflecting a two-phase cognitive elaboration-tedium process. In the first phase, repeated message exposure should increase recipients’ opportunities to cognitively elaborate the message’s arguments; therefore repeated exposure increases persuasion for high quality messages but decreases persuasion for low quality messages. When repetition
reaches a "tedious" level, however, a second level is initiated in which feelings of boredom or psychological reactance (Brehm, 1972) are presumably experienced. During this tedium level, recipients become motivated to reject the message regardless of the inherent quality of its arguments.

Recipient Factors

(1) Recipients’ Level of Issue Involvement

*Issue involvement* refers to the extent to which recipients perceive that a *message topic* is personally important or relevant (Johnson and Eagly, 1989; Petty and Cacioppo, 1979b, 1986a, 1990). In other words, when *involvement* increases, people’s motivation to engage in message- and issue-relevant thinking increases, and therefore the message content becomes a more important determinant of persuasion than source credibility (Petty and Cacioppo, 1979). Several experimental studies have supported the cognitive response hypothesis that increased issue involvement enhances persuasion with strong messages but inhibits persuasion with weak messages (Petty and Cacioppo, 1981b, 1984a; Petty, Cacioppo and Goldman, 1981; Petty, Cacioppo and Schumann, 1983).

(2) Recipients’ Level of Ego-Involvement

Johnson and Eagly (1989) distinguish between *issue involvement* and *ego-involvement*. They term the former *outcome-relevant* involvement and the latter *value-relevant* involvement. Johnson and Eagly’s (1989) meta-analysis on value-relevant involvement (or ego-involvement) shows that this type of involvement tends to reduce persuasion, regardless of *argument quality*. Given counterattitudinal messages, social judgment theory predicts less persuasion for ego-involved recipients, for example, experts in a particular domain. As counterattitudinal messages become more extreme, involvement differences in persuasibility increase in magnitude. Overall susceptibility to influence as a function of
ego-involvement can be described by McGuire's information-processing model, with ego-involvement being negatively related to yielding.

(3) Recipients' Amount of Knowledge
According to ELM, the more knowledge the recipient of a message has on the topic or issue under discussion, the more likely it is for the recipient to engage in the central route to persuasion, and thus the more difficult it will be to find any effects of source credibility (Rhine and Severance, 1970). In other words, when prior knowledge increases, people's motivation to engage in message- and issue-relevant thinking increases, and therefore the message content becomes a more important determinant of persuasion than source credibility (Petty and Cacioppo, 1979).

As predicted by ELM, high-knowledge recipients process message content more extensively than low-knowledge subjects do (Wood and Kallgren, 1985). High knowledge recipients also generate more negative thoughts about the messages and are less persuaded by them than low knowledge recipients (Wood and Kallgren, 1985). This effect is more pronounced for weak messages than for strong messages. Research evidence supports the heightened criticality hypothesis that high knowledge recipients have greater ability to critically evaluate even strongly argued messages and, as a result, tend to do so (Biek and Wood, 1996). This explains why high knowledge recipients tend to be less persuaded than low knowledge recipients.

Source Factor
(1) Source Expertise
The effect of source credibility has been discussed earlier under "expert power". However, in the earlier discussion, the moderating role of prior knowledge (ability) and issue involvement (motivation) was not explicitly discussed. Other than prior knowledge, time
pressure, message comprehensibility, and recipient mood are also presumed to influence ability for processing. As illustrated by ELM, *ability* and *motivation* moderate the route to persuasion. The first half of Petty and Cacioppo’s key hypothesis that elaboration likelihood moderates the route to persuasion shows that argument quality determines persuasion when motivation and ability for message processing are high. The other half of the hypothesis states that peripheral cues determine persuasion when motivation and ability for processing are low. However, these peripheral cues are relatively unimportant determinants when motivation and ability are high.

In the low relevance (i.e., low ability and low motivation) condition, source expertise has a significant influence on recipients’ post-message attitudes, presumably through some peripheral mechanism such as heuristic processing (e.g., “Experts’ statements can be trusted”). In the high relevance (i.e., high ability and high motivation) condition, source expertise has no impact on attitudes; only argument quality influences persuasion (Petty, Cacioppo, and Goldman, 1981). Other studies that have manipulated motivation for processing in conjunction with other source variables (e.g., communicator likability, attractiveness) yield virtually identical findings (Chaiken, 1980; Petty, Cacioppo, and Schumann, 1983). Whereas source variables exert little persuasive impact under high ability conditions, their effect on persuasion is typically substantial when recipients lack ability for extensive processing.

Finally, the effect of source credibility can also be explained using social judgment theory. Heightened social credibility increases the width of the latitude of acceptance or, alternatively, narrows the latitude of rejection. Therefore, maximum persuasion occurs at a higher level of message discrepancy when communicators are credible and a lower level of message discrepancy when communicators lack credibility (Aronson, Turner, and Carlsmith, 1963).
Summary

In summary, the characteristics of the message, source, and recipient exert differing effects on information processing, attitude change, and persistence. These effects can be analyzed using McGuire's information processing paradigm, Greenwald's cognitive response model, Petty and Cacioppo's elaboration likelihood model, and Sherif and Hovland's social judgment theory.

3.3 Expert-Novice Differences

In this research, a novice is defined as one who has sufficient basic knowledge about the problem domain but no working experience. On the other hand, an expert is one who has both basic knowledge about the problem domain and experience in performing the task. Experts differ from novices in that the experts often know more, and in a more elaborate way, about the subject than the novices. Compared to novices, experts not only have more complete knowledge (facts, laws, principles, heuristics) but also have better cross-referencing and memory organization, and a superior mechanism for relating problems to appropriate knowledge and courses of action (Bedard, 1991). Davis (1996) and Etenson, Shanteau, and Krogstad (1987) have identified one major difference in problem solving between experts and novices. Experts exhibit a higher level of selective attention to relevant information and based their judgments on fewer cues than novices.

Kail and Bisanz (1980) suggest that a fundamental difference between expert and novice decision making appears to be caused by the expert's ability to take fundamental strategies (templates) that have been taught, and through practice, modify these into more efficient and powerful procedures. The resulting expert strategies may be too complex to be taught; therefore, the only method of acquiring them is through experience. In this sense, a strategy refers to the structures or rules that underlie performance on cognitive tasks.
Miller (1956) and Chase and Simon (1973) examined an organization strategy known as “chunking” that is utilized by experts. The experts’ ability to perceive information about a problem in “chunks” of structural knowledge enhanced later recall and efficiency of performance. Experts can account for large amounts of information internally and can integrate different pieces of information that are not simultaneously available perceptually. They have superior memory skills in recognizing patterns in their domains of expertise. Their internal representation of the available information is sufficiently precise to allow extensive reasoning and evaluation of consistency, and sufficiently flexible to allow reinterpretation as new information becomes available (Lesgold, 1984).

Expert-novice differences can be summarized as follows. In comparison to novices, domain experts 1) perceive large meaningful patterns in their domain through chunking (Chase and Simon, 1973), 2) have superior short- and long-term memory for domain-relevant information (Chase and Ericsson, 1982), 3) are faster at performing the basic skills of the domain (Chi, Glaser, and Farr, 1988), 4) represent problems at a deeper (principled) and more elaborate level (Weiser and Shertz, 1983), 5) make judgments based on fewer and more relevant cues (Davis, 1996), and 6) spend a great deal of time analyzing a problem before attempting a solution (Voss and Post, 1988).

The following sections review the differences between experts’ and novices’ cognitive information processing systems.
3.3.1 Architecture of Adaptive Control of Thought (ACT)¹

One of the most popular and well-known human information-processing models is the Adaptive Control of Thought (ACT) architecture proposed by Anderson (1985, 1993, 1995). The ACT architecture consists of three types of memories — declarative, production (or procedural), and working — as shown in Figure 3-3.

![Figure 3-3: The ACT Architecture](image)

Declarative and production memories are long term memories that store declarative and production knowledge respectively. On the other hand, working memory is short term. Declarative memory contains factual knowledge that humans can report or describe, whereas production memory is knowledge that can lead to performance (Anderson 1993).

---

¹ This section is adapted, with permission, from the research work of Keng L. Siau.
In plain language, declarative knowledge is knowing that something is the case, and production knowledge is knowing how to do something. Working memory contains information to which the system currently has access. It consists of information retrieved from long-term declarative memory as well as temporary structures deposited by encoding processes and the action of productions that are stored in the production (or procedural) memory. Experts generally have access to both production (or procedural) and declarative knowledge. Novices generally have the declarative knowledge to solve a problem, but they lack abstracted procedural knowledge (Chase and Simon, 1973; de Groot, 1966).

Production knowledge can be differentiated in two dimensions. The first is the domain-specificity dimension and the second is the automated versus controlled dimension. The former, domain-general versus domain-specific dimension, refers to the degree to which production knowledge is tied to a specific domain. Domain-general knowledge is applicable across domains, and domain-specific knowledge is specialized because it is specific to a particular domain. In this research, expertise refers to domain-specific knowledge in financial statement analysis. The second dimension of production knowledge is "degree of automation" with the end points of the continuum being labeled automatic and controlled (or conscious) (Gagne et al. 1985). An automated process or procedure is one that consumes few cognitive resources of the information-processing system. A controlled process, on the other hand, is knowledge that underlies deliberate thinking because it is under the conscious control of the thinker. Experts rely more on automated problem solving processes (Shiffrin and Schneider, 1977). Automated processes are often parallel and function independently, somewhat like visual perception or pattern recognition. Controlled processes, on the other hand, are linear and sequential, more like deductive reasoning. With practice, some control processes may become automatized over time (Larkin et al., 1980). As they gain experience, experts come to rely less on deductive
thinking and more on pattern recognition-like thinking. Related to this concept of “degree of automation” is the three-stage learning model which is described next.

3.3.2 *Three-Stage Learning Model*\(^2\)

A theory that describes the cognitive changes that occur in the evolution of expertise is the "Three-Stage Learning Model" proposed by Anderson (1982, 1995) and illustrated in Figure 3-4. The three stages in the model are (Fitts 1964, Anderson 1982): cognitive stage, associative stage, and automatic stage.

![Three-Stage Learning Model](image)

*Figure 3-4: Three-Stage Learning Model*

The "cognitive stage" is characterized by the discovery of relevant aspects of the task and the storage of declarative knowledge about the skills. It is an effort to understand the task and its demand and to learn which information one must attend to. The "associative stage" involves making the cognitive processes efficient, allowing rapid retrieval and perception of required information. Thus, during the associative stage, skills are chunked, or compiled, into procedural knowledge. At the "autonomous stage", performance is

\(^2\) This section is adapted, with permission, from the research work of Keng L. Siau.
automatic, and conscious cognition is minimal. The procedures of the basic skills undergo a process of continual refinement (i.e., tuning) and strengthening, which increases performance, speed, and accuracy. Novices perform at the cognitive or associative level, and require conscious cognitive effort in carrying out the task. On the other hand, expert performance is automatized because they have reached the autonomous stage.

Conceptual understanding is housed in declarative knowledge. In solving a problem, conceptual understanding helps the problem solver develop a meaningful representation of the problem and narrows the search for solutions by matching the schema with conditions of productions in the procedural memory. Domain-specific skills and domain-specific strategies are housed in procedural knowledge. Domain-specific strategies help make both the search process and the evaluation of the outcome faster than would otherwise be the case. Automated basic skills allow the problem solver to perform necessary, routine mental operations without thinking much about them.

The three-stage learning model is fairly similar in concept to the typology of strategy transformations proposed by Neches and Hayes (1978) as a basis for understanding how individuals modify their strategies through experience. Neches and Hayes (1978) identify three strategy transformations: unit building, reduction to a rule, and deletion of unnecessary parts. Unit building allows the combination of groups of operations into a set that can be accessed as a single unit. Reduction to a rule replaces a procedure with a rule describing its results. This rule is constructed through the experience of observing constant relations within ordered sets of results or across the pairs of inputs and results. Deletion of unnecessary parts simplifies the flow of control by eliminating nonessential operations. Through practice, the decision maker can determine the operations that are minimally sufficient to solve the problem at hand. All three of these strategy transformations result in more efficient and effective decision making by reducing or combining the information.
(cues) that the decision maker needs to address. The result of strategy transformation, more efficient and effective decision making, is equivalent to the factual definition of learning.

3.4 The Lens Model Framework

The lens model is a generic decision making framework that originates from the psychology literature and is widely used in the accounting literature. It is used as the conceptual framework for this research because the different task supports provided by the ESS map onto the lens model (Dhaliwal and Benbasat, 1996). The lens model is typically used to derive the relationships between the criterion, cues, and judgment. However, as this research only considered one experimental case, it would not be possible to determine these relationships. As such, the lens model framework is only used for illustration purposes, i.e., as a conceptual framework, to explain the different levels of decision support in this research.

The lens model framework is based on Brunswik's (1952, 1956) theory of perception or the so-called cue theory. According to the theory, decision makers do not have direct access to information about the objects in the environment. Instead, perception is an indirect process, mediated by a set of proximal cues. In accordance with this view, Hammond et al. (1975) define judgment as a process which involves the integration of information from a set of cues into a judgment about some distal state of affairs.

The lens model, as illustrated in Figure 3-5, defines the unit for psychological analysis as a system consisting of two subsystems. These subsystems have a common interface which consists of the proximal cues in perception. The two subsystems in the model are the task system and the cognitive (or judgmental) system. The task system is defined in terms of the relations between the cues ($X_i$) and the distal variable ($Y_e$) of interest to the person, as well as the relations among the cues ($X_i$). The cognitive system is defined in terms of the
relations between the cues ($X_i$) and the judgment ($Y_s$). A review of the “multiple-cue probability learning” literature by Balzer, Doherty, and O’Connor (1989) indicates that it is the task information (i.e., information about the relations in the task environment) rather than the cognitive information (i.e., relations perceived by the decision maker) that influences performance.

![Brunswik's Standard Lens Model](image)

**Figure 3-5: Brunswik's Standard Lens Model**

As an example, in graduate admissions decision making, the distal variable or criterion ($Y_e$) may be graduate grades or the attainment of a Ph.D. because the decision makers base their decisions mainly on the projected academic performance of the applicants. The cues ($X_i$) that are typically used to make the prediction or judgment ($Y_s$) include the GMAT score, GPA, and the quality of undergraduate institution.

### 3.4.1 Mapping between ESS Support and Lens Model

The support provided by an ESS corresponds to information about the task environment (Dhaliwal, 1993; Dhaliwal and Benbasat, 1996). Put into the framework of the lens model, the two fundamental types of task support provided by the ESS – the ability to draw conclusions and give advice, and the ability to explain its knowledge, reasoning and
conclusions — complement one another in providing a more complete set of information about the task environment. In this case, where the task support is provided by an ESS, the task support refers to the cognitive information of the expert(s) whose knowledge was acquired in the development of the system. The ESS, named FINALYZER, used in this research was developed based on the consensus of five experts (Dhaliwal, 1993). Thus, the task support provided by FINALYZER refers to the cognitive information that the group of experts agreed upon through consensus. The cognitive system in this case refers to the cognitive information of the decision maker(s).

Figure 3-6 illustrates the different levels of support investigated in this study using the lens model framework. Figure 3-6(a) illustrates the scenario where no ESS support is available. In this case, the decision maker has to make judgments solely from cues available in the environment. The lens model in Figure 3-6(b) illustrates the situation where the decision maker is supported by ESS analyses without any explanations support. In this case, only advice about the criterion ($Y_e$) is provided (illustrated in Figure 3-6(b) using the rectangle that enclosed the criterion). Figure 3-6(c) depicts the context where the decision maker is supported by both ESS analyses and explanations support. In this case, in addition to advice about the criterion, explanations about how the criterion ($Y_e$) relates to the cues ($X_i$) as well as how the cues relate to one another are also provided (represented in Figure 3-6(c) by the thick lines connecting the cues and the criterion, and the cues themselves). Figure 3-6(a,b,c) thus illustrates increasing level of ESS support.
In this research, since the true value of the criterion \((Y_e)\) is not known, it was determined by the consensus of the group of experts involved in developing FINALYZER (Dhaliwal, 1993). Thus, the term \textit{"consistency with original experts\textquotedblright} is used throughout this dissertation to describe the deviation between the judgments of the decision makers and those of the original experts. Also, \textit{FINALYZER does not provide the value of the criterion \((Y_e)\) to the decision maker(s); instead, the decision maker(s) will use the analyses and explanations support provided by FINALYZER to make a judgment \((Y_s)\) of the criterion.}

\textit{Relationship to Current Research}

Since the level of ESS support corresponds conceptually with task support in the lens model, we draw upon the literature on the lens model to hypothesize that higher levels of ESS support create greater consistency of judgments between the decision makers and the original experts who developed the ESS.
3.5 Summary of Chapter 3

This chapter reviewed the literature on four main areas: 1) relevant concepts in group decision making, 2) theories on persuasion and factors influencing persuasion, 3) expert-novice differences, and 4) the lens model framework which is the conceptual framework for this research. The literature on group decision making is useful for explaining the group processes that are covered in the qualitative analysis (Chapter 8). Four theories of persuasion, namely, McGuire's information processing paradigm, Greenwald's cognitive response model, Petty and Cacioppo's elaboration likelihood model, and Sherif and Hovland's social judgment theory were reviewed. These theories are used in Chapter 4 to explain the derivation of hypotheses, and in Chapters 6, 7, and 8 to explain the quantitative and qualitative results. The literature on expert-novice differences is used in Chapter 4 to explain the derivation of hypotheses for this research. The lens model framework is reviewed because it is the conceptual framework for this research.
CHAPTER 4: RESEARCH FRAMEWORK, DESIGN, AND HYPOTHESES

This chapter presents the research framework and design and uses the theories covered in Chapter 3 to derive the hypotheses for this research. Section 4.1 presents the research framework. Section 4.2 discusses the research design and the subject recruitment process. Section 4.3 derives the research hypotheses on the effects of ESS analyses and explanations support on group decision making by experts and novices.

4.1 Research Framework

Three levels of decision support were proposed for investigation in this experimental study: 1) no ESS support, 2) ESS analyses support only, and 3) ESS analyses plus explanations support. With ESS analyses support only, the system provided advice and analyses on the case, but not explanations relating to these advice or the task. With ESS analyses plus explanations support, the subjects were given the option to access explanations relating to the analyses and the task.

Several studies have found that domain novices and experts not only utilize the expert support technology in different ways (Dhaliwal, 1993; Gregor, 1996; Mao, 1995) but are also influenced by the system to different extent (Lamberti and Wallace, 1990; Peterson, 1988). As such, a second independent variable, domain expert versus novice users, was introduced into this study. This study differs from earlier studies (e.g., Dhaliwal, 1993; Mao, 1995) in that it examines expert versus novice decision making by groups. Figure 4-1 shows the research framework.

As discussed earlier, we are interested in evaluating the usefulness of ESS analyses and explanations support as a whole, as well as the relative usefulness of each of the two components — ESS analyses (i.e., conclusions) and ESS explanations. We are also
interested in examining whether experts would benefit as much as novices from the ESS analyses and explanations support.

<table>
<thead>
<tr>
<th>Input Variables</th>
<th>Group Process Variables</th>
<th>System Related Perceptions</th>
</tr>
</thead>
</table>
| I. Degree of Support  
  - no support  
  - ESS analyses support  
  - ESS analyses and explanations support | I. ESS Usage Characteristics  
  - uses of ESS conclusions  
  - uses of ESS explanations  
  - problems encountered without ESS support  
  - preferences for explanation types | I. Perceived Usefulness of ESS  
  II. Trust in ESS |
| II. Domain Expertise  
  - novice user  
  - expert user | II. Decision Characteristics  
  - disagreement with ESS  
  - disagreement among group members | Task Related Outcomes  
  I. Consistency with Original Experts  
  II. Attitude of Group Members toward Group Judgments  
  - satisfaction  
  - consensus |
| | | Group Related Perception  
  I. Satisfaction with Group Process |

*Figure 4-1: Research Framework*

Both quantitative and qualitative analyses were employed in this experimental research. As shown in Figure 4-1, three sets of variables were analyzed: 1) perception variables, 2) outcome variables, and 3) group process variables. Quantitative analysis of perception and outcome variables is presented in Chapters 6 and 7. The perception and outcome variables were measured using questionnaires, except for consistency with original experts and consensus with group judgments, which were computed directly from the subjects' judgments. The measurement for these perception and outcome variables is discussed in Section 5.6. The group process variables were examined in the qualitative analysis reported in Chapter 8. The group process variables are enclosed in a *dotted* rather than solid box in
Figure 4-1 because their impact on the dependent variables was not directly studied or analyzed. The qualitative analysis was carried out to provide a richer explanation and understanding of the impact of ESS analyses and explanations support on group decision making processes and outcomes. It focuses mainly on identifying the ways in which the availability of the ESS conclusions and explanation facilities influences the group process.

Outcome and Perception Variables

Performance, satisfaction, and consensus have been investigated extensively in the group support literature (Benbasat and Lim, 1993; Dennis, Haley, and Vandenberg, 1996; McLeod, 1992). In this research, performance refers to the consistency of the decision makers' judgments with the judgments of the original experts who were involved in developing FINALYZER. Using the literature from the lens model and the elaboration likelihood model, we predict that the higher the level of ESS support provided, the greater the consistency of the novices' judgments with those of the original experts. This hypothesized effect, which will be described in greater detail in Section 4.3, is consistent with the GDSS literature in that, in general, users' performance increases with GDSS use (Benbasat and Lim, 1993; Dennis, Haley, and Vandenberg, 1993; McLeod, 1992).

As explained earlier, the higher the level of ESS support provided, the greater the influence of the ESS on the novices' judgments, and therefore we would expect the judgments of the group members to be more similar. We, therefore, hypothesize that greater levels of ESS support produce a higher level of consensus in judgments. Benbasat and Lim (1993) and McLeod (1992) found GDSS use to produce negative effects on satisfaction. In this research, we are interested not only in finding out the effects of ESS support in itself, but also in determining whether ESS support would produce similar effects as those reported in the GDSS literature.
According to the *heightened criticality* hypothesis of experts, experts are expected to be more critical than novices towards the ESS conclusions and explanations and, hence, are less influenced by the ESS conclusions and explanations than the novices. Thus, the judgments of the experts are expected to be further away from the judgments of the original experts than the judgments of the novices are. The *heightened criticality* hypothesis of experts also applies among the experts. In other words, experts are not only critical with the recommendations of the ESS but also among themselves. Thus, the level of consensus among experts is expected to be lower than the level of consensus among novices.

Perceived usefulness is an important variable in MIS research as it strongly influences users' intention to use a system (Davis, Bagozzi, and Warshaw, 1989) which in turn determines actual usage of the system (Ajzen and Fishbein, 1980; Davis, Bagozzi, and Warshaw, 1989). Trust has been recognized by many (e.g., Byrd, 1992; Hayes-Roth and Jacobstein, 1994; Lerch, Prietula, and Kim, 1993; Mao, 1995; Scheier, 1996) as an important and necessary attribute of expert systems. Hayes-Roth and Jacobstein (1994) claim that "expert systems have won the acceptance and trust of end users, partially because of their ability to provide explanations". However, they did not provide support for their claim. Ye and Johnson (1995) have empirically demonstrated that explanation facilities increase users' acceptance of system's advice. Other researchers, such as Lamberti and Wallace (1990) and Oz, Fedorowicz, and Stapleton (1993), have evaluated confidence which is another related construct of trust. Mao (1995) found feedback explanations to be a stronger determinant of trust than feedforward explanations. Since the intention of providing explanation facilities is to increase users' trust in the system's advice, we are interested in testing if the explanation facilities serve its intended purpose in the group context.
Group Process Variables

Qualitative analysis of the group process variables strengthens our understanding by providing explanations on the “why” and the “how” of the relationships between the independent and dependent (outcome and perception) variables. The richer insights and understanding gained from qualitative analysis helps us to explain phenomena occurring in the groups. For instance, we are interested in finding out from the qualitative analysis if the amount of disagreement with the advice and analyses given by the ESS is higher among the experts or the novices. Such information may help us to explain how consensus is reached in the group setting or why true consensus is not reached. From the qualitative analysis, the preferences for use of explanation types by experts and novices and a comparison of the total number of explanations accessed by experts versus novices were also assessed.

To better understand how ESS analyses and explanations have benefited or hindered group decision making, we identify, from the qualitative analysis, 1) why, when, and under what circumstances the groups accessed explanations, 2) how the groups that were not provided with ESS explanations reacted to the absence of explanations supporting ESS conclusions, and 3) supporting evidence to explain whether or not the ESS analyses were helpful and in what ways they influenced the group processes. In addition, any observable group process gains and losses, and any other interesting or unique observations are presented and discussed.

Disagreement with ESS is expected to be lower for the experts than the novices because of experts’ heightened criticality characteristic. On the other hand, since ESS provides analyses and explanations support for group decision making, we expect it to produce positive outcomes in the form of process gains and informational influence. The experts’ and novices’ preferences for use of explanation types were assessed to find out if they are potential contributing factors to any observed difference in the dependent variables.
between experts and novices. For instance, if the experts’ versus novices’ preferences for use of explanation types do not differ, they can be ruled out as a possible explanation for any differences in the outcome and perception variables between the experts and novices.

### 4.2 Research Design

Two groups of subjects participated in this study — domain experts and novices. As mentioned in an earlier chapter, both the domain experts and novices in this study have sufficient basic domain knowledge to carry out the task. The domain experts and novices differ mainly in their working experience. (Note: Section 5.3 discusses the subjects’ characteristics.)

A group of three was adopted in this study for two reasons: 1) the optimal group size for decision making is between three to five (Shaw, 1981), and 2) to maximize the number of groups that can be formed from the limited number of subjects who were available for this study. This task involves financial statement analysis and commercial loan decision making. Therefore, special skills, i.e., skills related to advanced accounting and financial statement analysis, are required to carry out the task. Thus, to increase generalizability of the study, only subjects who understood accounting were recruited. This criterion severely restricts the number of available subjects for the study.

**Recruitment of Expert and Novice Subjects**

The experts are professional credit loan officers working in three major financial institutions in the western part of Canada. The novices were recruited from final year undergraduate Commerce students majoring in Accounting as well as undergraduate and graduate students taking the Financial Statement Analysis course in the Faculty of Commerce, University of British Columbia.
The recruitment of novice subjects was handled separately from the recruitment of expert subjects. Those invited to participate in the study were final year undergraduate accounting students (approximately 120 students) and the undergraduate and graduate Commerce students who were taking the financial statement analysis course (approximately 25 undergraduate and 60 graduate students). At the beginning of the class, an information sheet was distributed to each of the prospective subjects (refer to Appendix C for the information sheet). One of the researchers made a short five minute presentation to introduce the students to the study and highlight the benefits of participating in the study. The novice subjects were promised $25 upon the completion of their participation which was estimated to take 2-3 hours. The top 20% performing groups in each experimental condition received additional cash awards of $90 per group or $30 per individual in the group. A total of 75 novice subjects took part in the study.

Several problems were encountered in the recruitment of expert subjects. Many available professionals who qualify as experts for this study have taken part in earlier studies carried out by Dhaliwal (1993) and Mao (1995), thus restricting the number of expert subjects available for this study. Dhaliwal (1993) recruited expert subjects through the distribution of information packages to members of the Society of Financial Analysts, the Financial Executives Institute, and various financial and lending institutions in Vancouver. Mao (1996) recruited expert subjects through the distribution of information packages to members of the Vancouver Society of Financial Analysts (VSFA) and all the Canadian Certified General Accountants (CGAs) in the province of British Columbia.

Every effort was made to recruit expert subjects. The Dean of the Commerce Faculty identified four financial institutions that have maintained regular contact with our Faculty and requested their participation in this study. The presidents of these four financial institutions were contacted and an initial meeting was set up with the vice president of the
commercial credit loan department of each of these institutions. Through the initial meetings, 21 professional experts were recruited from three of these institutions to participate in the study – 9 from each of two institutions and 3 from the other. The fourth institution was not interested in participating in the study. In order to attract as many experts as possible to participate in the study, we tried to minimize inconvenience and disruptions to the expert subjects by setting up the study at their institutions.

With only 21 expert subjects available, there would not be enough statistical power to study the effects of different levels of ESS support on the experts. Since the main focus of this study is on ESS explanation facilities and one of our research objectives is to evaluate the usefulness of ESS analyses plus explanations support for domain expert versus novice users, all of the 21 available expert subjects (7 groups of 3) were assigned to the ESS analyses plus explanations (or full ESS) support condition to maximize the power for statistical comparisons. This enables us to make statistical comparison between domain experts and novices. Based on the rule-of-thumb, the recommended minimum number of units per cell for experimental studies is 10 although 5 is acceptable when non-parametric statistics are applied (Siegel and Castellan, 1988). Due to the constraint on the number of available expert subjects, the final research design is an unbalanced one as illustrated in Table 4-1.

<table>
<thead>
<tr>
<th>Expert Support by User Expertise</th>
<th>No ESS Support</th>
<th>ESS Analyses Support without Explanations</th>
<th>ESS Analyses Support with Explanations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Novice (Students)</td>
<td>Cell 1</td>
<td>Cell 2</td>
<td>Cell 3</td>
</tr>
<tr>
<td>Expert (Professionals)</td>
<td>—</td>
<td>—</td>
<td>Cell 4</td>
</tr>
</tbody>
</table>

*Table 4-1: Research Design*
4.3 Derivation of Hypotheses on Effects of ESS Support

The focus of this research is on group (rather than individual) decision making. Although this research analyzes the judgment data at both the individual and group levels, the focus is in evaluating the effects of ESS analyses and explanations support on these judgments in the group decision making context. The perception measures were, however, captured at the individual rather than the group level. (Hypotheses dealing with group judgments have subscripts starting with g, while those dealing with individual judgments have subscript beginning with i. Hypotheses dealing with perceptions have subscripts starting with p.)

4.3.1 Effects of ESS Analyses Support on Judgments and Consensus

When ESS provides analyses or advice about the criterion (Y_e), it exerts expert power (French and Raven, 1959) over the decision makers, thus increasing the consistency of their group judgments with the judgments of the original experts who were involved in developing the system. Expert power refers to the influence of group members accepting the truth of facts provided by the source (in this case, the ESS) based on the credibility or expertness of the source that delivers it. The concept of expert power is supported by findings from the persuasion literature which indicate that persuasive impact of communicator characteristics, such as expertise, is mediated most typically by differential acceptance of the recommendations contained in the messages (Hovland, Janis, and Kelley, 1953; Hovland and Weiss, 1951). Thus, we hypothesize that the ESS analyses support, by itself, will influence the group judgments toward those of the original experts involved in developing the system:

H_{g:1a}: ESS analyses support increases the consistency of group judgments with the judgments of the original experts who were involved in developing the system.
Furthermore, the members’ level of consensus with their group judgments is expected to increase as they are likely to trust and be influenced by the ESS analyses and advice (due to expert power or the perceived credibility of the ESS, as discussed earlier), thus converging toward the recommendations given by the ESS. In this way, the ESS analyses support helps in the conflict resolution process by indirectly resolving and reducing the number of conflicts in the group. Therefore, providing ESS advice to a group of decision makers is expected to increase the members’ consensus with their group judgments:

\[ H_{g2a} \]: \textit{ESS analyses support increases members’ consensus with group judgments.} \\

In some context, the individuals involved in the group decision making process may be asked to make or recommend their final individual judgments (based on their prior knowledge and any additional information and knowledge acquired or learned during the discussion). Even in situations where group members do not explicitly make final individual judgments, the preferred judgments of the individuals still exist. As such, a secondary research focus investigates if the hypothesized effects of ESS analyses are carried over to the individual judgments. Thus:

\[ H_{i1a} \]: Providing groups with \textit{ESS analyses support} increases the consistency of \textit{individual} judgments with the judgments of the original experts who were involved in developing the system.

Similar to the discussion on consensus with group judgments, the consensus among members’ individual judgments is also expected to increase as they are likely to trust and be influenced by the ESS analyses and advice toward the same direction:

\[ H_{i2a} \]: \textit{ESS analyses support increases members’ consensus with \textit{individual} judgments.}
4.3.2 Effects of ESS Explanations Support on Judgments and Consensus

When ESS provides explanations, a more complete set of task information becomes available (refer to Figure 3-6 for a diagrammatic representation of the additional task support). ESS explanations provide task information on the relations between the cues and the criterion, and on the relations among the cues. According to the “multiple cue probability learning” literature, providing such task support improves the mapping of the cognitive system to the task system (Balzer, Doherty, and O'Connor, 1989). In other words, we hypothesize that ESS explanations support, which is a form of task support, influences the group judgments of the decision makers in favor of the judgments of the original experts who were involved in developing the system. This is consistent with findings from the persuasion literature that the provision of supporting arguments by an expert would increase the persuasive impact of a message (Petty and Cacioppo, 1981a). These explanations exert expert power over the decision makers, increase the comprehensibility of the analyses given by the ESS, and enhance decision makers’ “cognitive processing and responding” of the ESS analyses and explanations, thus providing a greater opportunity for the group members to be persuaded by the system (Eagly and Chaiken, 1993; Perloff, 1993). According to Greenwald's cognitive response approach, the “cognitive responses” that recipients generate — and, thus, rehearse and learn — as they receive and reflect upon persuasive communications play the mediating role in persuasion. According to McGuire’s information processing paradigm, persuasion increases when one comprehends the message (in this case, ESS conclusions and explanations). By providing explanations on these analyses, it prompts users to mentally rehearse the arguments of the advice and the advice itself, thus establishing a link between the cues and the advice if one does not already exist. It is, therefore, hypothesized that:
$H_{g1b}$: *ESS explanations support* increases the consistency of *group* judgments with the judgments of the original experts who were involved in developing the system.

In addition, consensus with the group judgments is likely to increase because the set of *reasons* provided by the ESS for its analyses and advice not only increases the group members' joint understanding of the advice, but also serves as a common frame of reference for reconciling the differences among the multiple judgments of the individual group members. Thus:

$H_{g2b}$: *ESS explanations support* increases members' consensus with *group* judgments.

Similarly, we are also interested in investigating if the hypothesized effects of ESS explanations are carried over to the individual judgments. Thus:

$H_{t1b}$: Providing groups with *ESS explanations support* increases the consistency of *individual* judgments with the judgments of the original experts who were involved in developing the system.

Similar to the discussion on consensus with group judgments, the consensus among members' individual judgments is also expected to increase:

$H_{t2b}$: *ESS explanations support* increases members' consensus with *individual* judgments.

### 4.3.3 Effects of ESS Analyses and Explanations Support on Perceptions

With increasing level of ESS support, group members are likely to perceive improvement in their group judgments, thus increasing their satisfaction with the group judgments.
H_{p1}: The greater the level of ESS support provided, the greater the satisfaction with group judgments.

Since ESS is a form of level 2 GDSS, we expect its effect on satisfaction with group decision making process to be similar to that found in GDSS research. Increasing the level of ESS support may create confusion among group members in the way the system is being appropriated (Gallupe, DeSanctis, and Dickson, 1988). Group members may be unsure of how to appropriately integrate the system into their group decision making process. Furthermore, the additional influence or expert power exerted by the ESS may cause group members to perceive themselves to be less important in, and to contribute less to, the decision making process. For these reasons, increased levels of ESS support are hypothesized to decrease group members' satisfaction with the group process. This is consistent with the findings in the GDSS literature (Benbasat and Lim, 1993; McLeod, 1992; Gallupe, DeSanctis, and Dickson, 1988). Thus:

H_{p2}: The greater the level of ESS support provided, the lower the satisfaction with group decision making process.

The persuasion theories and the “multiple cue probability learning” literature predict that explanation facilities will lead to greater knowledge transfer from the ESS to the users (refer to the first paragraph in Section 4.2.2 for the discussion). The explanation facilities provide justifications to assure users that the system's reasoning is sound and appropriate to the task at hand. They increase users' understanding of the ESS conclusions and advice, and the task itself. Thus, the explanation facilities are hypothesized to increase users' perceived usefulness of the ESS.
$H_{p3}$: The explanation facilities increase users' *perceived usefulness* of the ESS.

Users' trust in an ESS is believed to depend on the degree to which the technical competence of the ESS is made apparent to its users. This function is typically achieved through the provision of the explanation facilities. More specifically, the explanation facilities are likely to increase users' trust in the ESS because they increase users' perceptions of the level of expert knowledge and technical capability demonstrated by the ESS.

$H_{p4}$: The explanation facilities increase users' *trust* in the ESS.

### 4.4 Derivation of Hypotheses on Expert-Novice Differences

This section derives the hypotheses relating to the effects of using ESS analyses and explanations support for group decision making by experts and novices. The initial (before system-use) and final (after system-use) judgments of the experts and novices were evaluated with respect to the *consensus judgments* arrived by the group of experts who were involved in developing FINALYZER. The perceptions of the experts and novices were also captured through the use of questionnaires.

#### 4.4.1 Effects of ESS Analyses and Explanations Support on Experts-Novice Judgments

The experts in this study are professional financial analysts whose major responsibilities include making commercial loan decisions on a daily or frequent basis. The novices are final year undergraduate students majoring in Accounting, and graduate or final year undergraduate students who have taken the Financial Statements Analysis course offered in the Faculty of Commerce, University of British Columbia. Thus, both the novice and expert subjects have the declarative knowledge to solve the experimental task on Financial Statements Analysis. However, since the experts are more experienced in performing the
task, we expect them to have more refined procedural knowledge than the novices. Since there are no unique solutions to the problems, we use the consensus judgments of the original experts (who were involved in the development of FINALYZER) as the yardstick to evaluate the experts’ and novices’ judgments.

Although both the experts and novices have the basic declarative knowledge to perform the task, we hypothesize that the experts would produce judgments that are closer to those of the original experts involved in developing the system because they possess more refined procedural knowledge and better capabilities to identify relationships in the task domain:

$H_{3a}$: The initial judgments of the experts (that are made without any form of ESS support) are more similar to the judgments of the original experts than it is for the initial judgments of the novices.

As the experts in our study are professional financial analysts whose skills are highly valued in the financial industry, we expect them to be highly ego-involved in their area of specialization (Sherif and Cantril, 1947) and highly critical in carrying out financial statements analysis. This is in line with the heightened criticality hypothesis (Biek and Wood, 1996) and the social judgment theory (Sherif and Hovland, 1953, 1961) which predict that people with high knowledge and experience not only have greater ability to critically evaluate even strongly argued messages, but also have a stronger tendency to reject or oppose them (Biek and Wood, 1996; Johnson and Eagly, 1989; Sherif and Hovland, 1961; Sherif and Sherif, 1967). We, therefore, hypothesize that the experts will be less willing than the novices to accept the analyses and advice given by the ESS. In other words, we expect the experts’ judgments to be farther away from the judgments of the original experts (involved in developing FINALYZER) than it is for the novices’ judgments:
H₃: The *group* judgments of the experts are farther away from the judgments of the original experts than it is for the *group* judgments of the novices.

We are also interested in evaluating whether the above hypothesis will hold when the level of analysis is individual. Since the theories discussed earlier are also applicable at the individual context, we would expect the same result to hold for individual judgments:

H₁₀: The final *individual* judgments of the experts are farther away from the judgments of the original experts than it is for the final *individual* judgments of the novices.

As discussed earlier, the experts are, in general, more critical and less likely to agree with others’ judgments (Biek and Wood, 1996; Johnson and Eagly, 1989). Although group judgments are derived from consensus decision making, we expect the true consensus among the experts to be lower than that among the novices:

H₄: Experts’ consensus with *group* judgments is lower than novices’ consensus with *group* judgments.

H₄: Experts’ consensus among *individual* judgments is lower than novices’ consensus among *individual* judgments.

4.4.2 *Effects of ESS Analyses and Explanations Support on Expert-Novice Perceptions*

Since the ESS analyses and explanations are likely to be of greater help to, and exert greater influence on, the novices than the experts, we expect the novices to perceive greater improvement in the quality of their group judgments, and to be, therefore, more satisfied with their group judgments:
Hₚ₅: The experts are less satisfied with the group judgments than the novices.

Since the experts are experienced in financial analysis, they may not perceive the ESS support to be helping them in making the judgments. On the other hand, the novices lack experience and, thus, are more likely to find the ESS support useful. Therefore, the ESS support may cause experts to be less satisfied with the group decision making process because it is being viewed as unnecessary or interfering. In contrast, novices are more likely to accept the ESS support than the experts. Thus:

Hₚ₆: The experts are less satisfied with the group decision making process than the novices.

As discussed earlier, we expect the ESS analyses and explanations to exert greater influence on the novices than the experts. The novices, therefore, perceive not only greater improvement in the quality of their group judgments, but also a higher level of usefulness of ESS than the experts:

Hₚ₇: The experts’ perceived usefulness of ESS is lower than that of the novices.

Since experts not only have greater ability to critically evaluate the analyses and explanations in the ESS but also have a stronger tendency to reject or oppose them (Biek and Wood, 1996; Johnson and Eagly, 1989; Sherif and Hovland, 1961; Sherif and Sherif, 1967), we hypothesize that the experts will have lesser trust in the ESS than the novices have.

Hₚ₈: The experts’ trust in ESS is lower than that of the novices.
4.5 Summary of Chapter 4

This chapter presents the research framework and draws upon the persuasion theories and the lens model or “multiple cue probability learning” literature to derive the hypotheses for this research. From these theories, it is hypothesized that consensus in judgments and knowledge transfer from the ESS to the users increase with the level of ESS support. We hypothesize that higher levels of ESS support will increase decision makers’ satisfaction with group judgments but lower their satisfaction with the group decision making process. We also expect the availability and use of explanation facilities to increase both users’ trust in the system and their perceived usefulness of the system. Without any form of ESS support, the judgments of the experts and the judgments of the original experts who were involved in developing the ESS are expected to be more similar than it is between the judgments of the novices and the original experts. However, with ESS analyses and explanations support, the judgments of the experts are expected to be farther away from the judgments of the original experts than it is for the judgments of the novices. The level of consensus in judgments among the experts is expected to be lower than that among the novices. It is hypothesized that, when working with ESS analyses and explanations support, experts are less satisfied than novices with their group judgments and decision making processes. Furthermore, novices perceive the ESS to be more useful and have greater trust in the ESS than experts do.
CHAPTER 5: RESEARCH METHODOLOGY

This chapter describes the methodology used for the experiment that was conducted to examine the effects of ESS analyses and explanations support on group decision making processes and outcomes. The research framework and design were discussed in Chapter 4. This chapter describes the research procedures, the subject characteristics, the experimental task as well as the ESS that was used in this study. The operationalization of the dependent variables is also discussed.

The experimental study was carried out with the novice and expert subjects during the period from October 1995 to May 1996. A total of six pilots were completed during the period of March 1995 to August 1995, with 3 groups in the full (ESS analyses and explanations) support treatment (cell 1 in Table 4-1), 1 group in the partial (ESS analyses without explanations) support treatment (cell 2 in Table 4-1), and 2 groups in control (no ESS support) group (cell 3 in Table 4-1). These pilots were carried out before conducting the actual experiment to thoroughly test the experimental procedures, materials, and task, the instruments in the questionnaire, the equipment setup, as well as the ESS, FINALYZER.

5.1 Experimental Procedures

The experiment comprises three phases: pre-experiment, experiment, and post-experiment. During the pre-experiment phase, the subjects filled out a consent form and a background questionnaire (see Appendix A). Figure 5-1 shows the research procedure in the experiment phase. During this phase, subjects were given the relevant financial information of the company to be evaluated. They then assessed the case individually and without any form of ESS support (see Appendix A for the judgment sheet). Following that, subjects in the ESS-supported groups (cells 2-4 in Table 4-1) received the appropriate training to
familiarize them with the system features. Another system, known as the CREDIT-ADVISOR system, was used for the training. The CREDIT-ADVISOR evaluates consumer credit applications. The subjects were told that the FINALYZER system that they would be using later had an interface similar to the CREDIT-ADVISOR system. The subjects in the control group (cell 1 in Table 4-1) received no training. Next, the subjects worked in their groups of three under the assigned experimental condition until a group consensus was reached on the same set of judgments they had made earlier. Finally, they were asked to make the same set of judgments again individually taking into account what they had learned from FINALYZER and their group discussions. These post-discussion individual judgments allow us to evaluate consensus achieved in the groups. Thus, a total of three sets of judgments were made by each subject (pre-discussion individual, group, post-discussion individual judgments) during the experiment phase.

Figure 5-1: Research Procedure
Basic aids — calculators, papers, and pens — were provided to all groups. Only one ESS was provided to each ESS-supported group. If the ESS could be directly accessed by the groups, one or more group members may dominate the process by controlling the mouse all or most of the time. In order not to incidentally introduce dominant members into the groups, the researcher, who took the role of the chauffeur during the experiment, carried out the groups' requests to access the system and its explanations (see experimental setup in Figure 5-2).

![Experimental Setup](image)

*Figure 5-2: Experimental Setup*

The researcher had two major responsibilities during the experiment: 1) to carry out the experimental procedures (including training) and to ensure that the experiment progressed according to the procedures, and 2) to operate the mouse upon the subjects' requests (chauffeur’s role) without participating in the group decision making process. Two monitors were used in the experiment, one for the subjects or decision makers, and the other for the chauffeur. The two monitors were connected by a synchronizer so that whatever appeared in the screen of one monitor would appear in the other. However, only one mouse was available for operating the system which was handled by the chauffeur. The
group discussions were transcribed from the audio- and video-recordings. Both the transcripts and the video-recordings were used to analyze the data.

Finally, during the post-experiment phase, subjects filled out questionnaires and were debriefed about the experiment. If time permitted, interviews with the subjects were conducted.

5.2 Group Effect and ESS Analyses and Explanations Support Effect

Referring back to Figure 5-1, the group effect (i.e., effect due to group discussion) accounts for the difference between the pre-discussion judgments and the group judgments in the control group. That is, for the control group, the ESS support effect was neither present nor relevant. On the other hand, for the partial support (i.e., ESS analyses support only) treatment, the difference between the pre-discussion judgments and the group judgments arises from the aggregate effect of group and ESS analyses support. In other words, to determine if the ESS analyses support effect is significant, this difference (between the pre-discussion judgments and the group judgments in the partial group) will be compared with the group effect, which is determined by the difference between the pre-discussion judgments and the group judgments in the control group. Similarly, to determine the effect due to ESS explanations support only, the difference that arises between the pre-discussion judgments and group judgments (i.e., due to group and ESS analyses support effect) of the partial support treatment will be compared with the corresponding difference (due to group and ESS analyses plus explanations support) in the full support treatment. Lastly, to determine the aggregate effect of ESS analyses and explanations support, the group effect will be compared with the difference between the pre-discussion judgments and group judgments of the full support (ESS analyses plus explanations support) treatment.
5.3 Subject Characteristics

To ensure that the subjects satisfied our selection criterion, we captured the subjects' self-rating of their competence as a financial analyst of corporate loan decisions. This serves as a validation check. One of the questions in the background information questionnaire is “How do you rate yourself as a financial analyst (of a corporate loan decision)?” (1—Excellent, 2—Good, 3—Somewhat good, 4—Fair, 5—Somewhat poor, 6—Poor, 7—Bad). The average rating of the novice subjects is 3.8 and that of the expert subjects is 1.9, and the difference is statistically significant at $p<.01$. This supports our operationalization of the expertise construct. For the experts, the average number of years of financial analysis related working experience is 13.3 years.

5.4 Experimental Task

A commercial loan decision task was chosen for this study. Not only is this a realistic task, but the composition of financial statements and the expertise required to evaluate them are also sufficiently complex to justify the use of ESS support. The task involves evaluating the financial position, performance, and potential of a company, and determining an appropriate loan amount. Financial statement analysis usually entails the review of a company's financial data to evaluate various aspects of its financial standing and performance. It is conducted by comparing a firm's financial ratios to the same ratios in earlier years, and to the ratios of other firms in the same industry, often summarized into industry composites. The use of ratios to derive judgments in financial analysis is an unstructured process, characterized by the use of specialized domain knowledge. As this is an unstructured task where no computerized decision procedure or algorithm exists to determine an appropriate loan amount, rule-based procedures for evaluating a company's financial position, performance, and potential can be implemented into the ESS as indirect support for the task. For instance, the ESS may prompt its users to certain problematic areas or concerns that are reflected in the financial statements of the company, as well as
highlight favorable aspects where the company has performed well. In this case, the decision makers used the support provided by FINALYZER to carry out an evaluation of various financial aspects of the case followed by the determination of an appropriate loan amount.

A financial analysis case was prepared which involved the evaluation of an application for senior borrowing by a hypothetical firm. Subjects were told to assume the role of corporate loan evaluation officers working for a large financial institution in Western Canada. They were provided with five years of financial statements of the hypothetical firm — "Canacom" — and a complete set of common-size statements and financial ratios. The financial statements and case description were described and used in previous studies by Dhaliwal (1993) and Mao (1995). The company was applying for a senior borrowing of $800 million for streamlining its operations. Subjects were asked to use an ESS designed for loan evaluation to assess various aspects of the company's financial health. Then, based on the assessment, they would make a recommendation regarding whether the loan should be approved and, if yes, the amount.

The financial statements and descriptions of the case were revised in scale to a loan request of $800 thousand to make it in line with the cases that are handled by the professionals in our sample on a daily basis. This was done to maintain experimental realism (Swieringa and Weick, 1982) which is particularly important for practitioner subjects. Two accounting professors and one MIS professor, who has a background in accounting, were consulted to ensure that the case was appropriately revised in scale and were equivalent, in such a way that the performance of experts and novices would be comparable. Since the ratios in the case remained unchanged and the task concerned risk (or ratio) analysis, the change in scale should not affect the subjects' ratings of the various financial aspects of the company.
A total of eight judgments were to be made by the subjects (refer to Appendix A for the judgment sheet). Questions 1-6 required the subjects to rate on a scale of 1-10 the liquidity, long-term solvency, asset utilization, the value of stock as collateral, and the quality of financial and operating management of the company, respectively. Question 7 asked for an estimate of the predicted net income of the company for the coming year and question 8 concerned the amount of loan to be granted. Questions 1 to 6 were used to evaluate the consistency of the subjects’ judgments with those of the original experts who were involved in developing FINALYZER. Questions 7 and 8 were not included in the evaluation because they are subjective judgments.

5.5 Experimental ESS
The ESS, named FINALYZER was used, tested, revised, and validated in two earlier studies (Dhaliwal, 1993; Mao, 1995). Before describing FINALYZER, we will examine its explanation facilities.

5.5.1 Explanation Facilities of FINALYZER
Two main categories of explanations are available in FINALYZER — feedforward and feedback. Feedforward and feedback explanations provided by FINALYZER were designed based on the cognitive feedforward and feedback paradigm for learning in the context of problem solving, which emphasizes a particular order among events (Bjorkman, 1972). The concept of feedforward and feedback explanations was introduced into the expert systems literature by Dhaliwal and Benbasat (1996) based on the two learning operators from the cognitive learning perspective, cognitive feedback and cognitive feedforward (Bjorkman, 1972). Feedforward knowledge was presented as cognitive feedforward prior to analysis, and feedback explanations were accessible as cognitive feedback after the system has presented its analyses and advice.
Dhaliwal and Benbasat (1996) also integrate the Why, How, and Strategic explanation types (Buchanan and Shortliffe, 1984; Hasling, Clancey, and Rennels, 1984) with the concept of feedforward and feedback explanation provision strategies. Why, How, and Strategic explanations can be presented both as feedforward and feedback, giving six (2x3) explanation types whose definitions are presented in Table 5-1. These six types of explanations were adopted for this study. The correspondence between these six types of explanations and the three approaches to explanations reviewed in Chapter 2 is as follows:

1. **Reasoning trace explanations**: Feedback WHY and HOW;
2. **Deep explanations**: Feedforward WHY and HOW;
3. **Strategic explanations**: Feedforward STRATEGIC and Feedback STRATEGIC.

| Feedforward Why explanations justify the importance of, and the need for, input information to be used. |
| Feedforward How explanations detail the manner in which input information is derived for use. |
| Feedforward Strategic explanations clarify the overall manner in which input information to be used is organized or structured, and specify the manner in which each input cue to be used fits into the overall plan of assessment that is to be performed. |
| Feedback Why explanations justify the importance, and clarify the implications, of a particular conclusion that is reached by the system. |
| Feedback How explanations present a trace of the evaluations performed and intermediate inferences made in getting to a particular conclusion. |
| Feedback Strategic explanations clarify the overall goal structure used by a system to reach a particular conclusion, and specify the manner in which each particular assessment leading to the conclusion fits into the overall plan of assessments that were performed. |

Table 5-1: Definition of Explanation Types (Adapted from Dhaliwal and Benbasat, 1996)

### 5.5.2 Description of FINALYZER

FINALYZER is a simulated system that provides five sub-analyses — funds flow analysis, liquidity analysis, capital structure analysis, profitability analysis, and market value
analysis. Subjects were asked to use FINALYZER by going through all of the five sub-analyses to support their group judgments. For each of the five sub-analyses, FINALYZER provides three basic types of screens, in the order of information screen, data screen, and conclusion screen (see Figure 5-3 for a flow chart of FINALYZER and Appendix B for examples of these screens.) An information screen contains an index of relevant domain concepts (financial terms or ratios) and procedures to be used as inputs to the current subanalysis. A data screen contains the relevant financial ratios calculated from the financial statements of the firm to be evaluated. A recommendation screen presents results of FINALYZER's "evaluation" of the financial statements and ratios.

Upon selecting a sub-analysis, feedforward explanations are made available through hypertext links provided on the information and data screens (refer to Appendix B for examples of these screens). Feedforward explanations are independent of the case or context. They not only explain the relationships among the inputs that will be used by the system to carry out its operations, but also relate these inputs to the task. For example, feedforward explanations on "current ratio" can be accessed from FINALYZER. Next, conclusions or analyses/advice on the sub-analysis are displayed on the recommendation screens. Feedforward and feedback explanations on these conclusions are made available.
to the users upon request. Feedback explanations are case-specific. They explain the conclusions or recommendations reached by the system. A recommendation given by FINALYZER could be “Canacom is in a very favorable working capital position, indicating little risk in the short term of financial disaster. It would be an optimal client for a short term loan”. Both feedforward and feedback explanations in FINALYZER are provided via hypertext (Mao, 1995). For example, users may seek the feedback explanation of the above recommendation or the feedforward explanation of “working capital” in the recommendation directly from the recommendation screen. In addition, feedforward explanations (called deep explanations by Mao, 1995) are inter-linked, i.e., they can be accessed from one to another. They are also accessible from feedback explanations (note the hypertext linkages in Figure 5-3). Finally, a summary of the overall analysis will be presented to the users before they exit the system.

The five sub-analyses provide feedforward explanations on 42 domain concepts (i.e., financial terms and ratios) and 18 conclusions. Each domain concept has one HOW and one WHY explanation. There is only one STRATEGIC explanation for each sub-analysis, common for all domain concepts involved in the sub-analysis. Therefore, the total number of feedforward explanations is 89 (42x2+5). Similarly, each conclusion has one HOW and one WHY explanation, and a common STRATEGIC explanation for all the conclusions of that sub-analysis. The total number of feedback explanations is thus 41 (18x2+5). (To view examples of feedforward and feedback explanations, please refer to Appendix B.)

5.6 Dependent Variables

Both quantitative and qualitative data were gathered and analyzed. Quantitative analyses were carried out using data collected from the judgment recording sheets, post-study questionnaires, and computer logs of the experimental sessions. Case and interpretive
analyses were carried out by analyzing conversations and interactions among the group members. Table 5-2 lists the dependent variables in the quantitative analysis.

<table>
<thead>
<tr>
<th>Quantitative Analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. consistency with judgments of original experts</td>
</tr>
<tr>
<td>2. actual and perceived consensus in decision making</td>
</tr>
<tr>
<td>3. satisfaction with group decision making process</td>
</tr>
<tr>
<td>4. satisfaction with group judgments</td>
</tr>
<tr>
<td>5. trust in ESS</td>
</tr>
<tr>
<td>6. perceived usefulness of ESS</td>
</tr>
</tbody>
</table>

*Table 5-2: Dependent Variables in Quantitative Analysis*

5.6.1 Quantitative Analysis

By “consistency with judgments of original experts”, we mean the degree of similarity between the subjects’ judgments and the judgments of the original experts who were involved in developing FINALYZER. Thus, it can be computed directly from the subjects’ judgments. Similarly, the actual consensus measure can also be computed directly from the subjects’ judgments. Perceived consensus (with group judgments), satisfaction with group process, satisfaction with group judgments, trust in ESS, and perceived usefulness of ESS are perception measures that were captured through the use of questionnaires.

5.6.1.1 Consistency with Original Experts

As mentioned earlier, performance or consistency with original experts was assessed using the subjects' ratings on a scale of 1-10 for the liquidity, long-term solvency, asset utilization, the value of stock as collateral, and the quality of financial and operating management of the company. The most prevalent definition of consistency with original experts is the correspondence between the judgment (of decision maker(s)) and the criterion (consensus judgments of original experts) in the lens model. The subjects'
judgments were assessed in relation to a set of expert consensus estimates agreed upon by a panel of five expert judges who were involved in developing FINALYZER (Dhaliwal, 1993). Consistency with original experts was assessed by the sum of the absolute difference between the subjects' and original experts' judgments. As an example, with the original experts' consensus judgments given in Table 5-3, the maximum numerical value that can be obtained for the total absolute deviation of a set of judgments from the original experts' consensus judgments is: 7+7+5+5+6+6=36.

For example, if the following judgments were made in the experiment phase:

<table>
<thead>
<tr>
<th></th>
<th>Pre-discussion judgments</th>
<th>Group's Consensus</th>
<th>Post-discussion Judgments</th>
<th>Original Experts' Consensus</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q1</td>
<td>S1</td>
<td>6</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>S2</td>
<td>7</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>S3</td>
<td>7</td>
<td>6</td>
<td>8</td>
</tr>
<tr>
<td>Q4</td>
<td></td>
<td>9</td>
<td>5</td>
<td>7</td>
</tr>
<tr>
<td>Q5</td>
<td></td>
<td>8</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Q6</td>
<td></td>
<td>8</td>
<td>7</td>
<td>4</td>
</tr>
</tbody>
</table>

Table 5-3: An Example for Illustration Purposes

* — discrepancies between group and post-discussion individual judgments
Consistency with Judgments of Original Experts

<table>
<thead>
<tr>
<th>Source</th>
<th>Deviation (the smaller, the more consistent)</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1's pre-discussion individual judgments</td>
<td>2+2+2+4+1+1=12</td>
</tr>
<tr>
<td>S2's pre-discussion individual judgments</td>
<td>1+0+1+0+2+0=4</td>
</tr>
<tr>
<td>S3's pre-discussion individual judgments</td>
<td>1+0+3+2+2+3=11</td>
</tr>
<tr>
<td>Group's consensus judgments</td>
<td>0+1+1+2+1+1=6</td>
</tr>
<tr>
<td>S1's post-discussion individual judgments</td>
<td>0+1+1+3+0+1=6</td>
</tr>
<tr>
<td>S2's post-discussion individual judgments</td>
<td>0+1+1+2+1+1=6</td>
</tr>
<tr>
<td>S3's post-discussion individual judgments</td>
<td>0+1+1+2+2+3=9</td>
</tr>
</tbody>
</table>

Table 5-4: An Example to Illustrate Computation of Absolute Deviation from Original Experts' Judgments

5.6.1.2 Actual and Perceived Consensus

Group consensus can be measured in terms of 1) actual (or objective) consensus, which evaluates the degree of agreement with the group judgments and degree of agreement among the group members' post-discussion individual judgments, and 2) perceived consensus, which assesses the degree to which members believe they agree with the group's judgments.

Actual consensus was measured in two ways: 1) the sum of the difference between the post-discussion individual judgments and the group judgments; and 2) the total difference between all pair-wise post-discussion individual judgments. In the above example (Table 5-3),
Computation for Consensus with Group Judgments is:
Sum of the difference between S1’s post-discussion and group judgments = 1+1=2
Sum of the difference between S2’s post-discussion and group judgments = 0
Sum of the difference between S3’s post-discussion and group judgments = 1+2=3
*Sum of the difference between post-discussion individual and group judgments*  
(= sum of the above) = 2+0+3=5

Computation for Consensus Among Individual Judgments is:
Difference between S1’s and S2’s post-discussion individual judgments = 1+1=2
Difference between S2’s and S3’s post-discussion individual judgments = 1+2=3
Difference between S3’s and S1’s post-discussion individual judgments = 1+2+2=5
*Total difference between post-discussion individual judgments*  
(= sum of the above) = 2+3+5=10

The perceived consensus in group judgments was adapted from Wheeler (1993). This scale was captured through the post-study questionnaire administered to each group member and analyzed at the individual level. The original instrument comprises 8 six-point Likert-scale items but was subsequently reduced to 7 items. The scale's reliability was .89 (Wheeler, 1993). As most of these 7 items are variations and repetitions of one another, we have reduced the number of items to three, and administered them on 7-point Likert-scale to maintain consistency with items of other constructs. These three items are given in Table 5-5.
1. How different or similar are your final individual decisions from your group's decisions?
   
   Very different: 1 - 2 - 3 - 4 - 5 - 6 - 7 : Very similar

2. Do you disagree or agree with your group's solution?
   
   Strongly disagree: 1 - 2 - 3 - 4 - 5 - 6 - 7 : Strongly agree

3. To what extent do you oppose or support your group's solution?
   
   Strongly oppose: 1 - 2 - 3 - 4 - 5 - 6 - 7 : Strongly support

Table 5-5: Items to Measure Perceived Consensus of Group Judgments

5.6.1.3 Satisfaction with Group Process

Satisfaction with group process was measured using the instrument developed by Green and Taber (1980). It was captured by the post-study questionnaire administered to the individual group members. The original instrument comprises 5 items on a 5-point scale with a median alpha coefficient of .88. In this study, we administered the 5 items on a 7-point Likert scale to maintain consistency with items measuring other constructs (i.e., trust in ESS and perceived usefulness of ESS). These 5 items are shown in Table 5-6.

Table 5-6: Items to Measure Satisfaction with Group Process

1. How would you describe your group's problem solving process?
   
   Efficient: 1 - 2 - 3 - 4 - 5 - 6 - 7 : Inefficient
   Coordinated: 1 - 2 - 3 - 4 - 5 - 6 - 7 : Uncoordinated
   Fair: 1 - 2 - 3 - 4 - 5 - 6 - 7 : Unfair
   Confusing: 1 - 2 - 3 - 4 - 5 - 6 - 7 : Understandable
   Satisfying: 1 - 2 - 3 - 4 - 5 - 6 - 7 : Dissatisfying

5.6.1.4 Satisfaction with Group Judgments

Satisfaction with group judgments or decision outcomes was measured using the instrument developed by Green and Taber (1980). It was captured by the post-study questionnaire administered to the individual group members. The original instrument comprised 5 items on a 5-point scale with a median alpha coefficient of .88. These 5 items
were administered on a 7-point Likert scale to maintain consistency with items of other constructs (i.e., trust in ESS and perceived usefulness of ESS). These 5 items are shown in Table 5-7.

<table>
<thead>
<tr>
<th></th>
<th>How satisfied or dissatisfied are you with the quality of your group's solution?</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Very dissatisfied: 1 - 2 - 3 - 4 - 5 - 6 - 7 : Very satisfied</td>
</tr>
<tr>
<td>1.</td>
<td>To what extent does the group solution reflect your contributions?</td>
</tr>
<tr>
<td></td>
<td>Not at all: 1 - 2 - 3 - 4 - 5 - 6 - 7 : To a very great extent</td>
</tr>
<tr>
<td>2.</td>
<td>To what extent do you feel committed to your group's solution?</td>
</tr>
<tr>
<td></td>
<td>Not at all: 1 - 2 - 3 - 4 - 5 - 6 - 7 : To a very great extent</td>
</tr>
<tr>
<td>3.</td>
<td>To what extent are you confident that your group's solution is correct?</td>
</tr>
<tr>
<td></td>
<td>Not at all: 1 - 2 - 3 - 4 - 5 - 6 - 7 : To a very great extent</td>
</tr>
<tr>
<td>4.</td>
<td>To what extent do you feel personally responsible for the correctness of your group's solution?</td>
</tr>
<tr>
<td></td>
<td>Not at all: 1 - 2 - 3 - 4 - 5 - 6 - 7 : To a very great extent</td>
</tr>
</tbody>
</table>

Table 5-7: Items to Measure Satisfaction with Group Judgments

5.6.1.5 Trust in ESS

User trust in ESS was measured with an instrument used by Mao (1995). The instrument was adapted from Lerch, Prietula, and Kim (1993). The scale was slightly modified to suit the group decision making context, and captured by the post-study questionnaire administered to the individual group members. The alpha coefficient for the scale is .85. The eight items used to measure this construct are presented in Table 5-8.
1. FINALYZER provided good advice across different situations.
   Strongly disagree: 1 - 2 - 3 - 4 - 5 - 6 - 7 : Strongly agree
2. FINALYZER is dependable in important decisions.
   Strongly disagree: 1 - 2 - 3 - 4 - 5 - 6 - 7 : Strongly agree
3. When FINALYZER gave unexpected advice, my group is confident that the advice is correct.
   Strongly disagree: 1 - 2 - 3 - 4 - 5 - 6 - 7 : Strongly agree
4. FINALYZER is a reliable source of knowledge for financial analysis.
   Strongly disagree: 1 - 2 - 3 - 4 - 5 - 6 - 7 : Strongly agree
5. I think users with little expertise would trust the advice given by FINALYZER.
   Strongly disagree: 1 - 2 - 3 - 4 - 5 - 6 - 7 : Strongly agree
6. FINALYZER gave the same advice for the same situation over time.
   Strongly disagree: 1 - 2 - 3 - 4 - 5 - 6 - 7 : Strongly agree
7. FINALYZER behaved in a very consistent manner.
   Strongly disagree: 1 - 2 - 3 - 4 - 5 - 6 - 7 : Strongly agree
8. FINALYZER helped my group make good decisions.
   Strongly disagree: 1 - 2 - 3 - 4 - 5 - 6 - 7 : Strongly agree

Table 5-8: Items to Measure User Trust in ESS

5.6.1.6 Perceived Usefulness of ESS

Perceived usefulness of ESS refers to the degree to which users perceived the ESS to enhance their task performance. This scale was measured using the instrument developed by Dhaliwal (1993) which was adapted from Moore and Benbasat (1991) and Davis (1986). It was also slightly modified to suit the group decision making context. The instrument was part of the post-study questionnaire administered to individual members. The alpha coefficient for the scale is .85. Table 5-9 presents the items in this instrument.
1. The use of FINALYZER greatly enhanced the quality of my group's judgments.
   Strongly disagree: 1 - 2 - 3 - 4 - 5 - 6 - 7 : Strongly agree
2. Using FINALYZER gave my group more control over the financial analysis task.
   Strongly disagree: 1 - 2 - 3 - 4 - 5 - 6 - 7 : Strongly agree
3. Using FINALYZER made the financial analysis task carried out by my group easier to perform.
   Strongly disagree: 1 - 2 - 3 - 4 - 5 - 6 - 7 : Strongly agree
4. Using FINALYZER enabled my group to accomplish the financial analysis task more quickly.
   Strongly disagree: 1 - 2 - 3 - 4 - 5 - 6 - 7 : Strongly agree
5. Using FINALYZER improved the quality of the analysis my group performed.
   Strongly disagree: 1 - 2 - 3 - 4 - 5 - 6 - 7 : Strongly agree
6. FINALYZER supported all types of analysis needed by my group to make its decisions.
   Strongly disagree: 1 - 2 - 3 - 4 - 5 - 6 - 7 : Strongly agree
7. Using FINALYZER increased my group's productivity.
   Strongly disagree: 1 - 2 - 3 - 4 - 5 - 6 - 7 : Strongly agree
8. Overall, I found FINALYZER useful in helping my group analyze the financial statements.
   Strongly disagree: 1 - 2 - 3 - 4 - 5 - 6 - 7 : Strongly agree
9. Using FINALYZER enhanced my group's effectiveness in completing the financial analysis task.
   Strongly disagree: 1 - 2 - 3 - 4 - 5 - 6 - 7 : Strongly agree
10. Using FINALYZER allowed my group to accomplish more analysis than would otherwise have been possible.
    Strongly disagree: 1 - 2 - 3 - 4 - 5 - 6 - 7 : Strongly agree

Table 5-9: Items to Measure Perceived Usefulness of ESS

5.6.2 Reliability and Validity of Perception Measures

The task-, group-, and system-related perception measures in this study are: satisfaction with group process, satisfaction with group judgments, perceived consensus in group judgments, perceived usefulness of ESS, and trust in ESS. This section assesses the reliability and validity of these perception measures.
A multi-item instrument in the form of a post-study questionnaire was used for measuring the subjects' perceptions. The overall reliability of each of the five scales (satisfaction with group process, satisfaction with group judgments, perceived consensus in group judgments, perceived usefulness of ESS, and trust in ESS) was assessed using Cronbach's Alpha (Cronbach, 1990). The scales for "satisfaction with group process" and "satisfaction with group judgments" were previously developed and validated by Green and Taber (1980). The scale for "perceived consensus in group judgments" was adapted from Wheeler (1993). The scale for "perceived usefulness of ESS" was measured using Dhaliwal's (1993) instrument adapted from Moore and Benbasat (1991) and Davis (1986). Lastly, the scale for "trust in ESS" was measured using the instrument Mao (1995) adapted from Lerch, Prietula, and Kim (1993) and McCroskey (1985). Since the control group did not utilize the ESS, their post-study questionnaire comprised only three multi-item scales measuring the perception constructs: "satisfaction with group process", "satisfaction with group judgments", and "perceived consensus in group judgments". In other words, the scales for "perceived usefulness of ESS" and "trust in ESS" were excluded because of their irrelevance to the control group. The other two experimental groups utilized the ESS support and therefore were given the complete instrument comprising all five scales.

The reliability and validity of the instrument were assessed at two different levels. At the scale level, indicators of reliability and validity were obtained for each scale. The overall reliability of each of the five scales was assessed using Cronbach's Alpha (Cronbach, 1990). The construct validity of the scales was assessed by performing Principal Components Factor Analysis utilizing both the Varimax (orthogonal) and Direct Oblimin (oblique) rotations\(^1\) to obtain the eigenvalues and the percentage of variance explained.

\(^{1}\) I thank the Purdue University Statistical Consulting Services, especially Colin Ho, and my Ph.D. colleague, Andrew Gemino, for directing me to use the appropriate factor analysis method. Varimax rotation is the most
(Section 5.6.2.2 reports the results of the factor analyses.) At the item level, the items comprising each scale were evaluated using various item reliability statistics, including the standard deviation score, the effect on Cronbach’s Alpha if an item was deleted, the item-to-total scale correlation, as well as the rotated factor loadings. Examining item reliability statistics and rotated factor loadings help to identify items that reduced either the reliability or construct validity of the scales. The elimination of such items could enhance subsequent statistical analysis conducted to test the research hypotheses. An item reduces the scale reliability if (1) its deletion helps to improve Cronbach’s Alpha, (2) it has a low correlation to the total scale, and (3) it has a low standard deviation score. On the other hand, an item reduces the construct validity if it does not load strongly on any factor.

5.6.2.1 Reliability of Perception Measures

Tables 5-10 to 5-17 show the item reliability statistics for the five scales — “satisfaction with group process”, “satisfaction with group judgments”, “perceived consensus in group judgments”, “perceived usefulness of ESS”, and “trust in ESS”. Note that the first column of the tables gives the question number of the post-study questionnaire. A copy of the post-study questionnaire is included in Appendix A.

The reliability statistics were generated both using the entire data set (cells 1-4 in Table 4-1, comprising both the expert and novice groups) and the data from the novice groups only (cells 1-3 in Table 4-1). For the scales measuring “satisfaction with group process”, “satisfaction with group judgments”, and “perceived consensus in group judgments”, the reliability statistics generated in both cases differ little, indicating that the inclusion of the expert group (cell 4 in Table 4-1) into the analyses does not distort the overall reliability.
statistics. As such, Tables 5-10 to 5-12 present the reliability statistics for the three scales generated using the entire data set.

<table>
<thead>
<tr>
<th>Qn. No.</th>
<th>Standard Deviation</th>
<th>Scale Mean if Item Deleted</th>
<th>Scale Variance if Item Deleted</th>
<th>Item-to-Total Correlation</th>
<th>Alpha if Item Deleted</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q1a</td>
<td>1.28</td>
<td>10.65</td>
<td>16.43</td>
<td>.67</td>
<td>.83</td>
</tr>
<tr>
<td>Q1b</td>
<td>1.26</td>
<td>10.65</td>
<td>16.51</td>
<td>.67</td>
<td>.83</td>
</tr>
<tr>
<td>Q1c</td>
<td>1.16</td>
<td>11.10</td>
<td>16.81</td>
<td>.72</td>
<td>.82</td>
</tr>
<tr>
<td>Q1d</td>
<td>1.26</td>
<td>10.90</td>
<td>16.28</td>
<td>.70</td>
<td>.82</td>
</tr>
<tr>
<td>Q1e</td>
<td>1.29</td>
<td>10.82</td>
<td>16.74</td>
<td>.62</td>
<td>.85</td>
</tr>
</tbody>
</table>

Cronbach's alpha = .86

Table 5-10: Item Reliability Statistics of “Satisfaction with Group Process” Scale

<table>
<thead>
<tr>
<th>Qn. No.</th>
<th>Standard Deviation</th>
<th>Scale Mean if Item Deleted</th>
<th>Scale Variance if Item Deleted</th>
<th>Item-to-Total Correlation</th>
<th>Alpha if Item Deleted</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q3</td>
<td>.85</td>
<td>22.41</td>
<td>5.22</td>
<td>.46</td>
<td>.71</td>
</tr>
<tr>
<td>Q5</td>
<td>.89</td>
<td>21.96</td>
<td>4.80</td>
<td>.55</td>
<td>.67</td>
</tr>
<tr>
<td>Q6</td>
<td>.72</td>
<td>22.08</td>
<td>5.24</td>
<td>.58</td>
<td>.66</td>
</tr>
<tr>
<td>Q7</td>
<td>.76</td>
<td>22.26</td>
<td>5.54</td>
<td>.46</td>
<td>.71</td>
</tr>
<tr>
<td>Q8</td>
<td>.74</td>
<td>22.31</td>
<td>5.63</td>
<td>.44</td>
<td>.71</td>
</tr>
</tbody>
</table>

Cronbach’s alpha = .74

Table 5-11: Item Reliability Statistics of “Satisfaction with Group Judgments” Scale

As presented in Tables 5-10 and 5-11, the Cronbach’s alpha of the scales for “satisfaction with group process” and “satisfaction with group judgments” would not increase with the deletion of any of the items. As such, the current scales comprising the original five items will be kept and used in the statistical analyses.
Table 5-12: Item Reliability Statistics of “Perceived Consensus in Group Judgments” Scale

<table>
<thead>
<tr>
<th>Qn. No.</th>
<th>Standard Deviation</th>
<th>Scale Mean if Item Deleted</th>
<th>Scale Variance if Item Deleted</th>
<th>Item-to-Total Correlation</th>
<th>Alpha if Item Deleted</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q2</td>
<td>1.26</td>
<td>11.49</td>
<td>2.30</td>
<td>.58</td>
<td>.80</td>
</tr>
<tr>
<td>Q4</td>
<td>.87</td>
<td>11.03</td>
<td>3.36</td>
<td>.63</td>
<td>.67</td>
</tr>
<tr>
<td>Q9</td>
<td>.79</td>
<td>11.04</td>
<td>3.43</td>
<td>.70</td>
<td>.63</td>
</tr>
</tbody>
</table>

Cronbach’s alpha = .77

The reliability analysis of the scale measuring “perceived consensus in group judgments” indicates that Q2 differs slightly from Q4 and Q9. Question 2 prompted the respondents as to how different or similar the respondents’ final individual judgments were from their group’s judgments. Question 4 asked if the respondents agreed or disagreed with their group’s solution, whereas Question 9 asked for the extent to which the respondents support their group’s solution. The data set indicates that subjects may agree and support their group’s solution even when their individual judgments differ from group judgments. This indicates that the subjects were using different criteria to evaluate their own judgments and those of the group. Since the “perceived consensus in group judgments” construct includes both factors (Wheeler, 1993), all three of its items were retained.

The reliability statistics for the scales measuring “perceived usefulness of ESS” and “trust in ESS” were also generated both using the complete data set for these perception measures (cells 2-4 in Table 4-1, comprising both the expert and novice groups) and the data from the novice groups only (cells 2 and 3 in Table 4-1). Table 5-13 presents the reliability statistics for “perceived usefulness of ESS” using data from the novice groups only (cells 2 and 3 in Table 4-1) while Table 5-14 presents the reliability statistics for “perceived usefulness of ESS” using data from all of the groups, i.e., the expert group and the novice groups (cells 2-4 in Table 4-1).

117
<table>
<thead>
<tr>
<th>No.</th>
<th>Deviation</th>
<th>Scale Mean if Item Deleted</th>
<th>Scale Variance if Item Deleted</th>
<th>Item-to-Total Correlation</th>
<th>Alpha if Item Deleted</th>
</tr>
</thead>
<tbody>
<tr>
<td>P1</td>
<td>.88</td>
<td>47.08</td>
<td>39.11</td>
<td>.68</td>
<td>.83</td>
</tr>
<tr>
<td>P2</td>
<td>1.04</td>
<td>47.29</td>
<td>38.97</td>
<td>.56</td>
<td>.83</td>
</tr>
<tr>
<td>P6</td>
<td>.99</td>
<td>47.00</td>
<td>39.24</td>
<td>.58</td>
<td>.83</td>
</tr>
<tr>
<td>P7</td>
<td>1.32</td>
<td>47.33</td>
<td>37.87</td>
<td>.48</td>
<td>.85</td>
</tr>
<tr>
<td>P9</td>
<td>.76</td>
<td>46.96</td>
<td>40.80</td>
<td>.62</td>
<td>.83</td>
</tr>
<tr>
<td>P10</td>
<td>1.52</td>
<td>48.25</td>
<td>35.59</td>
<td>.52</td>
<td>.85</td>
</tr>
<tr>
<td>P11</td>
<td>.99</td>
<td>47.31</td>
<td>38.66</td>
<td>.63</td>
<td>.83</td>
</tr>
<tr>
<td>P13</td>
<td>.68</td>
<td>46.82</td>
<td>42.23</td>
<td>.53</td>
<td>.84</td>
</tr>
<tr>
<td>P14</td>
<td>.83</td>
<td>47.08</td>
<td>38.99</td>
<td>.74</td>
<td>.82</td>
</tr>
<tr>
<td>P16</td>
<td>1.25</td>
<td>47.27</td>
<td>38.68</td>
<td>.46</td>
<td>.85</td>
</tr>
</tbody>
</table>

Cronbach's alpha = .85

Table 5-13: Item Reliability Statistics of “Perceived Usefulness of ESS” Scale (Novices only)

<table>
<thead>
<tr>
<th>No.</th>
<th>Deviation</th>
<th>Scale Mean if Item Deleted</th>
<th>Scale Variance if Item Deleted</th>
<th>Item-to-Total Correlation</th>
<th>Alpha if Item Deleted</th>
</tr>
</thead>
<tbody>
<tr>
<td>P1</td>
<td>1.15</td>
<td>45.87</td>
<td>55.53</td>
<td>.73</td>
<td>.87</td>
</tr>
<tr>
<td>P2</td>
<td>1.09</td>
<td>45.88</td>
<td>57.43</td>
<td>.65</td>
<td>.87</td>
</tr>
<tr>
<td>P6</td>
<td>1.17</td>
<td>45.67</td>
<td>56.28</td>
<td>.67</td>
<td>.87</td>
</tr>
<tr>
<td>P7</td>
<td>1.41</td>
<td>46.00</td>
<td>54.24</td>
<td>.63</td>
<td>.87</td>
</tr>
<tr>
<td>P9</td>
<td>.91</td>
<td>45.58</td>
<td>59.13</td>
<td>.67</td>
<td>.87</td>
</tr>
<tr>
<td>P10</td>
<td>1.63</td>
<td>46.93</td>
<td>55.51</td>
<td>.46</td>
<td>.89</td>
</tr>
<tr>
<td>P11</td>
<td>1.20</td>
<td>46.03</td>
<td>54.50</td>
<td>.76</td>
<td>.86</td>
</tr>
<tr>
<td>P13</td>
<td>0.99</td>
<td>45.43</td>
<td>58.10</td>
<td>.68</td>
<td>.87</td>
</tr>
<tr>
<td>P14</td>
<td>.92</td>
<td>45.65</td>
<td>57.47</td>
<td>.80</td>
<td>.87</td>
</tr>
<tr>
<td>P16</td>
<td>1.21</td>
<td>45.70</td>
<td>60.74</td>
<td>.38</td>
<td>.89</td>
</tr>
</tbody>
</table>

Cronbach's alpha = .89

Table 5-14: Item Reliability Statistics of “Perceived Usefulness of ESS” Scale (Novices and Experts)

As can be seen from Tables 5-13 and 5-14, removing items 10 and 16 could possibly increase the Cronbach’s alpha of the scale for measuring “perceived usefulness of ESS.”
With these two items removed, the Cronbach’s alpha for “perceived usefulness of ESS” remains the same, i.e., at .85, using data from the novice groups, but it increases from .89 to .91 using the complete data set, i.e., data from all of the groups. Table 5-15 shows the reliability statistics for “perceived usefulness of ESS” using the complete data set with items 10 and 16 removed. These remaining eight items will be used for subsequent statistical analyses.

<table>
<thead>
<tr>
<th>Qn. No.</th>
<th>Standard Deviation</th>
<th>Scale Mean if Item Deleted</th>
<th>Scale Variance if Item Deleted</th>
<th>Item-to-Total Correlation</th>
<th>Alpha if Item Deleted</th>
</tr>
</thead>
<tbody>
<tr>
<td>P1</td>
<td>1.15</td>
<td>36.55</td>
<td>36.49</td>
<td>.76</td>
<td>.89</td>
</tr>
<tr>
<td>P2</td>
<td>1.09</td>
<td>36.57</td>
<td>38.60</td>
<td>.63</td>
<td>.90</td>
</tr>
<tr>
<td>P6</td>
<td>1.17</td>
<td>36.35</td>
<td>36.49</td>
<td>.74</td>
<td>.89</td>
</tr>
<tr>
<td>P7</td>
<td>1.41</td>
<td>36.68</td>
<td>35.16</td>
<td>.67</td>
<td>.90</td>
</tr>
<tr>
<td>P9</td>
<td>.91</td>
<td>36.26</td>
<td>40.08</td>
<td>.65</td>
<td>.90</td>
</tr>
<tr>
<td>P11</td>
<td>1.20</td>
<td>36.71</td>
<td>35.50</td>
<td>.80</td>
<td>.89</td>
</tr>
<tr>
<td>P13</td>
<td>0.99</td>
<td>36.12</td>
<td>39.28</td>
<td>.65</td>
<td>.90</td>
</tr>
<tr>
<td>P14</td>
<td>.92</td>
<td>36.33</td>
<td>38.46</td>
<td>.80</td>
<td>.89</td>
</tr>
</tbody>
</table>

Cronbach’s alpha = .91

Table 5-15: Item Reliability Statistics of Revised “Perceived Usefulness of ESS” Scale (Novices and Experts)

Table 5-16 presents the reliability statistics for “trust in ESS” using data from the novice groups only (cells 2 and 3 in Table 4-1) and Table 5-17 presents the reliability statistics for “trust in ESS” using data from all of the groups (cells 2-4 in Table 4-1).
### Table 5-16: Item Reliability Statistics of “Trust in ESS” Scale (Novices only)

<table>
<thead>
<tr>
<th>Qn. No.</th>
<th>Standard Deviation</th>
<th>Scale Mean if Item Deleted</th>
<th>Scale Variance if Item Deleted</th>
<th>Item-to-Total Correlation</th>
<th>Alpha if Item Deleted</th>
</tr>
</thead>
<tbody>
<tr>
<td>P3</td>
<td>.86</td>
<td>34.92</td>
<td>26.67</td>
<td>.47</td>
<td>.80</td>
</tr>
<tr>
<td>P4</td>
<td>1.11</td>
<td>35.73</td>
<td>22.76</td>
<td>.71</td>
<td>.76</td>
</tr>
<tr>
<td>P5</td>
<td>1.40</td>
<td>36.22</td>
<td>22.21</td>
<td>.56</td>
<td>.79</td>
</tr>
<tr>
<td>P8</td>
<td>.99</td>
<td>35.33</td>
<td>24.83</td>
<td>.58</td>
<td>.78</td>
</tr>
<tr>
<td>P12</td>
<td><strong>.99</strong></td>
<td><strong>34.55</strong></td>
<td><strong>29.69</strong></td>
<td><strong>.08</strong></td>
<td><strong>.84</strong></td>
</tr>
<tr>
<td>P15</td>
<td>1.18</td>
<td>35.75</td>
<td>23.43</td>
<td>.59</td>
<td>.75</td>
</tr>
<tr>
<td>P17</td>
<td>1.15</td>
<td>35.12</td>
<td>22.51</td>
<td>.71</td>
<td>.76</td>
</tr>
<tr>
<td>P18</td>
<td>.77</td>
<td>35.14</td>
<td>26.20</td>
<td>.60</td>
<td>.78</td>
</tr>
</tbody>
</table>

Cronbach’s alpha = .81

### Table 5-17: Item Reliability Statistics of “Trust in ESS” Scale (Novices and Experts)

<table>
<thead>
<tr>
<th>Qn. No.</th>
<th>Standard Deviation</th>
<th>Scale Mean if Item Deleted</th>
<th>Scale Variance if Item Deleted</th>
<th>Item-to-Total Correlation</th>
<th>Alpha if Item Deleted</th>
</tr>
</thead>
<tbody>
<tr>
<td>P3</td>
<td>.93</td>
<td>34.74</td>
<td>22.93</td>
<td>.47</td>
<td>.74</td>
</tr>
<tr>
<td>P4</td>
<td>1.12</td>
<td>35.59</td>
<td>20.68</td>
<td>.59</td>
<td>.72</td>
</tr>
<tr>
<td>P5</td>
<td>1.43</td>
<td>36.14</td>
<td>19.75</td>
<td>.49</td>
<td>.74</td>
</tr>
<tr>
<td>P8</td>
<td>.97</td>
<td>35.03</td>
<td>21.97</td>
<td>.56</td>
<td>.73</td>
</tr>
<tr>
<td>P12</td>
<td><strong>1.00</strong></td>
<td><strong>34.38</strong></td>
<td><strong>26.39</strong></td>
<td><strong>.06</strong></td>
<td><strong>.80</strong></td>
</tr>
<tr>
<td>P15</td>
<td>1.12</td>
<td>35.42</td>
<td>21.36</td>
<td>.52</td>
<td>.73</td>
</tr>
<tr>
<td>P17</td>
<td>1.09</td>
<td>34.79</td>
<td>21.37</td>
<td>.54</td>
<td>.73</td>
</tr>
<tr>
<td>P18</td>
<td>.82</td>
<td>34.97</td>
<td>22.65</td>
<td>.60</td>
<td>.73</td>
</tr>
</tbody>
</table>

Cronbach’s alpha = .77

Tables 5-16 and 5-17 indicate that item 12 is not a good measure of “trust in ESS”. It does not correlate well with the rest of the items (the item-to-total correlation is .08 in Table 5-16 and .06 in Table 5-17), and its deletion would increase the Cronbach’s alpha from .81 to .84 in Table 5-16 and from .77 to .80 in Table 5-17. As such, item 12 will be removed from subsequent statistical analyses.
Table 5-18 summarizes the results of the reliability tests for the five perception measures. The values of Alpha for the five perception scales are all within or above the 0.60 to 0.80 range recommended by Nunnally (1978) as a sufficient reliability level for “basic research”.

<table>
<thead>
<tr>
<th>Scale</th>
<th>No. of Items</th>
<th>Cronbach’s Alpha</th>
</tr>
</thead>
<tbody>
<tr>
<td>Satisfaction with group process</td>
<td>5</td>
<td>.86</td>
</tr>
<tr>
<td>Satisfaction with group judgments</td>
<td>5</td>
<td>.74</td>
</tr>
<tr>
<td>Perceived consensus in group judgments</td>
<td>3</td>
<td>.77</td>
</tr>
<tr>
<td>Perceived usefulness of ESS</td>
<td>8</td>
<td>.91</td>
</tr>
<tr>
<td>Trust in ESS</td>
<td>7</td>
<td>.80</td>
</tr>
</tbody>
</table>

Table 5-18: Overall Reliability of the Five Perception Scales

5.6.2.2 Validity of Perception Measures

The principal components factor analysis was carried out to examine the extent to which the individual items of each scale would converge and load on the underlying factor. All 28 items of the five scales were put through a confirmatory factor analysis using both the Varimax and Oblimin rotations by specifying the number of factors as five. As the “perceived consensus in group judgments” and “satisfaction with group judgments” load on the same factor, indicating that these two scales are either related or measuring the same construct, the principal components factor analysis was re-run by specifying the number of factors as four. Since both the Varimax and Oblimin rotations yield similar results, only the results of the rotated factor matrix of the Varimax rotation are presented in Table 5-19, with only loadings exceeding .4 displayed. A cut-off of .4 was chosen, as recommended by Stevens (1996, p. 372), which is in between the suggested cut-off of .3 by Child (1970) and .45 by Comrie (1973).
<table>
<thead>
<tr>
<th>Item</th>
<th>Intended Factor</th>
<th>Factor 1 (Usefulness of ESS)</th>
<th>Factor 2 (Consensus &amp; Satisfaction with Group Judgments)</th>
<th>Factor 3 (Satisfaction with Process)</th>
<th>Factor 4 (Trust in ESS)</th>
</tr>
</thead>
<tbody>
<tr>
<td>P14</td>
<td>1</td>
<td>.85</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P1</td>
<td>1</td>
<td>.83</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P11</td>
<td>1</td>
<td>.80</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P6</td>
<td>1</td>
<td>.76</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P2</td>
<td>1</td>
<td>.74</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P9</td>
<td>1</td>
<td>.73</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P13</td>
<td>1</td>
<td>.71</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P7</td>
<td>1</td>
<td>.67</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P18</td>
<td>4</td>
<td>.67</td>
<td></td>
<td></td>
<td>.49</td>
</tr>
<tr>
<td>Q9</td>
<td>2</td>
<td></td>
<td>.83</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Q5</td>
<td>2</td>
<td></td>
<td>.77</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Q4</td>
<td>2</td>
<td></td>
<td>.77</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Q6</td>
<td>2</td>
<td></td>
<td>.69</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Q2</td>
<td>2</td>
<td></td>
<td>.65</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Q7</td>
<td>2</td>
<td></td>
<td>.65</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Q3</td>
<td>2</td>
<td></td>
<td>.59</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Q8</td>
<td>2</td>
<td></td>
<td>.50</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Q1c</td>
<td>3</td>
<td></td>
<td></td>
<td>.81</td>
<td></td>
</tr>
<tr>
<td>Q1d</td>
<td>3</td>
<td></td>
<td></td>
<td>.81</td>
<td></td>
</tr>
<tr>
<td>Q1a</td>
<td>3</td>
<td></td>
<td></td>
<td>.79</td>
<td></td>
</tr>
<tr>
<td>Q1b</td>
<td>3</td>
<td></td>
<td></td>
<td>.76</td>
<td></td>
</tr>
<tr>
<td>Q1e</td>
<td>3</td>
<td></td>
<td></td>
<td>.73</td>
<td></td>
</tr>
<tr>
<td>P15</td>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td>.75</td>
</tr>
<tr>
<td>P17</td>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td>.73</td>
</tr>
<tr>
<td>P8</td>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td>.68</td>
</tr>
<tr>
<td>P4</td>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td>.62</td>
</tr>
<tr>
<td>P3</td>
<td>4</td>
<td></td>
<td>.41</td>
<td></td>
<td>.45</td>
</tr>
<tr>
<td>P5</td>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td>.43</td>
</tr>
</tbody>
</table>

*Table 5-19: Rotated Factor Loadings*
All of the items load more heavily on their intended factors except for P18, which loads more on the factor “usefulness of ESS” than the factor it is intended to measure, which is “trust in ESS”. A re-examination of P18 which reads “FINALYZER helped my group make good decisions” prompted us to conclude that it is more appropriate to include the item under the scale measuring “usefulness of ESS” than the scale measuring “trust in ESS”. As such, P18 was moved from the “trust in ESS” scale to the “usefulness of ESS” scale. Another observation of the factor loadings in Table 5-19 is that item P3 loads only slightly higher on factor 4 than factor 1. However, dropping it from the “trust in ESS” scale would reduce the Cronbach’s Alpha from .77 to .74, indicating that it is a good measure of “trust in ESS”. As such, item P3 is retained in the “trust in ESS” scale. With that, all of the 28 items loaded on their respective factors (refer to Table 5-19), indicating that the scales have a satisfactory level of convergent and discriminant validity.

However, since the items measuring “perceived consensus in group judgments” and “satisfaction with group judgments” load on the same factor, we carried out a closer examination of the three items intended to measure “perceived consensus in group judgments” (i.e., Q2, Q4 and Q9). These three items are “To what extent do you support your group’s solution?”, “Do you disagree or agree with your group’s solution?” and “How different or similar are your final individual decisions from your group’s decisions?”. They appealed to us as valid measures of “satisfaction with group judgments”, which prompted us to include these three items into the original five-item “satisfaction with group judgments” scale, giving a total of eight items. The correlation matrix of these eight items also suggests that these two sets of items are inseparable, i.e., there is no discriminant characteristic between them. A reliability test of the cohesiveness of these eight items was carried out and presented in Table 5-20. Table 5-20 indicates that the three items intended to measure “perceived consensus in group judgments” fit very well into the original “satisfaction with group judgments” scale. For instance, they have substantially greater
than .4 item-to-total correlations (Moore and Benbasat, 1991) and the removal of any of the three items from the scale would reduce the Cronbach’s alpha. As such, these eight items will be used to measure “satisfaction with group judgments”.

<table>
<thead>
<tr>
<th>Qn. No.</th>
<th>Standard Deviation</th>
<th>Scale Mean if Item Deleted</th>
<th>Scale Variance if Item Deleted</th>
<th>Item-to-Total Correlation</th>
<th>Alpha if Item Deleted</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q3</td>
<td>.85</td>
<td>39.19</td>
<td>19.05</td>
<td>.48</td>
<td>.83</td>
</tr>
<tr>
<td>Q5</td>
<td>.89</td>
<td>38.74</td>
<td>17.35</td>
<td>.70</td>
<td>.81</td>
</tr>
<tr>
<td>Q6</td>
<td>.72</td>
<td>38.86</td>
<td>19.12</td>
<td>.58</td>
<td>.82</td>
</tr>
<tr>
<td>Q7</td>
<td>.76</td>
<td>39.04</td>
<td>19.24</td>
<td>.52</td>
<td>.83</td>
</tr>
<tr>
<td>Q8</td>
<td>.74</td>
<td>39.10</td>
<td>20.04</td>
<td>.41</td>
<td>.84</td>
</tr>
<tr>
<td>Q2</td>
<td>1.26</td>
<td>39.25</td>
<td>15.91</td>
<td>.58</td>
<td>.83</td>
</tr>
<tr>
<td>Q4</td>
<td>.87</td>
<td>38.78</td>
<td>17.73</td>
<td>.66</td>
<td>.81</td>
</tr>
<tr>
<td>Q9</td>
<td>.79</td>
<td>38.80</td>
<td>17.62</td>
<td>.76</td>
<td>.80</td>
</tr>
</tbody>
</table>

Table 5-20: Item Reliability Statistics of Revised “Satisfaction with Group Judgments” Scale

Table 5-21 shows the Cronbach’s Alpha, the eigenvalues, and the percentage of variance explained by the four factors, taking into account the movement of item P18 from the “trust in ESS” scale to the “usefulness of ESS” scale. The four factors, comprising a total of 28 items, collectively accounted for 58.3% of the total variance. Each factor also had an eigenvalue that is significantly larger than the usual threshold of one.

In summary, the scales were found to be sound from the perspective of reliability and validity of measurement, and there was substantial evidence relating to the convergent and divergent validity of the items comprising the scales.
<table>
<thead>
<tr>
<th>Factor</th>
<th>Scale</th>
<th># of Items</th>
<th>Cronbach's Alpha</th>
<th>Eigenvalue</th>
<th>% of Variance Explained</th>
<th>Cumulative Variance (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Usefulness of ESS</td>
<td>9</td>
<td>.92</td>
<td>7.60</td>
<td>27.1</td>
<td>27.1</td>
</tr>
<tr>
<td>2</td>
<td>Satisfaction with group judgments</td>
<td>8</td>
<td>.84</td>
<td>4.31</td>
<td>15.4</td>
<td>42.5</td>
</tr>
<tr>
<td>3</td>
<td>Satisfaction with group process</td>
<td>5</td>
<td>.86</td>
<td>2.58</td>
<td>9.2</td>
<td>51.7</td>
</tr>
<tr>
<td>4</td>
<td>Trust in ESS</td>
<td>6</td>
<td>.77</td>
<td>1.84</td>
<td>6.6</td>
<td>58.3</td>
</tr>
</tbody>
</table>

*Table 5-21: Cronbach's Alpha, Eigenvalues, and Variance Explained by the Factors*

### 5.6.3 Qualitative Analysis

The analysis of conversations and interactions among group members were analyzed to identify the reasons for the use of explanation types and the groups' reactions to the ESS conclusions and explanations. The analysis was carried out to provide a richer understanding of, and greater insights into, the relationships between the independent and dependent (outcome and perception) variables. The analysis identified 1) why explanations were accessed and how they were used, 2) the groups' reactions to the absence of ESS explanations support, and 3) whether the ESS analyses were helpful to the groups, and if so, in what ways. In addition, any observable group process gains and losses, and any other interesting or unique observations were also presented and discussed. Conversations extracted from the group process were used as anecdotal evidence to provide support for the findings in the qualitative analysis. The results of the qualitative analyses are reported in Chapter 8.

### 5.7 Summary of Chapter 5

This chapter describes the research procedures, the research task, the subjects' characteristics, and the ESS, named FINALYZER, that was used in this study. The operationalization of the dependent variables was also discussed. Questionnaire
instruments were used for assessing user perceptions while conversations and interactions among group members were analyzed to understand the groups' interactions with the ESS. The reliability and validity of the perception variables were also evaluated. A combination of positivist (quantitative) and interpretive (qualitative) approaches was used for the study. Chapters 6 and 7 report the quantitative analysis of results while Chapter 8 reports the qualitative analysis.
CHAPTER 6: RESULTS OF QUANTITATIVE ANALYSES – PART I

Chapters 6, 7, and 8 report the results of the experimental study. Chapter 6 reports the results of quantitative analysis of outcome and perception variables for levels of ESS support, while Chapter 7 discusses the results from quantitative measures of outcome and perception variables between the experts and novices. Chapter 8 reports results of qualitative analyses of process variables by analyzing the conversations and interactions that were taking place in the groups.

Summary of Research Design

Table 6-1 shows the research design and the number of groups in each treatment. Section 6.1 highlights the statistical tests employed to analyze the quantitative results reported in Chapters 6 and 7. Section 6.2 compares 1) the consensus of the novice subjects' judgments and 2) the consistency of novice subjects' judgments with those of the original experts across levels of ESS support (cells 1 to 3). The analysis of consistency and consensus for the novices' and experts' judgments (cell 3 versus 4) are presented in Chapter 7. Section 6.3 compares the task-, group- and system-related perceptions (satisfaction with group process, satisfaction with group judgments, perceived usefulness of ESS, trust in ESS) of novices across levels of ESS support (cells 1 to 3). The difference in perceptions between the novice and expert groups that were provided with the full ESS support (cell 3 versus 4) is presented in Chapter 7.

<table>
<thead>
<tr>
<th>Expert Support by User Expertise</th>
<th>No ESS Support</th>
<th>ESS Analyses Support without Explanations</th>
<th>ESS Analyses Support with Explanations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Novice (Students)</td>
<td>Cell 1 (8 groups)</td>
<td>Cell 2 (8 groups)</td>
<td>Cell 3 (9 groups)</td>
</tr>
<tr>
<td>Expert (Professionals)</td>
<td>-</td>
<td>-</td>
<td>Cell 4 (6 groups)</td>
</tr>
</tbody>
</table>

*Table 6-1: Research Design and Number of Groups in Each Treatment*
6.1 Statistical Analyses Employed

Quantitative analyses were carried out to compare the performance and perceptions of:

(1) novices across the different levels of ESS support (cells 1 to 3), and
(2) novices versus experts that were provided with the full ESS support (cell 3 versus 4).

Quantitative analyses were carried out at both the individual and group levels. The nested (or hierarchical) ANOVA design (Anderson and Ager, 1978; Lindman, 1974; Myers, 1972; Neter, Kutner, Nachtsheim, and Wasserman, 1996; Winer, 1993) was used to analyze quantitative measures at the individual level. The Kruskal-Wallis ANOVA by ranks design and the Mann-Whitney U tests are two non-parametric tests (Siegel and Castellan, 1988) that were used when the assumptions of the parametric tests (e.g., F or t test) were not met. The validity of the parametric model (e.g., F or t test) requires three assumptions to be met (Neter, Kutner, Nachtsheim, and Wasserman, 1996; Stevens, 1996):

(1) The observations (or residuals) are normally distributed on the dependent variable in each group;
(2) The population variances of the dependent variable for the groups are equal (homogeneity of variance);
(3) The observations are independent.

The non-parametric or distribution-free statistical tests are more general as they do not make such assumptions. Although parametric tests tend to be more powerful than their non-parametric counterparts, their use requires the above assumptions to be met. Therefore, non-parametric statistics were used to overcome the problem of violations of assumptions. These assumptions are more likely to be violated when sample size is small (Siegel and Castellan, 1988; Tabachnick and Fidell, 1989).
The assumptions underlying the parametric model were checked before the \( F \) or \( t \) test was applied. If these assumptions were not met, the Kruskal-Wallis rank test or the Mann-Whitney \( U \) test was used to analyze the results. When the assumptions were satisfied, the Kruskal-Wallis rank test or the Mann-Whitney \( U \) test was used to corroborate the results of the \( F \) or \( t \) test. The Mann-Whitney \( U \) test is used to test whether two independent samples have been drawn from the same population, while the Kruskal-Wallis by rank test is used when more than two independent samples are involved. The Kruskal-Wallis ANOVA by ranks test (or non-parametric rank \( F \) test) is a widely used non-parametric test for testing the equality of treatment means. It ranks all \( N \) observations from 1 to \( N \) in ascending order and carries out the usual \( F \) test based on the ranks. Instead of using the \( F \) distribution approximation, the Kruskal-Wallis rank test uses a chi-square (\( \chi^2 \)) distribution approximation (Neter, Kutner, Nachtsheim, and Wasserman, 1996; Siegel and Castellan, 1988). The Mann-Whitney \( U \) test is equivalent to the Wilcoxon rank sum test, and the Kruskal-Wallis rank test for two groups. It tests whether two independent samples are equivalent, that is, from the same population. The observations from both groups are combined and ranked, with the average rank assigned in the case of ties. The number of times a score from group 1 precedes a score from group 2 and the number of times a score from group 2 precedes a score from group 1 are calculated. The Mann-Whitney \( U \) statistic is the smaller of these two numbers. The only requirement for the Kruskal-Wallis ANOVA by ranks and the Mann-Whitney \( U \) tests is the continuous distribution of the dependent variable.

The nested (or hierarchical) ANOVA design was used to analyze measures at the individual level (Anderson and Ager, 1978; Myers, 1972). In general, the nested design is used when there are more than one level of nestings; for instance, subjects may be nested within levels of a variable, which are in turn nested within the levels of another variable. In this case, subjects were assigned to (and therefore nested within) groups, which were
further assigned to (and nested within) experimental treatments. As such, it is a nested
design.

In nested designs, it is reasonable to assume that the total variability among subjects has
three potential sources. Subjects' scores may differ because of:

(1) *Treatment effects.* In this case, the level of decision support is a potential source;

(2) *Group effects.* Differences among the composition or characteristics of groups may
contribute to variability in the data.

(3) *Residual individual differences.* The scores of subjects within the same group may
vary due to such factors as attitude or ability.

The primary new aspect of the nested design is the assumption that an individual's score is
in part influenced by the social unit of which he or she is a member. Though the same
experimental treatment is applied to two individuals in two different groups, the responses
from these two individuals may differ, not merely because they are different individuals,
but also because they are subject to interactions with different sets of individuals and
events occurring in their groups. In small group research, group effects are typically
involved. It, therefore, becomes necessary to consider such group effects within the
statistical model. In other words, two sources of variation can be identified, one due to
treatment differences and the other due to group differences.

In this research, the treatment is *levels of ESS support* and the second factor, which is
nested within treatment, is *decision making groups* with three members in each group. As
such, this is a two-factor nested design (Neter, Kutner, Nachtsheim, and Wasserman,
1996), with one factor nested inside the other. The test statistic differs depending on
whether the factor effects are fixed or random. Table 6-2 shows the ANOVA table for
nested two-factor fixed, mixed, and random effects models. In this research, the treatment
factor effect is *fixed* and the group factor effect is *random*. Therefore, the appropriate test statistic given in the last column of Table 6-2 (A fixed and B random) was applied.

<table>
<thead>
<tr>
<th>Test for [Factors in this research]</th>
<th>A Fixed, B Fixed</th>
<th>A Fixed or Random, B Random</th>
</tr>
</thead>
<tbody>
<tr>
<td>Factor A [Treatment]</td>
<td>$F^* = \frac{MS_A}{MSE}$</td>
<td>$F^* = \frac{MS_A}{MSB(A)}$</td>
</tr>
<tr>
<td>Factor B(A) [Group(Treatment)]</td>
<td>$F^* = \frac{MSB(A)}{MSE}$</td>
<td>$F^* = \frac{MSB(A)}{MSE}$</td>
</tr>
</tbody>
</table>

*Table 6-2: Appropriate Test Statistic for Nested Two-Factor Designs with Fixed and Random Factor Effects (B nested within A)*

**6.1.1 Evaluation of Assumptions of Statistical Tests**

As mentioned earlier, the appropriateness of the parametric model depends on the validity of three assumptions: normality\(^1\), homogeneity of variances\(^2\), and independence\(^3\).

---

\(^1\) The normal probability plot, the detrended normal probability plot, the measures for skewness and kurtosis, the modified Kolmogorov-Smirnov (Lilliefors) test, and the Shapiro-Wilk test (if the number of observations per cell is less than 50) were used to examine the normality assumption (Hair, Anderson, Tatham and Black, 1995; Neter, Kutner, Nachtsheim and Wasserman, 1996; Tabachnick and Fidell, 1989). The normality assumption is satisfied when (1) the points on the normal probability plot cluster around a straight line, (2) the points on the detrended normal probability plot cluster around a horizontal line through zero, (3) the measures for skewness and kurtosis are not significantly different from zero, and (4) the modified Kolmogorov-Smirnov (Lilliefors) and the Shapiro-Wilk tests give insignificant $p$-values. The Shapiro-Wilk test shows good power in many situations compared to other tests of normality (such as the modified Kolmogorov-Smirnov test), especially when used in combination with the skewness and kurtosis coefficients (Wilk, Shapiro, and Chen, 1968; Conover, 1980).

\(^2\) The Levene test was used to assess the homogeneity of variances assumption (Neter, Kutner, Nachtsheim and Wasserman, 1996; Stevens, 1996). Unlike many of the frequently used tests for homogeneity of variance, such as Bartlett’s, Cochran’s, and Hartley’s $F_{\text{max}}$, which are quite sensitive to non-normality, the Levene test is less dependent on and more robust against non-normality (Stevens, 1996). Since the ANOVA model is pretty robust against departures from normality (Hays, 1994; Neter, Kutner, Nachtsheim, and Wasserman, 1996), the Levene test is particularly appropriate and useful with ANOVA. The Levene statistic is obtained by computing, for each case, the absolute difference from its cell mean and performing an ANOVA on these differences (Stevens, 1996). The homogeneity of variances assumption is satisfied when the null hypothesis that all group variances are equal is not rejected by the Levene test.

\(^3\) The validity of the independence assumption depends upon the experimental manipulations, that is, upon the care taken to ensure random assignment of treatments to subjects. In this research, the independence assumption was satisfied through the research design of the study, where subjects were randomly assigned to groups of three and the groups were in turn randomly assigned to treatments.
In general, the ANOVA model is quite robust to departures from normality, especially when the sample size is large; and it is quite robust to departures from homogeneity of variances, as long as the number of cases in each sample is equal or approximately equal (i.e., the ratio of the largest to the smallest sample (cell) size is less than 1.5) (Hays, 1994; Neter, Kutner, Nachtsheim, and Wasserman, 1996; Stevens, 1996). The independence assumption is by far the most important assumption, for even a small violation of it produces a substantial effect on both the level of significance and the power of the $F$ or $t$ statistic (Stevens, 1996).

6.2 Analysis of Consistency and Consensus of Judgments

A total of eight judgments were made by the subjects in each of three occasions, corresponding to individual pre-discussion judgments, group judgments, and individual post-discussion judgments. Questions 1 to 6 (see Appendix A) require the subjects to evaluate the liquidity, long-term solvency, asset utilization, the value of stock as loan collateral, and the quality of financial and operating management of the company, respectively, on a scale of 1-10. The answers to these six questions were used to evaluate their similarity with the judgments derived by a consensus of the five experts (so called original experts' consensus judgments) who were involved in developing FINALYZER (Dhaliwal, 1993). The consistency with judgments of original experts for each of the six judgments is assessed by its absolute deviation from the original experts' consensus judgment:

\[ D = |J - C| \]

where $D =$ is the deviation from consensus judgment, $J$ is the individual (or group) judgment made, and $C$ is the consensus judgment of the group of experts involved in developing FINALYZER. The sum of the absolute deviation, $D$, of these six judgments form the total deviation score. Therefore, the lower the total deviation score, $D$, the closer they are to the judgments of the original experts and the higher the consistency with
respect to the original experts' consensus judgments. The mean absolute error was selected as the measure since no a priori reason existed for viewing a positive error as more or less severe than a negative error. The use of the actual signed error as the measure would have allowed a subject's positive and negative errors to cancel each other out, which would have resulted in a mean error that was lower than the actual error incurred. In this research, consistency with the judgments of original experts can be achieved through a thorough analysis of the information given in the case or through knowledge transfer from the ESS to the decision makers.

The level of consensus of judgments were evaluated in two different ways: 1) consensus of individual judgments – the total deviation among the group members' individual post-discussion judgments, and 2) consensus with group judgments – the total deviation between the group judgments and the group members' individual post-discussion judgments. The consistency and consensus measures have a high degree of reliability as they were computed directly from the subjects' judgments.

6.2.1 Comparison of Novice Subjects' Performance Across Levels of ESS Support

The novice subjects' individual and group performance were compared across the different levels of ESS support. The validity of the assumptions underlying the ANOVA model was tested before statistical tests were applied.

6.2.1.1 Summary of Comparison of Novices' Judgments Across Levels of ESS Support

Tables 6-3 and 6-4 summarize the results of analyses of the novices' judgments across the different levels of ESS support conditions. Sections 6.2.1.2 to 6.2.1.7 present the detailed results.
The analyses of the results presented in the next few sections indicate that increasing level of ESS support increases knowledge transfer from the ESS to the users. More specifically, the results indicate that it is the combination of the ESS analyses and explanations support that contribute to the knowledge transfer. In other words, the addition of explanation features to ESS helps to further increase the amount of knowledge transfer. The ESS analyses and explanations support, however, do not lead to increased consensus in judgments.

6.2.1.2 Randomization Check—Consistency of Individual Pre-discussion Judgments
The consistency of individual pre-discussion judgments with judgments of the original experts who were involved in developing FINALYZER was assessed both at the individual and group levels, and compared across the experimental conditions to serve as a validation check for random assignment.
Evaluation of Assumptions of ANOVA Model

Individual as Level of Analysis. The absolute deviation of individual pre-discussion judgments from consensus judgments satisfies the homogeneity of variances and independence assumptions, but not the normality assumption. None of the commonly used transformations (i.e., cube, square, square root, logarithm, reciprocal of the square root, reciprocal) succeed in transforming the data to a normal distribution. As such, the Kruskal-Wallis rank test was used to perform the randomization check by comparing the consistency of individual pre-discussion judgments (pre-test) with the judgments of the original experts across the experimental groups.

Group as Level of Analysis. The absolute deviation of individual pre-discussion judgments from consensus judgments was also assessed at the group level by averaging the measure across the three members in a group. The normality, homogeneity of variances, and independence assumptions are all met. Therefore, both the ANOVA and Kruskal-Wallis rank tests were used to perform the randomization check.

Comparison of Consistency of Individual Pre-discussion Judgments Across Treatments

The descriptive statistics of the absolute deviation of individual pre-discussion judgments from consensus judgments are shown in Table 6-5. Table 6-8 shows the equivalent descriptive statistics aggregated at the group level.

---

4 The Levene statistic is .06 (p=.94), indicating that the homogeneity of variances assumption is not violated. The independence assumption is satisfied from the experimental procedure where subjects were randomly assigned to groups and then to treatments. Although the normal probability plot shows points that fall reasonably close to a straight line and the measure for skewness is .52 (p=.17), the measure for kurtosis is .76 (p=.00) and the Kolmogorov-Smirnov (Lillefors) statistic is .15 (p=.00), suggesting that the distribution does not satisfy the normality assumption.

5 The normality probability plot, the detrended normal probability plot, the measures of skewness (=.41; p=.19) and kurtosis (=.91; p=.16), the Kolmogorov-Smirnov (Lillefors) statistic (=.16; p=.13) and the Shapiro-Wilks' statistic (=.94, p=.24) support the normality assumption, and the Levene statistic (=.41; p=.67) supports the homogeneity of variances assumption.
**Individual Level**

<table>
<thead>
<tr>
<th>Treatment</th>
<th>N</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Std. Error</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control <em>(No ESS Support)</em></td>
<td>24</td>
<td>10.25</td>
<td>4.16</td>
<td>.85</td>
<td>3</td>
<td>20</td>
</tr>
<tr>
<td>Partial <em>(ESS Support w/o Explanations)</em></td>
<td>24</td>
<td>10.65</td>
<td>3.96</td>
<td>.81</td>
<td>4</td>
<td>20</td>
</tr>
<tr>
<td>Full <em>(ESS Support with Explanations)</em></td>
<td>27</td>
<td>9.15</td>
<td>3.48</td>
<td>.67</td>
<td>3</td>
<td>15</td>
</tr>
<tr>
<td>Total</td>
<td>75</td>
<td>9.98</td>
<td>3.87</td>
<td>.45</td>
<td>3</td>
<td>20</td>
</tr>
</tbody>
</table>

Table 6-5: Descriptive Statistics of Consistency of Individual Pre-discussion Judgments

There is no difference across the experimental groups in the consistency of individual pre-discussion judgments with those of the original experts when analyzed both at the individual level (Kruskal-Wallis $F_R$ test: $p=.39$) and the group level (ANOVA $F$ test: $p=.31$; Kruskal-Wallis $F_R$ test: $p=.31$), indicating *equivalence in pre-test performance across the experimental groups*. Table 6-6 summarizes the results. Table 6-7 shows the results of analysis at the individual level whereas Tables 6-9 and 6-10 show the results of analysis at the group level.

<table>
<thead>
<tr>
<th>Test \ Level of Analysis</th>
<th>Individual</th>
<th>Group</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>ANOVA (F)</em></td>
<td>N.A.</td>
<td>$p=.31$</td>
</tr>
<tr>
<td><em>Kruskal-Wallis (F$_R$)</em></td>
<td>$p=.39$</td>
<td>$p=.31$</td>
</tr>
</tbody>
</table>

Table 6-6: Results of Analysis of Consistency of Individual Pre-discussion Judgments
<table>
<thead>
<tr>
<th>Source</th>
<th>Chi-square ($x^2$)</th>
<th>DF</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Levels of ESS Support</td>
<td>1.91</td>
<td>2</td>
<td>.39</td>
</tr>
</tbody>
</table>

*Table 6-7: Results of Kruskal-Wallis Rank Test – Consistency of Individual Pre-discussion Judgments*

**Group Level**

<table>
<thead>
<tr>
<th>Treatment</th>
<th>N</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Std. Error</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control (No ESS Support)</td>
<td>8</td>
<td>10.25</td>
<td>2.54</td>
<td>.90</td>
<td>6.00</td>
<td>15.00</td>
</tr>
<tr>
<td>Partial (ESS Support w/o Expl.)</td>
<td>8</td>
<td>10.65</td>
<td>2.07</td>
<td>.73</td>
<td>8.67</td>
<td>14.00</td>
</tr>
<tr>
<td>Full (ESS Support with Expl.)</td>
<td>9</td>
<td>9.15</td>
<td>1.59</td>
<td>.53</td>
<td>6.00</td>
<td>10.67</td>
</tr>
<tr>
<td>Total</td>
<td>25</td>
<td>9.98</td>
<td>2.10</td>
<td>.42</td>
<td>6.00</td>
<td>15.00</td>
</tr>
</tbody>
</table>

*Table 6-8: Descriptive Statistics of Consistency of Individual Pre-discussion Judgments Aggregated at the Group Level*

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>DF</th>
<th>MS</th>
<th>F</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Levels of ESS Support</td>
<td>10.58</td>
<td>2</td>
<td>5.29</td>
<td>1.22</td>
<td>.31</td>
</tr>
<tr>
<td>Error</td>
<td>95.28</td>
<td>22</td>
<td>4.33</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Table 6-9: Results of ANOVA – Consistency of Individual Pre-discussion Judgments Analyzed at the Group Level*

<table>
<thead>
<tr>
<th>Source</th>
<th>Chi-square ($x^2$)</th>
<th>DF</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Levels of ESS Support</td>
<td>2.34</td>
<td>2</td>
<td>.31</td>
</tr>
</tbody>
</table>

*Table 6-10: Results of Kruskal-Wallis Rank Test – Consistency of Individual Pre-discussion Judgments Analyzed at the Group Level*
Although the consistency of individual pre-discussion judgments with those of the original experts does not satisfy the normality assumption when analyzed at the individual level, the ANOVA test produces a \( p \)-value of .36 (see Table 6-11) which is fairly close to the \( p \)-value of .39 produced by the Kruskal-Wallis rank test (refer to Table 6-7). Therefore, the ANOVA test seems to be fairly robust against departures from normality in this instance.

\[
\begin{array}{|c|c|c|c|c|}
\hline
\text{Source} & \text{SS} & \text{DF} & \text{MS} & \text{F} & \text{p-value} \\
\hline
\text{Levels of ESS Support} & 31.07 & 2 & 15.54 & 1.04 & .36 \\
\text{Error} & 1075.15 & 72 & 14.93 & & \\
\hline
\end{array}
\]

*Table 6-11: Results of ANOVA — Consistency of Individual Pre-discussion Judgments*

6.2.1.3 Consistency of Group Judgments

The consistency of group judgments, assessed using the absolute deviation of group judgments from the original experts' consensus judgments, was analyzed to determine if there is a treatment effect. The corresponding hypotheses to be tested are:

*Group as Level of Analysis — Consistency with Original Experts' Judgments*

\( H_{g1} \): The greater the level of ESS support provided, the greater the increase in the consistency of group judgments with the judgments of the original experts who were involved in developing the system.

\( H_{g1a} \): *ESS analyses support* increases the consistency of group judgments with the judgments of the original experts who were involved in developing the system.

\( H_{g1b} \): *ESS explanations support* increases the consistency of group judgments with the judgments of the original experts who were involved in developing the system.
An analysis of the group judgments only provides a measure of the differences in treatment groups. Examining the change between the individual pre-discussion judgments and the group judgments provides a measurement of the difference in performance that occurred. It also reduces the impact of any individual or group differences in ability at the start of the study. Hence, both the group judgments and the difference between the group and individual pre-discussion judgments will be compared across the treatment groups. The former is evaluated in this section while the latter is evaluated in the next section.

**Evaluation of Assumptions of ANOVA Model**

The absolute deviation of group judgments from consensus judgments satisfies the homogeneity of variances and independence assumptions, but not the normality assumption.\(^6\) As such, only the Kruskal-Wallis rank test was used to compare the consistency of group judgments across the experimental groups.

**Comparison of Consistency of Group Judgments Across Treatments**

Table 6-12 presents the descriptive statistics of the absolute deviation of group judgments from consensus judgments.

\(^6\) The Levene statistic of .37 \((p=.70)\) indicates that the homogeneity of variances assumption is not violated. The independence assumption is satisfied from the experimental procedure. However, the Kolmogorov-Smirnov (Lilliefors) and the Shapiro-Wilk statistics for the control group are .34 \((p=.01)\) and .76 \((p=.01)\) respectively, suggesting that the normality assumption is violated.
The descriptive statistics presented in Table 6-12 reveal that with increased levels of ESS support, not only was there a decrease in the absolute deviation of group judgments from original experts' consensus judgments (indicated by the decreasing mean with increased levels of ESS support), but the standard deviation and range of performance also decreased. The data in Table 6-12 suggests that increased levels of ESS support lower the maximum value of the absolute deviation from original experts' consensus judgments (as indicated by the "Max" column), indicating its influence on the groups whose judgments were furthest away from the original experts' consensus judgments.

The consistency of group judgments, assessed using the absolute deviation of group judgments from the original experts' consensus judgments, was analyzed using the non-parametric Kruskal-Wallis rank test. The results of the Kruskal-Wallis rank test, as shown in Table 6-13, indicate a significant difference across the different levels of ESS support (p=.04). The mean ranks of the consistency of group judgments in each of the experimental conditions are shown in Table 6-14.

<table>
<thead>
<tr>
<th>Treatment</th>
<th>N</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Std. Error</th>
<th>Min</th>
<th>Max</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control (No ESS Support)</td>
<td>8</td>
<td>8.63</td>
<td>4.10</td>
<td>1.45</td>
<td>5</td>
<td>18</td>
<td>13</td>
</tr>
<tr>
<td>Partial (ESS Support w/o Expl.)</td>
<td>8</td>
<td>7.63</td>
<td>3.29</td>
<td>1.16</td>
<td>4</td>
<td>14</td>
<td>10</td>
</tr>
<tr>
<td>Full (ESS Support with Expl.)</td>
<td>9</td>
<td>4.89</td>
<td>2.09</td>
<td>.70</td>
<td>3</td>
<td>8</td>
<td>5</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>25</strong></td>
<td><strong>6.96</strong></td>
<td><strong>3.49</strong></td>
<td><strong>.46</strong></td>
<td><strong>3</strong></td>
<td><strong>18</strong></td>
<td><strong>15</strong></td>
</tr>
</tbody>
</table>

*Table 6-12: Descriptive Statistics of Consistency of Group Judgments*

<table>
<thead>
<tr>
<th>Source</th>
<th>Chi-Square (χ²)</th>
<th>DF</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Levels of ESS Support</td>
<td>6.49</td>
<td>2</td>
<td>.04</td>
</tr>
</tbody>
</table>

*Table 6-13: Results of Kruskal-Wallis Rank Test – Consistency of Group Judgments*
A comparison of the magnitudes of these mean ranks was carried out using the multiple pairwise post-hoc comparisons procedure for the Kruskal-Wallis rank test (Neter, Kutner, Nachtsheim, and Wasserman, 1996; Siegel and Castellan, 1988). A similar analysis was also carried out using the Mann-Whitney U test (Siegel and Castellan, 1988). These tests were used to identify the specific treatment that caused the difference and to test the hypothesis that increased levels of ESS support would lead to increased consistency of group judgments with those of the original experts.

The results of the multiple pairwise comparisons are presented in Table 6-15 whereas the results of the Mann-Whitney U test are presented in Table 6-16.

### Table 6-14: Mean Ranks of Kruskal-Wallis Test — Consistency of Group Judgments

<table>
<thead>
<tr>
<th>Treatment</th>
<th>N</th>
<th>Mean Rank (R)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control (No ESS Support)</td>
<td>8</td>
<td>16.69</td>
</tr>
<tr>
<td>Partial (ESS Support w/o Expl.)</td>
<td>8</td>
<td>14.75</td>
</tr>
<tr>
<td>Full (ESS Support with Expl.)</td>
<td>9</td>
<td>8.17</td>
</tr>
<tr>
<td>Total</td>
<td>25</td>
<td></td>
</tr>
</tbody>
</table>

### Table 6-15: Results of Multiple Pairwise Comparisons for Kruskal-Wallis Test — Consistency of Group Judgments

<table>
<thead>
<tr>
<th>Treatment (I)</th>
<th>Treatment (J)</th>
<th>Absolute Mean Rank Difference:</th>
<th>p-value (2-tailed)</th>
<th>p-value (1-tailed)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control (no ESS)</td>
<td>Partial (ESS w/o Expl.)</td>
<td>1.94</td>
<td>1.79</td>
<td>.90</td>
</tr>
<tr>
<td>Partial (ESS w/o Expl.)</td>
<td>Full (ESS with Expl.)</td>
<td>6.58</td>
<td>.20</td>
<td>.10*</td>
</tr>
<tr>
<td>Control (no ESS)</td>
<td>Full (ESS with Expl.)</td>
<td>8.52</td>
<td>.05**</td>
<td>.03**</td>
</tr>
</tbody>
</table>

** significant at α=.05
* significant at α=.10
The results of the multiple pairwise post-hoc comparisons and the Mann-Whitney $U$ test suggest that the ESS explanation feature increases the consistency of group judgments with those of the original experts. Groups in the ESS with explanations support condition performed closer to the judgments of the original experts than groups in the other two treatment conditions (refer to results in Tables 6-15 and 6-16).

<table>
<thead>
<tr>
<th>Treatment (I)</th>
<th>N</th>
<th>Mean Rank</th>
<th>Sum of Ranks</th>
<th>Mann-Whitney $U$</th>
<th>$p$-value (2-tailed)</th>
<th>$p$-value (1-tailed)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control (no ESS)</td>
<td>8</td>
<td>9.19</td>
<td>73.50</td>
<td>26.50</td>
<td>.56</td>
<td>.28</td>
</tr>
<tr>
<td>Partial (ESS w/o Expl.)</td>
<td>8</td>
<td>7.81</td>
<td>62.50</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Partial (ESS w/o Expl.)</td>
<td>8</td>
<td>11.44</td>
<td>91.50</td>
<td>16.50</td>
<td>.06*</td>
<td>.03**</td>
</tr>
<tr>
<td>Full (ESS with Expl.)</td>
<td>9</td>
<td>6.83</td>
<td>61.50</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Control (no ESS)</td>
<td>8</td>
<td>12.00</td>
<td>96.00</td>
<td>12.00</td>
<td>.02**</td>
<td>.01**</td>
</tr>
<tr>
<td>Full (ESS with Expl.)</td>
<td>9</td>
<td>6.33</td>
<td>57.00</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 6-16: Results of Pairwise Comparisons using Mann-Whitney $U$ Test
- Consistency of Group Judgments

** significant at $\alpha=.05$
* significant at $\alpha=.10$

6.2.1.4 Change in Deviation Score from Individual Pre-discussion to Group Judgments

The change in deviation score from individual pre-discussion (group average) to group judgments was assessed at the group level by comparing it across the experimental groups.

Evaluation of Assumptions of ANOVA Model

The change in deviation score from individual pre-discussion to group judgments was assessed at the group level. It satisfies the normality, homogeneity of variances, and independence assumptions. The normality probability plot, its measures of skewness and
kurtosis, and the Kolmogorov-Smirnov (Lillefors) and Shapiro-Wilk statistics for the three experimental groups support the normality assumption, whereas the Levene statistic (=1.36; p=.28) supports the homogeneity of variances assumption. Therefore, both the ANOVA and Kruskal-Wallis rank tests were used for the analysis.

**Comparison of Change in Deviation Score from Pre-discussion to Group Judgments**

The descriptive statistics of the change in deviation score from individual pre-discussion (group average) to group judgments are shown in Table 6-17.

<table>
<thead>
<tr>
<th>Treatment</th>
<th>N</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Std. Error</th>
<th>Min</th>
<th>Max</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control (No ESS Support)</td>
<td>8</td>
<td>1.63</td>
<td>2.53</td>
<td>.89</td>
<td>-3</td>
<td>6</td>
<td>9</td>
</tr>
<tr>
<td>Partial (ESS Support w/o Expl.)</td>
<td>8</td>
<td>3.04</td>
<td>2.87</td>
<td>1.02</td>
<td>-1.33</td>
<td>6.67</td>
<td>8</td>
</tr>
<tr>
<td>Full (ESS Support with Expl.)</td>
<td>9</td>
<td>4.26</td>
<td>1.51</td>
<td>.50</td>
<td>2.33</td>
<td>6.33</td>
<td>4</td>
</tr>
<tr>
<td>Total</td>
<td>25</td>
<td>3.03</td>
<td>2.50</td>
<td>.50</td>
<td>-3</td>
<td>6.67</td>
<td>9.67</td>
</tr>
</tbody>
</table>

*Table 6-17: Descriptive Statistics of Change in Deviation Score from Individual Pre-discussion to Group Judgments Analyzed at the Group Level*

A preliminary analysis of the descriptive statistics indicates that the *mean* change in deviation score from individual pre-discussion (group average) to group judgments is largest in the ESS with explanations support condition followed by the ESS without explanations support condition, and it is smallest in the no ESS support condition. This indicates that the judgments of groups in the ESS with explanations support condition moved closer towards the consensus judgments of the original experts than groups in the other two conditions, suggesting a larger amount of knowledge transfer from the ESS to the groups that were provided with the ESS and its explanations support.
The results of analysis of improvement in deviation score from individual pre-discussion to group judgments are summarized in Table 6-18, and presented in greater detail in Tables 6-19 and 6-20. The mean ranks of improvement in deviation score from individual pre-discussion to group judgments are shown in Table 6-21.

<table>
<thead>
<tr>
<th>Test</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>ANOVA (F)</td>
<td>p=.09</td>
</tr>
<tr>
<td>Kruskal-Wallis ($F_g$)</td>
<td>p=.09</td>
</tr>
</tbody>
</table>

Table 6-18: Results of Analysis of Change from Individual Pre-discussion to Group Judgments

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>DF</th>
<th>MS</th>
<th>F</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Levels of ESS Support</td>
<td>29.38</td>
<td>2</td>
<td>14.69</td>
<td>2.68</td>
<td>.09</td>
</tr>
<tr>
<td>Error</td>
<td>120.76</td>
<td>22</td>
<td>5.49</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 6-19: Results of ANOVA – Change from Individual Pre-discussion to Group Judgments Analyzed at the Group Level

<table>
<thead>
<tr>
<th>Source</th>
<th>Chi-square ($x^2$)</th>
<th>DF</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Levels of ESS Support</td>
<td>4.79</td>
<td>2</td>
<td>.09</td>
</tr>
</tbody>
</table>

Table 6-20: Results of Kruskal-Wallis Rank Test – Change from Individual Pre-discussion to Group Judgments Analyzed at the Group Level
Table 6-21: Mean Ranks of Kruskal-Wallis Test – Change from Individual Pre-discussion to Group Judgments Analyzed at the Group Level

<table>
<thead>
<tr>
<th>Treatment</th>
<th>N</th>
<th>Mean Rank (R)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control (No ESS Support)</td>
<td>8</td>
<td>8.88</td>
</tr>
<tr>
<td>Partial (ESS Support w/o Explanations)</td>
<td>8</td>
<td>13.00</td>
</tr>
<tr>
<td>Full (ESS Support with Explanations)</td>
<td>9</td>
<td>16.67</td>
</tr>
<tr>
<td>Total</td>
<td>25</td>
<td></td>
</tr>
</tbody>
</table>

When change in the deviation score from the individual pre-discussion to the group judgments was analyzed at the group level, a significant difference was found (ANOVA $F$ test: $p=.09$; Kruskal-Wallis $F_R$ test: $p=.09$) at $\alpha=.10$. A priori contrasts indicate that groups in the control condition did not change as much as the aggregate effect of groups in the other two experimental conditions ($p=.04$), whereas groups in the ESS with explanations support condition changed more than the aggregate effect of groups in the other two conditions ($p=.02$). On the other hand, direct comparisons between the control (no ESS support) and partial (ESS without explanations support) groups and between the partial and full (ESS with explanations support) groups fail to produce significant results ($p=.16$ and $p=.15$ respectively). These, therefore, suggest that it was the combined effect of the ESS analysis and explanation support that contributed to the improvement in the consistency of group judgments (with respect to the original experts’ judgments). Post-hoc comparisons\(^7\) using Tukey and Scheffe tests produce similar results (see Table 6-22).

---

\(^7\) Tukey and Scheffe tests assume equal variances among the experimental groups, while Dunnett T3 and Games-Howell tests take into account unequal variances. Since the results for all four tests are consistent, only results of Tukey and Scheffe tests are shown.
Table 6-22: Results of Post-hoc Comparisons - Change from Individual Pre-discussion to Group Judgments Analyzed at the Group Level

** significant at $\alpha=.05$
* significant at $\alpha=.10$

The results of the multiple pairwise post-hoc comparisons for the Kruskal-Wallis rank test are presented in Table 6-23. The results of comparisons using the Mann-Whitney U test are presented in Table 6-24.

Table 6-23: Results of Multiple Pairwise Comparisons for Kruskal-Wallis Test - Change from Individual Pre-discussion to Group Judgments Analyzed at the Group Level

** significant at $\alpha=.05$
* significant at $\alpha=.10$
<table>
<thead>
<tr>
<th>Treatment (I)</th>
<th>Treatment (J)</th>
<th>N</th>
<th>Mean Rank</th>
<th>Sum of Ranks</th>
<th>Mann-Whitney U</th>
<th>p-value (2-tailed)</th>
<th>p-value (1-tailed)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control (no ESS)</td>
<td>8.750</td>
<td>8</td>
<td>60.00</td>
<td>24.00</td>
<td>.40</td>
<td>.20</td>
<td></td>
</tr>
<tr>
<td>Partial (ESS w/o Expl.)</td>
<td>8.950</td>
<td>8</td>
<td>76.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Partial (ESS w/o Expl.)</td>
<td>8</td>
<td>8.00</td>
<td>64.00</td>
<td>28.00</td>
<td>.44</td>
<td>.22</td>
<td></td>
</tr>
<tr>
<td>Full (ESS with Expl.)</td>
<td>9</td>
<td>9.89</td>
<td>89.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Control (no ESS)</td>
<td>8</td>
<td>5.88</td>
<td>47.00</td>
<td>11.00</td>
<td>.02**</td>
<td>.01**</td>
<td></td>
</tr>
<tr>
<td>Full (ESS with Expl.)</td>
<td>9</td>
<td>11.78</td>
<td>106.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 6-24: Results of Pairwise Comparisons using Mann-Whitney U Test—Change from Individual Pre-discussion to Group Judgments Analyzed at the Group Level

** significant at α=.05

Further analyses using the multiple pairwise comparisons for the Kruskal-Wallis rank test and the Mann-Whitney U test support similar findings as those obtained earlier. A significant difference is found between the control (no ESS support) and ESS with explanations support conditions (Kruskal-Wallis’ post-hoc comparison: \( p=.04 \); Mann-Whitney test: \( p=.01 \)), but not between the control and ESS without explanations support conditions (Kruskal-Wallis’ post-hoc comparison: \( p=.39 \); Mann-Whitney test: \( p=.20 \)) or between ESS with and without explanation support conditions (Kruskal-Wallis’ post-hoc comparison: \( p=.45 \); Mann-Whitney test: \( p=.22 \)). Consistent with findings from the a priori contrasts and the post-hoc comparisons, the results indicate that it is the combined effect of ESS analyses and explanations support that leads to the observed improvement in the consistency of group judgments with those of the original experts.

6.2.1.5 Consistency of Individual Post-discussion Judgments

The absolute deviation of individual post-discussion judgments from consensus judgments can be analyzed at both the group and individual levels. Group mean is used to analyze the
results when the level of analysis is group (Stevens, 1996). The nested design can be used to analyze the consistency of individual post-discussion judgments at the individual level if the assumptions of the ANOVA model are satisfied (Ager and Anderson, 1978). The hypotheses for this analysis are:

**Individual as Level of Analysis — Consistency with Original Experts’ Judgments**

**H_{ii}**: The greater the level of ESS support provided, the greater the increase in the consistency of individual judgments with those of the original experts who were involved in developing the system.

**H_{iia}**: Providing groups with *ESS analyses support* increases the consistency of *individual* judgments with the judgments of the original experts who were involved in developing the system.

**H_{iib}**: Providing groups with *ESS explanations support* increases the consistency of *individual* judgments with the judgments of the original experts who were involved in developing the system.

An analysis of the individual post-discussion judgments only provides a measure of the differences in treatment groups. Examining the change between the individual pre- and post-discussion judgments provide a measurement of the difference in performance that occurred. It also reduces the impact of any individual or group differences in ability at the start of the study. Hence, both the individual post-discussion judgments and the difference between the individual pre- and post-discussion judgments will be compared across the treatment groups. The former is carried out in this section and the latter in the next section.
Evaluation of Assumptions of ANOVA Model

Individual as Level of Analysis. The normality of residuals assumption of the nested design was not satisfied. As such, the nested design was not used for the analysis.

Group as Level of Analysis. The absolute deviation of individual post-discussion judgments from consensus judgments was first computed for each individual subject and then averaged across the three members in a group. This aggregate measure was checked to see whether they satisfy the assumptions of the ANOVA model. The homogeneity of variances and independence assumptions are satisfied, but not the normality assumption. As such, only the Kruskal-Wallis rank test was used to compare the consistency of individual post-discussion judgments (i.e., analyzed at the group level) across the experimental groups.

Comparison of Consistency of Individual Post-discussion Judgments Across Treatments

The absolute deviation of individual post-discussion judgments from original experts' consensus judgments was assessed at the group level by averaging it across the three individuals in a group. The descriptive statistics are shown in Table 6-25.

---

8 Although the homogeneity of variances assumption is not violated, the Kolmogorov-Smirnov (Lillefors) statistics for the residuals of both the control and partial (ESS without explanations support) groups are .00.

9 The Levene statistic of .33 (p=.72) indicates that the homogeneity of variances assumption is not violated. However, the Kolmogorov-Smirnov (Lillefors) and the Shapiro-Wilk statistics for the control group are .33 (p=.01) and .80 (p=.03) respectively, suggesting that the distribution in the control group is not normal.
### Table 6-25: Descriptive Statistics of Consistency of Individual Post-Discussion Judgments Analyzed at the Group Level

A preliminary analysis of the descriptive statistics given in Table 6-25 reveals that with increased levels of ESS support, not only was there a decrease in the aggregate absolute deviation of individual post-discussion judgments from original experts’ consensus judgments (indicated by the decreasing mean with increased levels of ESS support), but the standard deviation and range of performance also decreased.

The results of the Kruskal-Wallis rank test, as shown in Table 6-26, indicate a difference in the consistency of individual post-discussion judgments with those of the original experts across the different levels of ESS support \( (p=.09) \) at \( \alpha=.10 \). The mean ranks of the consistency of individual post-discussion judgments with those of the original experts are given in Table 6-27.

### Table 6-26: Results of Kruskal-Wallis Rank Test – Consistency of Individual Post-discussion Judgments

---

10 Only 7 cases were considered in the ESS without explanations support condition because one of the subjects did not specify one of his/her individual post-discussion judgments.
<table>
<thead>
<tr>
<th>Treatment</th>
<th>N</th>
<th>Mean Rank (R)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control (No ESS Support)</td>
<td>8</td>
<td>15.69</td>
</tr>
<tr>
<td>Partial (ESS Support w/o Explanations)\textsuperscript{11}</td>
<td>7</td>
<td>14.00</td>
</tr>
<tr>
<td>Full (ESS Support with Explanations)</td>
<td>9</td>
<td>8.50</td>
</tr>
<tr>
<td>Total</td>
<td>24</td>
<td></td>
</tr>
</tbody>
</table>

\textit{Table 6-27: Mean Ranks of Kruskal-Wallis Test}
\textit{— Consistency of Individual Post-discussion Judgments}

Comparison of the magnitudes of these mean ranks were carried out using the multiple pairwise post-hoc comparisons procedure for the Kruskal-Wallis rank test (Neter, Kutner, Nachtsheim, and Wasserman, 1996; Siegel and Castellan, 1988) and the Mann-Whitney \textit{U} test (Siegel and Castellan, 1988). These tests were used to identify the specific treatment that caused the difference and to test the hypothesis that increased levels of ESS support lead to increased consistency of individual post-discussion judgments with the judgments of the original experts. The results of the multiple pairwise post-hoc comparisons are presented in Table 6-28 whereas the results of the Mann-Whitney \textit{U} test are presented in Table 6-29.

<table>
<thead>
<tr>
<th>Treatment (I)</th>
<th>Treatment (J)</th>
<th>Absolute Mean Rank Difference:</th>
<th>\textit{p}-value\textit{(2-tailed)}</th>
<th>\textit{p}-value\textit{(1-tailed)}</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control (no ESS)</td>
<td>Partial (ESS w/o Expl.)</td>
<td>1.69</td>
<td>1.94</td>
<td>.97</td>
</tr>
<tr>
<td>Partial (ESS w/o Expl.)</td>
<td>Full (ESS with Expl.)</td>
<td>5.50</td>
<td>.37</td>
<td>.19</td>
</tr>
<tr>
<td>Control (no ESS)</td>
<td>Full (ESS with Expl.)</td>
<td>7.19</td>
<td>.11</td>
<td>.05**</td>
</tr>
</tbody>
</table>

\textit{Table 6-28: Results of Multiple Pairwise Comparisons for Kruskal-Wallis Test}
\textit{— Consistency of Individual Post-discussion Judgments}

\textit{**} significant at \(\alpha=.05\)

\textsuperscript{11} Only 7 cases were considered in the ESS without explanations support condition because one of the subjects did not specify one of his/her individual post-discussion judgments.
<table>
<thead>
<tr>
<th>Treatment (I)</th>
<th>N</th>
<th>Mean Rank</th>
<th>Sum of Ranks</th>
<th>Mann-Whitney U</th>
<th>p-value (2-tailed)</th>
<th>p-value (1-tailed)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control (no ESS)</td>
<td>8</td>
<td>8.44</td>
<td>67.50</td>
<td>24.50</td>
<td>.69</td>
<td>.34</td>
</tr>
<tr>
<td>Partial (ESS w/o Expl.)</td>
<td>7</td>
<td>7.50</td>
<td>52.50</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Partial (ESS w/o Expl.)</td>
<td>7</td>
<td>10.50</td>
<td>73.50</td>
<td>17.50</td>
<td>.14</td>
<td>.07*</td>
</tr>
<tr>
<td>Full (ESS with Expl.)</td>
<td>9</td>
<td>6.94</td>
<td>62.50</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Control (no ESS)</td>
<td>8</td>
<td>11.75</td>
<td>94.00</td>
<td>14.00</td>
<td>.03**</td>
<td>.02**</td>
</tr>
<tr>
<td>Full (ESS with Expl.)</td>
<td>9</td>
<td>6.56</td>
<td>59.00</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 6-29: Results of Pairwise Comparisons using Mann-Whitney U Test
— Consistency of Individual Post-discussion Judgments

** significant at α=.05  
* significant at α=.10

The results of the multiple pairwise post-hoc comparisons and the Mann-Whitney U test indicate that ESS analyses and explanations support are helpful in improving the consistency of individual judgments with those of the original experts (Mann-Whitney U test: p=.02; Kruskal-Wallis pairwise comparison: p=.05). There is also some evidence that the addition of explanations to ESS support contributes to the improvement in consistency of judgments to those of the original experts (Mann-Whitney U test: p=.07).

6.2.1.6 Change in Deviation Score from Initial to Post-discussion Individual Judgments

The change in deviation score from initial individual to post-discussion individual judgments was analyzed at both the *individual* and *group* levels. The validity of the assumptions underlying the ANOVA model was tested before statistical tests were applied.
Evaluation of Assumptions of ANOVA Model

When the improvement in deviation score was analyzed at the individual level, the normality, homogeneity of variances, and independence assumptions\textsuperscript{12} were satisfied. Thus, the nested design was used to run the analysis.

When improvement in deviation score from initial individual to post-discussion individual judgments was assessed at the group level, it satisfies the normality, homogeneity of variances, and independence assumptions.\textsuperscript{13} Thus, both the ANOVA and Kruskal-Wallis rank tests were used to run the analysis. In this case, the Kruskal-Wallis rank test is also used to verify the results of the ANOVA test.

Comparison of Change in Consistency from Individual Pre- to Post-discussion Judgments

The descriptive statistics of change in deviation score from individual pre- to post-discussion judgments are shown in Table 6-30. The corresponding descriptive statistics with group as the unit of analysis are shown in Table 6-32.

---

\textsuperscript{12} The Kolmogorov-Smirnov (Lillefors) and Shapiro-Wilk statistics do not show strong departures from normality, and the Levene statistic (=.134; \(p=.88\)) supports the homogeneity of variances assumption.

\textsuperscript{13} The normality probability plot, the detrended normal probability plot, the measures of skewness and kurtosis, and the Kolmogorov-Smirnov (Lillefors) and Shapiro-Wilk statistics do not show strong departures from normality, and the Levene statistic (=1.51; \(p=.24\)) supports the homogeneity of variances assumption. As multiple tests (coefficients of skewness and kurtosis, the modified Kolmogorov-Smirnov (Lillefors) statistic, and the Shapiro-Wilk statistic) were performed on each experimental group which increased the chances of type I error, Stevens (1996) advises the adoption of a more stringent alpha level (e.g., .01) to keep the overall type I error rate (i.e., the probability of at least one false rejection) somewhat under control. As such, the Kolmogorov-Smirnov (Lillefors) and Shapiro-Wilk statistics of .29 (\(p=.07\)) and .80 (\(p=.05\)) for the partial (ESS without explanations support) group were considered acceptable in satisfying the normality assumption.
The descriptive statistics in Table 6-30 suggest an interesting phenomenon. The increased levels of ESS support seemed to have helped the weaker members in improving their judgments, but not the stronger members. As indicated in the “Min” and “Max” columns, increased levels of ESS support improve the “Min” score, but the “Max” score remains more or less the same. In short, increased levels of ESS support benefit the weaker members in the groups.

The results of the nested design indicate that the treatment effect is marginal, though not significant. Table 6-31 shows the results of analysis of nested design.

---

14 Recall from Table 6-2 that the error term for treatment effect is the group within treatment effect.
Group Level

<table>
<thead>
<tr>
<th>Treatment</th>
<th>N</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Std. Error</th>
<th>Min</th>
<th>Max</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control (No ESS Support)</td>
<td>8</td>
<td>1.67</td>
<td>2.07</td>
<td>.73</td>
<td>-2.33</td>
<td>5.00</td>
<td>7.33</td>
</tr>
<tr>
<td>Partial (ESS Support w/o Expl.)</td>
<td>7</td>
<td>3.14</td>
<td>2.58</td>
<td>.98</td>
<td>-1.00</td>
<td>5.34</td>
<td>6.34</td>
</tr>
<tr>
<td>Full (ESS Support with Expl.)</td>
<td>9</td>
<td>3.74</td>
<td>1.36</td>
<td>.45</td>
<td>2.00</td>
<td>6.00</td>
<td>4.00</td>
</tr>
<tr>
<td>Total</td>
<td>24</td>
<td>2.87</td>
<td>2.12</td>
<td>.43</td>
<td>-2.33</td>
<td>6.00</td>
<td>8.33</td>
</tr>
</tbody>
</table>

Table 6-32: Descriptive Statistics of Change in Deviation Score from Individual Pre- to Post-discussion Judgments Analyzed at the Group Level

The results of analyzing the difference in the deviation score from individual pre- to post-discussion judgments at the group level are summarized in Table 6-33, and presented in greater detail in Tables 6-34 and 6-35. The mean ranks are shown in Table 6-36.

<table>
<thead>
<tr>
<th>Test</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>ANOVA (F)</td>
<td>p=.12</td>
</tr>
<tr>
<td>Kruskal-Wallis (F&lt;sub&gt;k&lt;/sub&gt;)</td>
<td>p=.10</td>
</tr>
</tbody>
</table>

Table 6-33: Results of Analysis of Difference in Consistency of Individual Pre- and Post-discussion Judgments

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>DF</th>
<th>MS</th>
<th>F</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Levels of ESS Support</td>
<td>18.90</td>
<td>2</td>
<td>9.45</td>
<td>2.34</td>
<td>.12</td>
</tr>
<tr>
<td>Error</td>
<td>84.76</td>
<td>21</td>
<td>4.04</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 6-34: Results of ANOVA – Change from Individual Pre- to Post-discussion Judgments Analyzed at the Group Level
<table>
<thead>
<tr>
<th>Source</th>
<th>Chi-square (x²)</th>
<th>DF</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Levels of ESS Support</td>
<td>4.70</td>
<td>2</td>
<td>.10</td>
</tr>
</tbody>
</table>

Table 6-35: Results of Kruskal-Wallis Rank Test – Change from Individual Pre- to Post-discussion Judgments Analyzed at the Group Level

<table>
<thead>
<tr>
<th>Treatment</th>
<th>N</th>
<th>Mean Rank (R)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control (No ESS Support)</td>
<td>8</td>
<td>8.13</td>
</tr>
<tr>
<td>Partial (ESS Support w/o Explanations)</td>
<td>7</td>
<td>14.14</td>
</tr>
<tr>
<td>Full (ESS Support with Explanations)</td>
<td>9</td>
<td>15.11</td>
</tr>
<tr>
<td>Total</td>
<td>24</td>
<td></td>
</tr>
</tbody>
</table>

Table 6-36: Mean Ranks of Kruskal-Wallis Test – Change from Individual Pre- to Post-discussion Judgments Analyzed at the Group Level

The improvement in the deviation score from individual pre- to post-discussion judgments was analyzed at the group level using the Kruskal-Wallis $F_R$ and the ANOVA $F$ tests. The results obtained from the Kruskal-Wallis rank test is significant at $\alpha=.10$, but the difference is not significant using the ANOVA test ($p=.12$). A priori contrasts indicate that the control group did not change as much as the aggregate effect of groups in the other two experimental conditions ($p=.04$), while direct comparisons between the control (no ESS support) and partial (ESS without explanations support) groups and between the partial and full (ESS with explanations support) groups fail to produce significant results ($p=.13$ and $p=.30$ respectively). Post-hoc comparisons using Tukey and Scheffe tests (see Table 6-37) indicate that ESS analyses support alone does not significantly improve the consistency of individual judgments with those of the original experts; it is the combination of ESS analyses and explanations support that contributed to the improvement in consistency.

---

15 Results for all four tests, Tukey, Scheffe, Dunnett T3 and Games Howell, are consistent, therefore, only the results of Tukey and Scheffe tests are shown.
<table>
<thead>
<tr>
<th>Treatment (I)</th>
<th>Treatment (J)</th>
<th>Absolute Mean Difference:</th>
<th>Post-hoc Test</th>
<th>p-value (2-tailed)</th>
<th>p-value (1-tailed)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control (no ESS)</td>
<td>Partial (ESS w/o Expl.)</td>
<td>1.47</td>
<td>Tukey</td>
<td>.35</td>
<td>.18</td>
</tr>
<tr>
<td>Partial (ESS w/o Expl.)</td>
<td>Full (ESS with Expl.)</td>
<td>.60</td>
<td>Tukey</td>
<td>.83</td>
<td>.41</td>
</tr>
<tr>
<td>Control (no ESS)</td>
<td>Full (ESS with Expl.)</td>
<td>2.07</td>
<td>Tukey</td>
<td>.11</td>
<td>.05**</td>
</tr>
</tbody>
</table>

Table 6-37: Results of Post-hoc Comparisons — Change from Individual Pre- to Post-discussion Judgments Analyzed at the Group Level

** significant at $\alpha=.05$
* significant at $\alpha=.10$

The results of multiple pairwise post-hoc comparisons for the Kruskal-Wallis rank test are presented in Table 6-38, and the results of comparisons using the Mann-Whitney U test are presented in Table 6-39.

<table>
<thead>
<tr>
<th>Treatment (I)</th>
<th>Treatment (J)</th>
<th>Absolute Mean Rank Difference:</th>
<th>p-value (2-tailed)</th>
<th>p-value (1-tailed)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control (no ESS)</td>
<td>Partial (ESS w/o Expl.)</td>
<td>6.01</td>
<td>.30</td>
<td>.15</td>
</tr>
<tr>
<td>Partial (ESS w/o Expl.)</td>
<td>Full (ESS with Expl.)</td>
<td>.97</td>
<td>1.00</td>
<td>.50</td>
</tr>
<tr>
<td>Control (no ESS)</td>
<td>Full (ESS with Expl.)</td>
<td>6.98</td>
<td>.13</td>
<td>.06*</td>
</tr>
</tbody>
</table>

Table 6-38: Results of Multiple Pairwise Comparisons for Kruskal-Wallis Test — Change from Individual Pre- to Post-discussion Judgments Analyzed at the Group Level

* significant at $\alpha=.10$
<table>
<thead>
<tr>
<th>Treatment (I)</th>
<th>Mean Rank</th>
<th>Sum of Ranks</th>
<th>Mann-Whitney U</th>
<th>p-value (2-tailed)</th>
<th>p-value (1-tailed)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Control (no ESS)</strong></td>
<td>6.75</td>
<td>54.00</td>
<td>18.00</td>
<td>.25</td>
<td>.13</td>
</tr>
<tr>
<td>Partial (ESS w/o Expl.)</td>
<td>9.43</td>
<td>66.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Partial (ESS w/o Expl.)</td>
<td>8.71</td>
<td>61.00</td>
<td>30.00</td>
<td>.87</td>
<td>44</td>
</tr>
<tr>
<td>Full (ESS with Expl.)</td>
<td>8.33</td>
<td>75.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Control (no ESS)</strong></td>
<td>5.94</td>
<td>47.50</td>
<td>11.50</td>
<td>.02**</td>
<td>.01**</td>
</tr>
<tr>
<td>Full (ESS with Expl.)</td>
<td>11.72</td>
<td>105.50</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 6-39 Results of Pairwise Comparisons using Mann-Whitney U Test – Change from Individual Pre- to Post-discussion Judgments Analyzed at the Group Level

** significant at α=.05

Further analyses using the multiple pairwise post-hoc comparisons for the Kruskal-Wallis rank test and direct comparisons using the Mann-Whitney U test support similar findings as those obtained earlier. A significant difference is found between the control (no ESS support) and ESS with explanations support conditions (Kruskal-Wallis’ post-hoc comparison: \( p=.07 \); Mann-Whitney test: \( p=.01 \)), but not between the control and ESS without explanations support conditions (Kruskal-Wallis’ post-hoc comparison: \( p=.15 \); Mann-Whitney test: \( p=.13 \)) or between ESS with and without explanation support conditions (Kruskal-Wallis’ post-hoc comparison: \( p=.50 \); Mann-Whitney test: \( p=.44 \)). Consistent with findings from the a priori contrasts and the post-hoc comparisons, the results indicate that the combined effect of ESS analyses and explanations support leads to the observed improvement in the consistency of individual judgments (with respect to the original experts’ judgments).
6.2.1.7 Consensus in Judgments

Consensus was measured in two ways: 1) consensus among individual judgments — the total distance between the group members’ post-discussion judgments, and 2) consensus with group judgments — the distance between the group judgments and its members’ post-discussion individual judgments.

6.2.1.7.1 Consensus Among Individual Judgments

The absolute distance between the group members’ post-discussion individual judgments was computed and compared across the experimental conditions. The hypotheses are:

Consensus Among Individual Judgments

H_{12}:

The greater the level of ESS support provided, the greater the level of consensus among individual judgments.

H_{12a}:

ESS analyses support increases members’ consensus among individual judgments.

H_{12b}:

ESS explanations support increases members’ consensus among individual judgments.

Evaluation of Assumptions of ANOVA Model

The Kolmogorov-Smirnov (Lillefors) and the Shapiro-Wilk statistics indicate that the distributions of the total absolute deviation between the three group members’ post-discussion individual judgments for both of the ESS support conditions violate the normality assumption. As such, only the Kruskal-Wallis rank test was used for the analysis.

Comparison of Consensus Among Individual Judgments Across Experimental Conditions

The descriptive statistics of the total absolute distance between the three group members’ post-discussion individual judgments are presented in Table 6-40.
Treatment & N & Mean & Std. Dev. & Std. Error & Min & Max & Range \\
--- & --- & --- & --- & --- & --- & --- & --- \\
Control *(No ESS Support)* & 8 & 10.75 & 6.23 & 2.20 & 0 & 20 & 20 \\
Partial *(ESS Support w/o Expl.)* & 8 & 13.50 & 6.74 & 2.38 & 8 & 26 & 18 \\
Full *(ESS Support with Expl.)* & 9 & 9.56 & 7.60 & 2.53 & 2 & 28 & 26 \\
Total & 25 & 11.20 & 6.83 & 1.37 & 0 & 28 & 28 \\

*Table 6-40: Descriptive Statistics of Total Absolute Distance between Group Members’ Post-discussion Individual Judgments Analyzed at the Group Level*

Tables 6-41 and 6-42 show the mean ranks of the absolute distance between the three group members’ post-discussion individual judgments.

| Treatment & N & Mean Rank | 
|--- & --- & --- | 
| Control *(No ESS Support)* & 8 & 13.50 | 
| Partial *(ESS Support w/o Explinations)* & 8 & 15.63 | 
| Full *(ESS Support with Explinations)* & 9 & 10.22 | 
| Total & 25 & 

*Table 6-41: Mean Ranks of Kruskal-Wallis Test – Total Absolute Distance between Group Members’ Post-discussion Individual Judgments*

| Source & Chi-square (χ²) & DF & p-value | 
|--- & --- & --- & --- | 
| Levels of ESS Support & 2.38 & 2 & .31 | 

*Table 6-42: Results of Kruskal-Wallis Rank Test – Total Absolute Distance between Group Member’s Post-discussion Individual Judgments*

The results in Table 6-42 show no difference in the level of consensus among groups provided with different levels of ESS support.
6.2.1.7.2 Consensus with Group Judgments

The absolute distance of the group members' post-discussion individual judgments from group judgments was computed and compared across the experimental conditions. The hypotheses are:

Consensus with Group Judgments

\( H_{g2} \): The greater the level of ESS support provided, the greater the level of consensus with group judgments.

\( H_{g2a} \): ESS analyses support increases members' consensus with group judgments.

\( H_{g2b} \): ESS explanations support increases members' consensus with group judgments.

Evaluation of Assumptions of ANOVA Model

The Kolmogorov-Smirnov (Lillefors) and the Shapiro-Wilk statistics indicate that the distributions of the total absolute distance of the group members' post-discussion individual judgments from group judgments for both of the ESS support conditions violate the normality assumption. As such, only the Kruskal-Wallis rank test was used for the analysis.

Comparison of Consensus with Group Judgments Across Experimental Conditions

The descriptive statistics of the total absolute distance of the three group members' post-discussion individual judgments from the group judgments are presented in Table 6-43.
<table>
<thead>
<tr>
<th>Treatment</th>
<th>N</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Std. Error</th>
<th>Min</th>
<th>Max</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control (No ESS Support)</td>
<td>8</td>
<td>6.63</td>
<td>4.03</td>
<td>1.43</td>
<td>0</td>
<td>13</td>
<td>13</td>
</tr>
<tr>
<td>Partial (ESS Support w/o Expl.)</td>
<td>8</td>
<td>7.13</td>
<td>4.02</td>
<td>1.42</td>
<td>4</td>
<td>14</td>
<td>10</td>
</tr>
<tr>
<td>Full (ESS Support with Expl.)</td>
<td>9</td>
<td>4.89</td>
<td>4.11</td>
<td>1.37</td>
<td>1</td>
<td>15</td>
<td>14</td>
</tr>
<tr>
<td>Total</td>
<td>25</td>
<td>6.16</td>
<td>4.01</td>
<td>.80</td>
<td>0</td>
<td>15</td>
<td>15</td>
</tr>
</tbody>
</table>

Table 6-43: Descriptive Statistics of Total Absolute Distance of Group Members’ Post-discussion Individual Judgments from Group Judgments Analyzed at the Group Level

Tables 6-44 and 6-45 show the mean ranks of the absolute distance of the three group members’ post-discussion individual judgments from group judgments.

<table>
<thead>
<tr>
<th>Treatment</th>
<th>N</th>
<th>Mean Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control (No ESS Support)</td>
<td>8</td>
<td>14.50</td>
</tr>
<tr>
<td>Partial (ESS Support w/o Explanations)</td>
<td>8</td>
<td>15.06</td>
</tr>
<tr>
<td>Full (ESS Support with Explanations)</td>
<td>9</td>
<td>9.83</td>
</tr>
<tr>
<td>Total</td>
<td>25</td>
<td></td>
</tr>
</tbody>
</table>

Table 6-44: Mean Ranks of Kruskal-Wallis Test – Total Absolute Distance of Group Members’ Post-discussion Individual Judgments from Group Judgments

<table>
<thead>
<tr>
<th>Source</th>
<th>Chi-square ($\chi^2$)</th>
<th>DF</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Levels of ESS Support</td>
<td>2.67</td>
<td>2</td>
<td>.26</td>
</tr>
</tbody>
</table>

Table 6-45: Results of Kruskal-Wallis Rank Test – Total Absolute Distance of Group Member’s Post-discussion Individual Judgments from Group Judgments

The results in Table 6-45 indicate no difference in the level of consensus from group judgments among groups provided with the different levels of ESS support. In summary,
ESS analyses and explanations support seem to have no effect on the level of consensus in the novice groups.

6.2.1.8 Summary of Results

The results indicate that ESS analyses support alone (i.e., no explanations facilities) does not produce significant knowledge transfer from the ESS. However, the combination of ESS analyses and explanations support contributes to significant knowledge transfer from the ESS to the users. In other words, the addition of ESS explanation facilities to ESS analysis support is necessary for significant knowledge transfer to take place. The ESS analyses and explanations support have no effect on consensus among novices.

6.3 Comparison of Novice Subjects’ Perceptions Across Levels of ESS Support

The novice subjects’ perceptions were compared across the different levels of ESS support. The validity of the assumptions underlying the nested ANOVA model was tested before statistical tests were applied. Since these perception measures were captured at the individual level, the nested design would be most appropriate for the analysis. However, if the assumptions of the nested ANOVA model are violated, these measures will be averaged across the group and analyzed at the group level.

6.3.1 Summary of Comparison of Novices’ Perceptions Across Levels of ESS Support

Table 6-46 summarizes the results of comparison of the novices’ perceptions across the different levels of ESS support conditions.
ESS and its explanations support does not increase decision makers’ satisfaction with group judgments. Instead, they lower the decision makers’ satisfaction with group process. Interestingly, although the explanation facilities of ESS increase decision makers’ trust in the system, they are not perceived to be useful in supporting group decision making. The following sections present the detailed results and analyses.

### 6.3.1.1 Perception Measure — Satisfaction with Group Process

The novice subjects’ satisfaction with group process was compared across the different levels of ESS support conditions. The ratings were analyzed and presented here in the form where the higher the subjects’ ratings of the scale, the greater the subjects’ satisfaction with group process. The hypothesis to be tested is:

$$H_{p2}: \text{The greater the level of ESS support provided, the lower the satisfaction with group decision making process.}$$

**Evaluation of Assumptions of ANOVA Model**

**Testing Assumptions of Nested ANOVA Design with Individual as Level of Analysis.** The distributions of the residuals do not satisfy the homogeneity of variances and normality
assumptions. As such, the nested design was not used. Instead, the use of regular ANOVA and Kruskal-Wallis rank tests were considered next.

*Testing Assumptions of ANOVA Design with Group as Level of Analysis.* The novice subjects’ perceived satisfaction with group process can be assessed at the *group* level by averaging their individual measures across the three members in a group. In this case, the normality, homogeneity of variances, and independence assumptions are satisfied. Therefore, both the ANOVA and the Kruskal-Wallis rank tests were used to assess the subjects’ perceived satisfaction with group process.

*Comparison of Novice Subjects’ Satisfaction with Group Process Across Treatments*

The descriptive statistics of the novice subjects’ satisfaction with group process are presented in Table 6-47 (with individual as the unit of analysis) and Table 6-48 (with group as the unit of analysis).

<table>
<thead>
<tr>
<th>Treatment</th>
<th>N</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Std. Error</th>
<th>Min</th>
<th>Max</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control <em>(No ESS Support)</em></td>
<td>24</td>
<td>23.83</td>
<td>3.57</td>
<td>.73</td>
<td>15</td>
<td>30</td>
<td>15</td>
</tr>
<tr>
<td>Partial <em>(ESS Support w/o Expl.)</em></td>
<td>24</td>
<td>20.71</td>
<td>6.22</td>
<td>1.27</td>
<td>6</td>
<td>28</td>
<td>22</td>
</tr>
<tr>
<td>Full <em>(ESS Support with Expl.)</em></td>
<td>27</td>
<td>21.11</td>
<td>4.09</td>
<td>.79</td>
<td>12</td>
<td>28</td>
<td>16</td>
</tr>
<tr>
<td>Total</td>
<td>75</td>
<td>21.85</td>
<td>4.88</td>
<td>.56</td>
<td>6</td>
<td>30</td>
<td>24</td>
</tr>
</tbody>
</table>

*Table 6-47: Descriptive Statistics of Perceived Satisfaction with Group Process (Individual Level)*

---

16 The Levene statistic is 6.42 (*p*=.00), indicating that the homogeneity of variances assumption is violated. The Kolmogorov-Smirnov (Lilliefors) and the Shapiro-Wilk statistics for the ESS with explanations support group are .20 (*p*=.01) and .91 (*p*=.02) respectively, suggesting that the distribution does not satisfy the normality assumption.

17 The Kolmogorov-Smirnov (Lilliefors) and the Shapiro-Wilk's statistics support the normality assumption, and the Levene statistic (=.10; *p*=.91) supports the homogeneity of variances assumption.
<table>
<thead>
<tr>
<th>Treatment</th>
<th>N</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Std. Error</th>
<th>Min</th>
<th>Max</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control (No ESS Support)</td>
<td>8</td>
<td>23.83</td>
<td>2.54</td>
<td>.90</td>
<td>20.33</td>
<td>27.00</td>
<td>6.67</td>
</tr>
<tr>
<td>Partial (ESS Support w/o Expl.)</td>
<td>8</td>
<td>20.71</td>
<td>2.66</td>
<td>.94</td>
<td>16.33</td>
<td>24.33</td>
<td>8.00</td>
</tr>
<tr>
<td>Full (ESS Support with Expl.)</td>
<td>9</td>
<td>21.11</td>
<td>2.64</td>
<td>.88</td>
<td>17.33</td>
<td>26.67</td>
<td>9.33</td>
</tr>
<tr>
<td>Total</td>
<td>25</td>
<td>21.85</td>
<td>2.87</td>
<td>.57</td>
<td>16.33</td>
<td>27.00</td>
<td>10.67</td>
</tr>
</tbody>
</table>

*Table 6-48: Descriptive Statistics of Perceived Satisfaction with Group Process (Group Level)*

Table 6-49 summarizes the results of the analysis, which are presented in greater detail in Tables 6-50 and 6-51. The mean ranks of the Kruskal-Wallis test are shown in Table 6-52.

<table>
<thead>
<tr>
<th>Test \ Level of Analysis</th>
<th>Group</th>
</tr>
</thead>
<tbody>
<tr>
<td>ANOVA (F)</td>
<td><em>p</em>=.05</td>
</tr>
<tr>
<td>Kruskal-Wallis (Fₚ)</td>
<td><em>p</em>=.05</td>
</tr>
</tbody>
</table>

*Table 6-49: Results of Analysis of Novice Subjects' Satisfaction with Group Process*

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>DF</th>
<th>MS</th>
<th>F</th>
<th><em>p</em>-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Levels of ESS Support</td>
<td>46.82</td>
<td>2</td>
<td>23.41</td>
<td>3.42</td>
<td>.05</td>
</tr>
<tr>
<td>Error</td>
<td>150.54</td>
<td>22</td>
<td>6.84</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Table 6-50: Results of ANOVA – Satisfaction with Group Process*

<table>
<thead>
<tr>
<th>Source</th>
<th>Chi-square (x²)</th>
<th>DF</th>
<th><em>p</em>-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Levels of ESS Support</td>
<td>5.85</td>
<td>2</td>
<td>.05</td>
</tr>
</tbody>
</table>

*Table 6-51: Results of Kruskal-Wallis Rank Test – Satisfaction with Group Process*
Treatment | N | Mean Rank
--- | --- | ---
Control (No ESS Support) | 8 | 18.13
Partial (ESS Support w/o Explanations) | 8 | 9.94
Full (ESS Support with Explanations) | 9 | 11.17
Total | 25 | 

Table 6-52: Mean Ranks of Kruskal-Wallis Test — Satisfaction with Group Process

The novice subjects’ satisfaction with group process differs significantly across the treatment groups (ANOVA $F$ test: $p=.05$; Kruskal-Wallis $F_r$ test: $p=.05$). A priori contrasts indicate that groups in the control condition were more satisfied with group process than groups in the other two experimental conditions, i.e., groups provided with some form of ESS support ($p=.02$). This indicates that satisfaction with the group decision making process is decreased with ESS support. This is an interesting finding because ESS support has shown strong evidence in improving the consistency of group judgments with those of the original experts (refer to Sections 6.2.1.3 and 6.2.1.4). Thus, in considering whether to use ESS support in the group context, two opposing factors have to be considered – the consistency of group judgments with those of the original experts and the members’ satisfaction with group process. If the former is more important, then the use of ESS support should be considered. However, if the latter (i.e., members’ satisfaction with group process) is more important, then the use of ESS support may not be appropriate.

Post-hoc comparisons\(^8\) using Tukey and Scheffe tests also produce similar results (see Table 6-53), indicating that ESS support decreases group members’ satisfaction with their group decision making process.

\(^8\) Results for all four tests, Tukey, Scheffe, Dunnett T3, and Games Howell, are consistent; therefore, only the results of Tukey and Scheffe tests were shown.
<table>
<thead>
<tr>
<th>Treatment (I)</th>
<th>Treatment (J)</th>
<th>Absolute Mean Difference:</th>
<th>Post-hoc Test</th>
<th>p-value (2-tailed)</th>
<th>p-value (1-tailed)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control (no ESS)</td>
<td>Partial (ESS w/o Expl.)</td>
<td>3.13</td>
<td>Tukey</td>
<td>.06*</td>
<td>.03**</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Scheffe</td>
<td>.08*</td>
<td>.04**</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Dunnett T3</td>
<td>.09*</td>
<td>.04**</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Games-H</td>
<td>.07*</td>
<td>.04**</td>
</tr>
<tr>
<td>Partial (ESS w/o Expl.)</td>
<td>Full (ESS with Expl.)</td>
<td>.40</td>
<td>Tukey</td>
<td>.95</td>
<td>.47</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Scheffe</td>
<td>.95</td>
<td>.48</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Dunnett T3</td>
<td>.99</td>
<td>.49</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Games-H</td>
<td>.95</td>
<td>.47</td>
</tr>
<tr>
<td>Control (no ESS)</td>
<td>Full (ESS with Expl.)</td>
<td>2.72</td>
<td>Tukey</td>
<td>.10*</td>
<td>.05**</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Scheffe</td>
<td>.12</td>
<td>.06*</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Dunnett T3</td>
<td>.13</td>
<td>.06*</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Games-H</td>
<td>.11</td>
<td>.06*</td>
</tr>
</tbody>
</table>

Table 6-53: Results of Post-hoc Comparisons — Satisfaction with Group Process

** significant at α=.05
* significant at α=.10

6.3.1.2 Perception Measure — Satisfaction with Group Judgments

The novice subjects' satisfaction with group judgments was also compared across the different levels of ESS support conditions. The ratings were analyzed and presented so that the higher the subjects' ratings of the scale, the greater the subjects' satisfaction with group judgments. The hypothesis to be tested is:

H_{p1}: The greater the level of ESS support provided, the greater the satisfaction with group judgments.
Evaluation of Assumptions of ANOVA Model

Testing Assumptions of Nested ANOVA Design with Individual as Level of Analysis. The distributions of the residuals do not satisfy the homogeneity of variances assumption\textsuperscript{19}. As such, the nested design was not applied. Instead, the use of regular ANOVA and Kruskal-Wallis rank tests were considered next.

Testing Assumptions of ANOVA Design with Group as Level of Analysis. The novice subjects' perceived satisfaction with group judgments can be assessed at the group level by averaging their individual measures across the three members in a group. In this case, the normality, homogeneity of variances, and independence assumptions are satisfied\textsuperscript{20}. Therefore, both the ANOVA and the Kruskal-Wallis rank tests were used to assess the subjects' perceived satisfaction with group judgments.

Comparison of Novice Subjects' Satisfaction with Group Judgments Across Treatments

The descriptive statistics of the novice subjects' satisfaction with group process are presented in Tables 6-54 (with individual as the unit of analysis) and 6-55 (with group as the unit of analysis).

\textsuperscript{19} Although the Kolmogorov-Smirnov (Lillefors) and the Shapiro-Wilk statistics support the normality assumption, the Levene statistic of 4.01 (p=.02) indicates that the homogeneity of variances assumption is violated.

\textsuperscript{20} The Kolmogorov-Smirnov (Lillefors) and the Shapiro-Wilks' statistics support the normality assumption, and the Levene statistic (=.71; p=.50) supports the homogeneity of variances assumption.
Table 6-54: Descriptive Statistics of Perceived Satisfaction with Group Judgments  
(Individual Level)

<table>
<thead>
<tr>
<th>Treatment</th>
<th>N</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Std. Error</th>
<th>Min</th>
<th>Max</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control (No ESS Support)</td>
<td>24</td>
<td>44.38</td>
<td>5.08</td>
<td>1.04</td>
<td>31</td>
<td>53</td>
<td>22</td>
</tr>
<tr>
<td>Partial (ESS Support w/o Expl.)</td>
<td>24</td>
<td>44.38</td>
<td>5.50</td>
<td>1.12</td>
<td>27</td>
<td>55</td>
<td>28</td>
</tr>
<tr>
<td>Full (ESS Support with Expl.)</td>
<td>27</td>
<td>44.26</td>
<td>4.82</td>
<td>.93</td>
<td>32</td>
<td>52</td>
<td>20</td>
</tr>
<tr>
<td>Total</td>
<td>75</td>
<td>44.33</td>
<td>5.06</td>
<td>.58</td>
<td>27</td>
<td>55</td>
<td>28</td>
</tr>
</tbody>
</table>

Table 6-55: Descriptive Statistics of Perceived Satisfaction with Group Judgments  
(Group Level)

<table>
<thead>
<tr>
<th>Treatment</th>
<th>N</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Std. Error</th>
<th>Min</th>
<th>Max</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control (No ESS Support)</td>
<td>8</td>
<td>44.38</td>
<td>3.98</td>
<td>1.41</td>
<td>36.00</td>
<td>48.00</td>
<td>12.00</td>
</tr>
<tr>
<td>Partial (ESS Support w/o Expl.)</td>
<td>8</td>
<td>44.38</td>
<td>2.36</td>
<td>.83</td>
<td>40.33</td>
<td>47.67</td>
<td>7.33</td>
</tr>
<tr>
<td>Full (ESS Support with Expl.)</td>
<td>9</td>
<td>44.26</td>
<td>3.91</td>
<td>1.30</td>
<td>36.67</td>
<td>50.67</td>
<td>14.00</td>
</tr>
<tr>
<td>Total</td>
<td>25</td>
<td>44.33</td>
<td>3.37</td>
<td>.67</td>
<td>36.00</td>
<td>50.67</td>
<td>14.67</td>
</tr>
</tbody>
</table>

Table 6-56 summarizes the results of the analysis, which are presented in greater detail in Tables 6-57 and 6-58. The mean ranks of the Kruskal-Wallis test are shown in Table 6-59.

<table>
<thead>
<tr>
<th>Test \ Level of Analysis</th>
<th>Group</th>
</tr>
</thead>
<tbody>
<tr>
<td>ANOVA (F)</td>
<td>$p=1.0$</td>
</tr>
<tr>
<td>Kruskal-Wallis ($F_s$)</td>
<td>$p=.91$</td>
</tr>
</tbody>
</table>

Table 6-56: Results of Analysis of Novice Subjects' Satisfaction with Group Judgments
<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>DF</th>
<th>MS</th>
<th>F</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Levels of ESS Support</td>
<td>.08</td>
<td>2</td>
<td>.04</td>
<td>.00</td>
<td>1.0</td>
</tr>
<tr>
<td>Error</td>
<td>272.15</td>
<td>22</td>
<td>12.37</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Table 6-57: Results of ANOVA — Satisfaction with Group Judgments*

<table>
<thead>
<tr>
<th>Source</th>
<th>Chi-square (χ²)</th>
<th>DF</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Levels of ESS Support</td>
<td>.19</td>
<td>2</td>
<td>.91</td>
</tr>
</tbody>
</table>

*Table 6-58: Results of Kruskal-Wallis Rank Test — Satisfaction with Group Judgments*

<table>
<thead>
<tr>
<th>Treatment</th>
<th>N</th>
<th>Mean Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control (No ESS Support)</td>
<td>8</td>
<td>13.94</td>
</tr>
<tr>
<td>Partial (ESS Support w/o Explanations)</td>
<td>8</td>
<td>12.50</td>
</tr>
<tr>
<td>Full (ESS Support with Explanations)</td>
<td>9</td>
<td>12.61</td>
</tr>
<tr>
<td>Total</td>
<td>25</td>
<td></td>
</tr>
</tbody>
</table>

*Table 6-59: Mean Ranks of Kruskal-Wallis Test — Satisfaction with Group Judgments*

The results clearly show that the novice subjects’ satisfaction with group judgments does not differ across the treatment groups (ANOVA F test: $p=.10$; Kruskal-Wallis $F_R$ test: $p=.91$). This indicates that ESS support has no effect on the group members’ satisfaction with group judgments. Interestingly, ESS support has been shown to be helpful in improving the consistency of group judgments with those of the original experts (refer to Sections 6.2.1.3 and 6.2.1.4), but, surprisingly, it does not lead to increased satisfaction with group judgments. One possible explanation is that the subjects did not perceive the ESS support to be helpful in bringing them closer to the judgments of the original experts, which explains why their satisfaction with group judgments did not increase.
6.3.1.3 Perception Measure—Perceived Usefulness of ESS

The novice subjects' perception of the usefulness of the ESS was compared across the two levels of ESS support conditions, i.e., with and without the explanation facilities. The ratings were analyzed and presented in the form where the higher the subjects' ratings of the scale, the greater the subjects' perceived usefulness of ESS. The hypothesis to be tested is:

\[ H_{p3}: \text{The explanation facilities increase users' perceived usefulness of the ESS.} \]

Evaluation of Assumptions of ANOVA Model

Testing Assumptions of Nested ANOVA Design with Individual as Level of Analysis. The distributions of the residuals show a slight violation of the normality assumption\(^{21}\). However, these statistics are acceptable, and they are not considered a strong violation of the assumption if Stevens' (1996) recommendation to adopt an alpha level of .01 is accepted. As such, the nested design will be applied. However, we also used both the \(t\) and the Mann-Whitney \(U\) tests to verify the results of the nested design.

Testing Assumptions of ANOVA Design with Group as Level of Analysis. The novice subjects’ perceived usefulness of the ESS can also be assessed at the group level by averaging their individual measures across the three members in a group. In this case, the normality, homogeneity of variances, and independence assumptions are satisfied\(^ {22}\).

\(^{21}\) Although the Levene statistic (=2.31; \(p=.14\)) indicates that the homogeneity of variances assumption is satisfied, the Kolmogorov-Smirnov (Lillefors) and the Shapiro-Wilk statistics for the partial (ESS without explanations) support group are .20 \( (p=.02) \) and .89 \( (p=.02) \) respectively, indicating that the normality assumption may be violated.

\(^{22}\) The Kolmogorov-Smirnov (Lillefors) and the Shapiro-Wilks' statistics support the normality assumption, and the Levene statistic (=.09; \(p=.77\)) supports the homogeneity of variances assumption.
Comparison of Novice Subjects' Perceived Usefulness of ESS Across Treatments

The descriptive statistics of the novice subjects' perceived usefulness of ESS are presented in Table 6-60 (with individual as the unit of analysis) and Table 6-61 (with group as the unit of analysis).

<table>
<thead>
<tr>
<th>Treatment</th>
<th>N</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Std. Error</th>
<th>Min</th>
<th>Max</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Partial (ESS Support w/o Expl.)</td>
<td>24</td>
<td>47.21</td>
<td>4.93</td>
<td>1.01</td>
<td>35</td>
<td>56</td>
<td>21</td>
</tr>
<tr>
<td>Full (ESS Support with Expl.)</td>
<td>27</td>
<td>49.30</td>
<td>6.52</td>
<td>1.25</td>
<td>39</td>
<td>62</td>
<td>23</td>
</tr>
<tr>
<td>Total</td>
<td>51</td>
<td>48.31</td>
<td>5.86</td>
<td>.82</td>
<td>35</td>
<td>62</td>
<td>27</td>
</tr>
</tbody>
</table>

Table 6-60: Descriptive Statistics of Perceived Usefulness of ESS (Individual Level)

<table>
<thead>
<tr>
<th>Treatment</th>
<th>N</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Std. Error</th>
<th>Min</th>
<th>Max</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Partial (ESS Support w/o Expl.)</td>
<td>8</td>
<td>47.21</td>
<td>3.83</td>
<td>1.35</td>
<td>40.67</td>
<td>52.67</td>
<td>12.00</td>
</tr>
<tr>
<td>Full (ESS Support with Expl.)</td>
<td>9</td>
<td>49.30</td>
<td>4.73</td>
<td>1.58</td>
<td>40.33</td>
<td>56.67</td>
<td>16.33</td>
</tr>
<tr>
<td>Total</td>
<td>17</td>
<td>48.31</td>
<td>4.33</td>
<td>1.05</td>
<td>40.33</td>
<td>56.67</td>
<td>16.33</td>
</tr>
</tbody>
</table>

Table 6-61: Descriptive Statistics of Perceived Usefulness of ESS (Group Level)

Table 6-62 summarizes the results of the analysis, which are presented in greater detail in Tables 6-63 to 6-65.

<table>
<thead>
<tr>
<th>Test \ Level of Analysis</th>
<th>Individual</th>
<th>Group</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nested ANOVA</td>
<td>p=.34</td>
<td>N.A.</td>
</tr>
<tr>
<td>Difference between Means (t)</td>
<td>N.A.</td>
<td>p=.34</td>
</tr>
<tr>
<td>Mann-Whitney (U)</td>
<td>N.A.</td>
<td>p=.34</td>
</tr>
</tbody>
</table>

Table 6-62: Results of Analysis of Novice Subjects' Perceived Usefulness of ESS
The nested ANOVA, $t$ and Mann-Whitney $U$ tests produce consistent results. There is no
difference in the novice subjects’ perception of the usefulness of the ESS across the two
levels of ESS support conditions (nested ANOVA test: $p=.34$; ANOVA $F$ test: $p=.34$;
Kruskal-Wallis $F_r$ test: $p=.34$). In other words, the novices did not perceive the
explanations provided by the ESS to be useful despite earlier findings that explanations
contribute to increased consistency of group judgments with those of the original experts
(refer to Sections 6.2.1.3 and 6.2.1.4). The results of the nested design, which are
presented in Table 6-63, also indicate that the novice subjects’ perception of the usefulness
of the ESS differs among the members in a group.

---

Recall from Table 5-2 that the error term for treatment effect is the group within treatment effect.
6.3.1.4 Perception Measure — Trust in ESS

The novice subjects' level of trust in the ESS was compared across the two levels of ESS support conditions, i.e., with and without the explanation facilities. The ratings were analyzed and presented in the form where the higher the subjects' ratings of the scale, the greater the subjects' trust in the ESS. The hypothesis to be tested is:

\[ H_{p4} \: \text{: The explanation facilities increase users' trust in the ESS.} \]

**Evaluation of Assumptions of ANOVA Model**

**Testing Assumptions of Nested ANOVA Design with Individual as Level of Analysis.** The distributions of the residuals satisfy the normality, homogeneity of variances and independence assumptions\(^\text{24}\). As such, it is not necessary to carry out the analysis at the group level.

**Comparison of Novice Subjects' Trust in ESS Across Treatments**

The descriptive statistics of the novice subjects' level of trust in ESS are presented in Table 6-66 (with individual as the unit of analysis) and the results of the nested ANOVA analysis are presented in Table 6-67.

<table>
<thead>
<tr>
<th>Treatment</th>
<th>N</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Std. Error</th>
<th>Min</th>
<th>Max</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Partial (ESS Support w/o Expl.)</td>
<td>24</td>
<td>27.33</td>
<td>5.16</td>
<td>1.05</td>
<td>15</td>
<td>34</td>
<td>19</td>
</tr>
<tr>
<td>Full (ESS Support with Expl.)</td>
<td>27</td>
<td>31.00</td>
<td>4.11</td>
<td>.79</td>
<td>24</td>
<td>39</td>
<td>15</td>
</tr>
<tr>
<td>Total</td>
<td>51</td>
<td>29.27</td>
<td>4.95</td>
<td>.69</td>
<td>15</td>
<td>39</td>
<td>24</td>
</tr>
</tbody>
</table>

*Table 6-66: Descriptive Statistics of Trust in ESS (Individual Level)*

\(^{24}\) The Kolmogorov-Smirnov (Lilliefors) and the Shapiro-Wilk statistics support the normality assumption, and the Levene statistic \(= .04; p=.85\) supports the homogeneity of variances assumption.
The results of the nested ANOVA design indicate that providing explanations in an ESS increases the novice subjects’ level of trust in the ESS \( (p=.05) \). This supports the findings of Ye and Johnson (1995) that explanation facilities are helpful in increasing users’ acceptability of the ESS advice. The results presented in Table 6-66 also indicate that trust in ESS differs among members in a group.

### 6.4 Summary of Chapter 6

This chapter reports the quantitative analysis of the effects of the different level of ESS support on consistency with judgments of original experts, consensus, and perceptions.

The influence of ESS on the judgments of novice users increases with increased levels of ESS support. The influence is more apparent in the group judgments than the individual judgments. However, the level of ESS support, i.e., ESS conclusions and explanations, has no effect on consensus. The use of ESS for group decision making decreases satisfaction with group process. However, the use of ESS analyses and explanations support has no effect on satisfaction with group judgments. The availability of the explanation facilities increases the novice users’ trust in the system, but it does not increase the novice users’ perceived usefulness of the system.

---

**Table 6-67: Results of Nested ANOVA – Trust in ESS**

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>DF</th>
<th>MS</th>
<th>F</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Levels of ESS Support(^\text{25})</td>
<td>170.82</td>
<td>1</td>
<td>170.82</td>
<td>4.36</td>
<td>.05</td>
</tr>
<tr>
<td>Group within Treatment</td>
<td>588.00</td>
<td>15</td>
<td>39.20</td>
<td>2.86</td>
<td>.01</td>
</tr>
<tr>
<td>Error</td>
<td>465.33</td>
<td>34</td>
<td>13.69</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\(^{25}\) Recall from Table 5-2 that the error term for treatment effect is the group within treatment effect.
CHAPTER 7: RESULTS OF QUANTITATIVE ANALYSES – PART II

This chapter discusses the comparative results from quantitative measures of the outcome and the perception variables between the experts and novices. Chapter 8 reports the results of qualitative analyses of process variables using the observational analysis approach.

7.1 Comparison of Novice versus Expert Subjects’ Performance with ESS Support

The individual and group performance of the novice and expert subjects that were provided with ESS analyses and explanations support were compared. The validity of the assumptions underlying the ANOVA model was tested before statistical tests were applied.

7.1.1 Summary of Comparison of Novices’ versus Experts’ Judgments with ESS Support

Tables 7-1 and 7-2 summarize the analyses of expert versus novice judgments with ESS analyses and explanations support provided. Sections 7.1.2 to 7.1.7 present the detailed results.

<table>
<thead>
<tr>
<th>Section #</th>
<th>Consistency with Original Experts’ Judgments</th>
<th>Hypothesis</th>
<th>Supported?</th>
</tr>
</thead>
<tbody>
<tr>
<td>7.1.2</td>
<td>Individual Pre Judgment</td>
<td>Novice&lt;Expert</td>
<td>No</td>
</tr>
<tr>
<td>7.1.3</td>
<td>Group Judgment</td>
<td>Novice&gt;Expert</td>
<td>Yes</td>
</tr>
<tr>
<td>7.1.4</td>
<td>Group Judgment - Individual Pre Judgment</td>
<td>Novice&gt;Expert</td>
<td>Yes</td>
</tr>
<tr>
<td>7.1.5</td>
<td>Individual Post Judgment</td>
<td>Novice&gt;Expert</td>
<td>No</td>
</tr>
<tr>
<td>7.1.6</td>
<td>Individual Post - Individual Pre Judgment</td>
<td>Novice&gt;Expert</td>
<td>Yes</td>
</tr>
</tbody>
</table>

Table 7-1: Summary of Results on Consistency of Experts’ versus Novices’ Judgments with Respect to Judgments of Original Experts
The analyses presented in the next few sections indicate that, with ESS analyses and explanations support, the novices made judgments that were closer to those of the original experts than the experts did. This finding indicates that the novices were more willing to accept the analyses and explanations given by the ESS than the experts. In other words, knowledge transfer from the ESS to the users was more successful in the case of the novices than the experts. In addition, consensus was significantly higher among members in the novice group than the expert group.

7.1.2 Pre-test — Experts’ versus Novices’ Initial Judgments

The experts’ and novices’ individual pre-discussion judgments were compared with respect to the judgments of the original experts. The analysis was carried out at both the individual and group levels to find out if there is a difference in their initial judgments with respect to the judgments of the original experts. The hypothesis to be tested is:

\[ H_{3a} \]: The initial judgments of the experts (that are made without any form of ESS support) are more similar to the judgments of the original experts than it is for the initial judgments of the novices.

Evaluation of Assumptions of Parametric Model

*Individual as Level of Analysis.* The absolute deviation of individual pre-discussion judgments from consensus judgments of the expert and novice groups that were provided
with full ESS support satisfy the normality, homogeneity of variances and independence assumptions.\textsuperscript{1} As such, both the \( t \) test and the Mann-Whitney \( U \) test were used to perform the analysis at the individual level. The Mann-Whitney \( U \) test would also verify the results of the \( t \) test.

**Group as Level of Analysis.** The absolute deviation of individual pre-discussion judgments from consensus judgments of the expert and novice groups that were provided with full ESS support was also assessed at the group level, by averaging the deviation measure across the three members in a group. The normality, homogeneity of variances, and independence assumptions are all met. Therefore, both the \( t \) test and the Mann-Whitney \( U \) test were also used to perform the analysis at the group level.

**Comparison of Initial Individual Judgments of Experts versus Novices**

The descriptive statistics of the absolute deviation of individual pre-discussion judgments from consensus judgments of the expert and novice groups are shown in Table 7-3. Table 7-4 shows the equivalent descriptive statistics aggregated at the group level.

\textsuperscript{1} The normality probability plot, the detrended normal probability plot, the measures of kurtosis, the Kolmogorov-Smirnov (Lilliefors) statistics, and the Shapiro-Wilks’ statistics support the normality assumption. The Levene statistic is .26 (\( p=.62 \)), indicating that the homogeneity of variances assumption is not violated. The independence assumption is satisfied from the experimental procedure where subjects were randomly assigned to groups and then to treatments. Although the measure of skewness for the expert group is .91 (\( p=.05 \)), it is not considered a severe violation of normality according to Stevens’ (1996) recommendation, which is to adopt a more stringent alpha level to control for overall type I error rate.
Table 7-3: Descriptive Statistics of Consistency of Individual Pre-discussion Judgments of Experts versus Novices

<table>
<thead>
<tr>
<th>Source</th>
<th>N</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Std. Error</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experts/Professionals</td>
<td>17</td>
<td>8.47</td>
<td>3.48</td>
<td>.85</td>
<td>3</td>
<td>17</td>
</tr>
<tr>
<td>Novices/Students</td>
<td>27</td>
<td>9.15</td>
<td>3.48</td>
<td>.67</td>
<td>3</td>
<td>15</td>
</tr>
<tr>
<td>Total</td>
<td>44</td>
<td>8.89</td>
<td>3.46</td>
<td>.52</td>
<td>3</td>
<td>17</td>
</tr>
</tbody>
</table>

Table 7-4: Descriptive Statistics of Consistency of Individual Pre-discussion Judgments of Experts versus Novices Analyzed at the Group Level

<table>
<thead>
<tr>
<th>Source</th>
<th>N</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Std. Error</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experts/Professionals</td>
<td>5</td>
<td>8.73</td>
<td>1.84</td>
<td>.83</td>
<td>7.00</td>
<td>11.33</td>
</tr>
<tr>
<td>Novices/Students</td>
<td>9</td>
<td>9.15</td>
<td>1.59</td>
<td>.53</td>
<td>6.00</td>
<td>10.67</td>
</tr>
<tr>
<td>Total</td>
<td>14</td>
<td>9.00</td>
<td>1.63</td>
<td>.43</td>
<td>6.00</td>
<td>11.33</td>
</tr>
</tbody>
</table>

Although the experts' judgments are closer to those of the original experts (i.e., mean deviation of experts' judgments from original experts' consensus judgments is smaller than the mean deviation of novices' judgments from original experts' consensus judgments — refer to Tables 7-3 and 7-4), the difference is insignificant when analyzed at the individual level (t test: p=.53 (2-tailed); Mann-Whitney U test: p=.45 (2-tailed)) and at the group level (t test: p=.67 (2-tailed); Mann-Whitney U test: p=.55 (2-tailed)). Table 7-5 summarizes the results. Tables 7-6 and 7-7 show the detailed results of the analysis.

---

2 Only 17 instead of 18 (6x3) cases were considered because one of the subjects did not specify one of his/her individual pre-discussion judgments.
3 Only 5 cases were considered because one of the subjects did not specify one of his/her individual pre-discussion judgments.
<table>
<thead>
<tr>
<th>Test \ Level of Analysis</th>
<th>Individual</th>
<th>Group</th>
</tr>
</thead>
<tbody>
<tr>
<td>Difference between Means ($t$)</td>
<td>$p=0.53$</td>
<td>$p=0.67$</td>
</tr>
<tr>
<td>Mann-Whitney ($U$)</td>
<td>$p=0.45$</td>
<td>$p=0.55$</td>
</tr>
</tbody>
</table>

*Table 7-5: Results of Analysis of Consistency of Individual Pre-discussion Judgments of Experts versus Novices*

<table>
<thead>
<tr>
<th>Source</th>
<th>Level of Analysis</th>
<th>$t$</th>
<th>DF</th>
<th>$p$-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Expertise</td>
<td>Individual</td>
<td>0.63</td>
<td>42</td>
<td>0.53</td>
</tr>
<tr>
<td>Expertise</td>
<td>Group</td>
<td>0.42</td>
<td>12</td>
<td>0.67</td>
</tr>
</tbody>
</table>

*Table 7-6: Results of $t$ Test – Consistency of Individual Pre-discussion Judgments*

<table>
<thead>
<tr>
<th>Level of Analysis</th>
<th>Treatment (I)</th>
<th>Treatment (J)</th>
<th>N</th>
<th>Mean Rank</th>
<th>Sum of Ranks</th>
<th>Mann-Whitney $U$</th>
<th>$p$-value (2-tailed)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Individual</td>
<td>Experts/Professionals</td>
<td>17</td>
<td>20.68</td>
<td>351.50</td>
<td>198.50</td>
<td>.45</td>
<td></td>
</tr>
<tr>
<td>Novices/Students</td>
<td>27</td>
<td>23.65</td>
<td>638.50</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Group</td>
<td>Experts/Professionals</td>
<td>5</td>
<td>6.60</td>
<td>33.00</td>
<td>18.00</td>
<td>.55</td>
<td></td>
</tr>
<tr>
<td>Novices/Students</td>
<td>9</td>
<td>8.00</td>
<td>72.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Table 7-7: Results of Mann-Whitney $U$ Test – Consistency of Individual Pre-discussion Judgments*

There is no significant difference in the consistency of initial individual judgments between the experts and novices who participated in the study. This is probably because the novices, like the experts, are knowledgeable in the task domain; however, unlike the experts, the novices lack experience. The results of the pre-test suggests that the level of knowledge of the novice subjects is comparable to that of the professional experts who participated in the study.
7.1.3 Group Judgments

The absolute deviations of the group judgments of the experts and novices from the consensus judgments of the original experts were compared to find out if the ESS analyses and explanations support have greater influence on the novices than the experts. The hypothesis that is being tested is:

\( H_{g3}: \) The group judgments of the experts are farther away from the judgments of the original experts than it is for the group judgments of the novices.

Evaluation of Assumptions of Parametric Model

The absolute deviations of the experts’ and novices’ group judgments from the original experts’ consensus judgments satisfy the homogeneity of variances and independence assumptions, but they hardly satisfy the normality assumption.\(^4\) As such, the Mann-Whitney \( U \) test was used to compare the group judgments of the experts and novices with respect to the original experts’ consensus judgments. However, if Stevens’ (1996) recommendation, which is to adopt a more stringent alpha level (e.g., .01) to control for type I error rate, is considered, the distributions would be acceptable. Thus, the \( t \) test was also used for the analysis.

Comparison of Consistency of Group Judgments Across Expert and Novice Groups

Table 7-8 presents the descriptive statistics of the absolute deviation of the group judgments of the expert and novice groups from the original experts’ consensus judgments.

\(^4\) The Levene statistic of .73 (\( p = .41 \)) indicates that the homogeneity of variances assumption is not violated. The independence assumption is satisfied from the experimental procedure. However, the Kolmogorov-Smirnov (Lillefors) and the Shapiro-Wilk statistics for the novice group are .26 (\( p = .08 \)) and .82 (\( p = .05 \)) respectively, suggesting that the distribution is not normal. In addition, the skewness statistic of the expert group is -1.35 (\( p = .05 \)), indicating that the distribution is skewed toward the right.
The descriptive statistics presented in Table 7-8 indicate that the ESS analyses and explanations support have greater influence on the novice groups than the expert groups.

The results of the non-parametric Mann-Whitney $U$ test, as shown in Table 7-9, indicate a significant difference in the deviation of the group judgments of the expert and novice groups ($p=.08$; 2-tailed; $p=.04$; 1-tailed) from the judgments of the original experts. The mean ranks are also presented in Table 7-9.

Although the distributions for both the expert and novice groups do not satisfy the normality assumption very well, the $t$ test produces a $p$-value of .08 (2-tailed) (see Table 7-9) which is similar to the result of the Mann-Whitney $U$ test (refer to Table 7-10).
Table 7-10: Results of t Test — Consistency of Group Judgments of Experts versus Novices

<table>
<thead>
<tr>
<th>Source</th>
<th>t</th>
<th>DF</th>
<th>p-value (2-tailed)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Expertise</td>
<td>-1.93</td>
<td>13</td>
<td>.08</td>
</tr>
</tbody>
</table>

The above results indicate that the novices were more influenced by the ESS analyses and explanations support than the experts were. This provides some indication that, with ESS analyses and explanations support, the novices, being more receptive to the ESS advice and explanations, may be able to surpass the experts in performance.

Further analyses indicate that of the six judgments analyzed, three of them differed between the experts and novices, with one of them (Question 4) having a greater impact than the other two. The experts and novices differed in their rating on the value of Canacom’s (i.e., the borrowing company’s) stock as loan collateral (Question 4). A more detailed analysis of Question 4 indicates that many of the subjects had difficulties with the rating because of the economics of the transaction. They recognized that it is when the company defaults on its loan that the collateral will come into play, in which case the stock of the company would also be of low value. This limitation of using stock as collateral was identified particularly by the experts, as reflected by their lower ratings. The experts may have been more aware of this issue because it was typically against the policy of their financial institutions to accept the stock of a borrowing company as collateral on its loan. This may explain the difference in the rating of Question 4 by the experts and novices.

7.1.4 Change in Deviation Score from Individual Pre-discussion to Group Judgments

The change in the deviation score from the individual pre-discussion to the group judgments was assessed at the group level by comparing it across the expert and novice groups.
Evaluation of Assumptions of Parametric Model

The change in the deviation score from the individual pre-discussion (group average) to the group judgments was assessed at the group level. The distributions for both the expert and novice groups satisfy the normality, homogeneity of variances, and independence assumptions.\(^5\) However, since the normal probability plot does not show an exact straight line, we question whether the assumptions of the parametric model have been completely met. As such, the results of the Mann-Whitney \(U\) test were used to verify the results of the \(t\) test.

Comparison of Change in Deviation Score from Pre-discussion to Group Judgments

The descriptive statistics of the change in the deviation score from the individual pre-discussion (group average) to the group judgments are shown in Table 7-11.

<table>
<thead>
<tr>
<th>Source</th>
<th>N</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Std. Error</th>
<th>Min</th>
<th>Max</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experts/Professionals(^6)</td>
<td>5</td>
<td>1.73</td>
<td>1.51</td>
<td>.50</td>
<td>-1.00</td>
<td>3.67</td>
<td>4.67</td>
</tr>
<tr>
<td>Novices/Students</td>
<td>9</td>
<td>4.26</td>
<td>1.93</td>
<td>.87</td>
<td>2.33</td>
<td>6.33</td>
<td>4.00</td>
</tr>
<tr>
<td>Total</td>
<td>15</td>
<td>3.36</td>
<td>2.03</td>
<td>.54</td>
<td>-1.00</td>
<td>6.33</td>
<td>7.33</td>
</tr>
</tbody>
</table>

Table 7-11: Descriptive Statistics of Change in Deviation Score from Individual Pre-discussion to Group Judgments Analyzed at the Group Level

With ESS analyses and explanations support, the deviation of the novices' judgments from those of the original experts decreased by a magnitude of 4.26 while that of the experts decreased by a magnitude of 1.73.

\(^5\) Although the normal probability plots do not fall closely on a straight line, the measures of skewness and kurtosis, and the Kolmogorov-Smirnov (Lillefors) and Shapiro-Wilk statistics for both the expert and novice groups support the normality assumption. The Levene statistic of .23 (\(p=.64\)) also supports the homogeneity of variances assumption.

\(^6\) Only 5 cases were considered because one of the subjects did not specify one of his/her individual pre-discussion judgments.
decreased by a magnitude of only 1.73. The descriptive statistics, therefore, indicate that the ESS analyses and explanations support have greater influence on the novices than the experts.

The results of the analysis of the change in the deviation score from the individual pre-discussion to the group judgments of the expert and novice groups are summarized in Table 7-12, and presented in greater detail in Tables 7-13 and 7-14.

<table>
<thead>
<tr>
<th>Test</th>
<th>p-value (2-tailed)</th>
<th>p-value (1-tailed)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Difference between Means (t)</strong></td>
<td>p=.02</td>
<td>p=.01</td>
</tr>
<tr>
<td><strong>Mann-Whitney (U)</strong></td>
<td>p=.05</td>
<td>p=.03</td>
</tr>
</tbody>
</table>

*Table 7-12: Results of Analysis of Change in Deviation Score from Individual Pre-discussion to Group Judgments of Experts versus Novices*

<table>
<thead>
<tr>
<th>Source</th>
<th>t</th>
<th>DF</th>
<th>p-value (2-tailed)</th>
<th>p-value (1-tailed)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Expertise</td>
<td>2.72</td>
<td>12</td>
<td>.02</td>
<td>.01</td>
</tr>
</tbody>
</table>

*Table 7-13: Results of t Test – Change in Deviation Score from Individual Pre-discussion to Group Judgments*

<table>
<thead>
<tr>
<th>Source</th>
<th>N</th>
<th>Mean Rank</th>
<th>Sum of Ranks</th>
<th>Mann-Whitney U</th>
<th>p-value (2-tailed)</th>
<th>p-value (1-tailed)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experts/Professionals</td>
<td>5</td>
<td>4.60</td>
<td>23.00</td>
<td>8.00</td>
<td>.05</td>
<td>.03</td>
</tr>
<tr>
<td>Novices/Students</td>
<td>9</td>
<td>9.11</td>
<td>82.00</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Table 7-14: Results of Mann-Whitney U Test – Change in Deviation Score from Individual Pre-discussion to Group Judgments*
When the change in the deviation score from the individual pre-discussion to the group judgments was compared across the expert and novice groups, a significant difference is found \( t \text{ test: } p=.02 \text{ (2-tailed), } p=.01 \text{ (1-tailed); Mann-Whitney } U \text{ test: } p=.05 \text{ (2-tailed), } p=.03 \text{ (1-tailed)} \). The analyses indicate that the ESS analyses and explanations support more greatly influence the novices than the experts.

7.1.5 Individual Post-discussion Judgments

The absolute deviation of individual post-discussion judgments from the original experts’ consensus judgments can be analyzed at both the group and individual levels. Group mean is used to analyze the results when the level of analysis is group (Stevens, 1996). The nested design can be used to analyze the deviation or consistency measure at the individual level if the assumptions of the ANOVA model are satisfied (Ager and Anderson, 1978). The hypothesis being tested is:

\[ H_{3b}: \text{ The final individual judgments of the experts are farther away from the judgments of the original experts than it is for the final individual judgments of the novices.} \]

Evaluation of Assumptions of ANOVA Model

Individual as Level of Analysis. The normality of residuals assumption of the nested design was not satisfied.\(^7\) As such, the nested design cannot be and was not used for the analysis.

Group as Level of Analysis. The absolute deviation of individual post-discussion judgments from consensus judgments was analyzed at the group level by computing it for each individual subject and averaging it across the three members in a group. The distributions for both the expert and novice groups satisfy the normality and independence assumptions.

\(^7\) Although the homogeneity of variances assumption is not violated, the Kolmogorov-Smirnov (Lillefors) statistic for the novice group is .00, thus violating the normality assumption.
assumptions, but not the homogeneity of variances assumption. As such, only the Mann-Whitney $U$ test was used to analyze the results.

Comparison of Experts' versus Novices' Individual Post-discussion Judgments

The absolute deviation of the individual post-discussion judgments from the original experts' consensus judgments was assessed at the group level by averaging it across the three individuals in a group. The descriptive statistics are shown in Table 7-15.

<table>
<thead>
<tr>
<th>Source</th>
<th>N</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Std. Error</th>
<th>Min</th>
<th>Max</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experts/Professionals</td>
<td>6</td>
<td>6.83</td>
<td>1.21</td>
<td>.49</td>
<td>5.00</td>
<td>8.67</td>
<td>3.67</td>
</tr>
<tr>
<td>Novices/Students</td>
<td>9</td>
<td>5.41</td>
<td>2.05</td>
<td>.68</td>
<td>3.00</td>
<td>8.67</td>
<td>5.67</td>
</tr>
<tr>
<td>Total</td>
<td>15</td>
<td>5.98</td>
<td>1.85</td>
<td>.48</td>
<td>3.00</td>
<td>8.67</td>
<td>5.67</td>
</tr>
</tbody>
</table>

Table 7-15: Descriptive Statistics of Consistency of Individual Post-Discussion Judgments of Experts versus Novices Analyzed at the Group Level

A preliminary analysis of the descriptive statistics in Table 7-15 indicates that, with group discussion and ESS analyses and explanations support, the novices made individual judgments that were closer to those of the original experts than the judgments of the experts were. In addition, the standard deviation and range of novices' performance were larger than those of the experts.

The results of the Mann-Whitney $U$ test and the mean ranks of the experts' and novices' judgments are shown in Table 7-16. The results indicate that the difference is not significant. In other words, the novices and experts did not differ in the deviation of their individual judgments from the consensus judgments of the original experts.

---

$^8$ The Levene statistic of $5.36 (p=.04)$ indicates that the homogeneity of variances assumption is violated.
7.1.6 Change in Deviation Score from Initial to Post-discussion Individual Judgments

The change in deviation score from initial individual to post-discussion individual judgments was analyzed at both the individual and group levels. The validity of the assumptions underlying the ANOVA model was tested before statistical tests were applied.

**Evaluation of Assumptions of Parametric Model**

*Individual as Level of Analysis.* When the change in the deviation score from the individual pre-discussion to the group judgments was assessed at the *individual* level, the distributions for both the expert and novice groups satisfy the normality, homogeneity of variances, and independence assumptions. Therefore, the nested design was used for the analysis.

*Group as Level of Analysis.* The distributions of the change in the deviation score from the initial individual judgments to the post-discussion individual judgments of both the expert and novice groups satisfy the normality, homogeneity of variances, and independence assumptions when assessed at the group level. Thus, both the *t* and the Mann-Whitney *U* tests were used for the analysis.

---

9 Although the detrended normal probability plot does not show a random pattern (i.e., deviation from normal distribution increases with observed value), the measures of skewness and kurtosis, and the Kolmogorov-Smirnov (Lillefors) and Shapiro-Wilk statistics for both the expert and novice groups support the normality assumption. The Levene statistic of 1.06 (*p*=.31) also supports the homogeneity of variances assumption.

10 The Levene statistic (=1.18; *p*=.30) supports the homogeneity of variances assumption. Although the distribution for the expert group is skewed slightly toward the right (skewness measure=−.13; *p*=.08) and the
Comparison of Change in Deviation from Individual Pre-to Post-discussion Judgments

The descriptive statistics of change in deviation score from individual pre-to post-discussion judgments are shown in Tables 7-17 (individual level) and 7-19 (group level).

**Individual Level**

<table>
<thead>
<tr>
<th>Source</th>
<th>N</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Std. Error</th>
<th>Min</th>
<th>Max</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experts/Professionals</td>
<td>17</td>
<td>1.82</td>
<td>2.83</td>
<td>.69</td>
<td>-3</td>
<td>6</td>
<td>9</td>
</tr>
<tr>
<td>Novices/Students</td>
<td>27</td>
<td>3.74</td>
<td>3.54</td>
<td>.68</td>
<td>-3</td>
<td>11</td>
<td>14</td>
</tr>
<tr>
<td>Total</td>
<td>44</td>
<td>3.00</td>
<td>3.38</td>
<td>.51</td>
<td>-3</td>
<td>11</td>
<td>14</td>
</tr>
</tbody>
</table>

Table 7-17: Descriptive Statistics of Change in Deviation Score from Individual Pre-to Post-discussion Judgments of Experts versus Novices Analyzed at the Individual Level

The descriptive statistics in Table 7-17 indicate that novices' judgments shifted toward the consensus judgments of the original experts more than those of the experts (3.74 versus 1.82). The results of the nested design (see Table 7-18) indicate that the difference is significant. In other words, with group discussion and ESS analyses and explanations support, the novices' judgments shifted significantly more toward the consensus judgments of the original experts than the corresponding judgment shift of the experts.

normality probability plot for the expert group does not fall very closely on a straight line, the Kolmogorov-Smirnov (Lilliefors) and Shapiro-Wilk statistics for both distributions support the normality assumption.

11 Only 17 cases were considered because one of the subjects did not specify one of his/her individual pre-discussion judgments.
Table 7-18: Results of Nested ANOVA — Change in Deviation Score from Individual Pre- to Post-discussion Judgments of Experts versus Novices

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>DF</th>
<th>MS</th>
<th>F</th>
<th>p-value (2-tailed)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Expertise</td>
<td>38.34</td>
<td>1</td>
<td>38.34</td>
<td>4.86</td>
<td>.05</td>
</tr>
<tr>
<td>Group within Treatment</td>
<td>102.49</td>
<td>13</td>
<td>7.88</td>
<td>.65</td>
<td>.79</td>
</tr>
<tr>
<td>Error</td>
<td>351.17</td>
<td>29</td>
<td>12.11</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Group Level*

Table 7-19: Descriptive Statistics of Change in Deviation Score from Individual Pre- to Post-discussion Judgments of Experts versus Novices Analyzed at the Group Level

<table>
<thead>
<tr>
<th>Source</th>
<th>N</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Std. Error</th>
<th>Min</th>
<th>Max</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experts/Professionals</td>
<td>5</td>
<td>1.87</td>
<td>2.19</td>
<td>.98</td>
<td>-1.67</td>
<td>4.00</td>
<td>5.67</td>
</tr>
<tr>
<td>Novices/Students</td>
<td>9</td>
<td>3.74</td>
<td>1.36</td>
<td>.45</td>
<td>2.00</td>
<td>6.00</td>
<td>4.00</td>
</tr>
<tr>
<td>Total</td>
<td>14</td>
<td>3.07</td>
<td>1.87</td>
<td>.50</td>
<td>-1.67</td>
<td>6.00</td>
<td>7.67</td>
</tr>
</tbody>
</table>

The descriptive statistics in Tables 7-19 and 7-17 are similar; they only differ in their unit of analysis. They both indicate that the shift toward the consensus judgments of the original experts was greater for the novices than the experts.

The results of the analysis are summarized in Table 7-20, and presented in greater detail in Tables 7-21 and 7-22.

---

12 Only 5 cases were considered because one of the subjects did not specify one of his/her individual pre-discussion judgments.
The change in deviation score from individual pre- to post-discussion judgments was analyzed both at the individual and group levels. The results are significant for both analyses (refer to Table 7-20). The results indicate that the ESS analyses and explanations support influenced the novices more than the experts.
7.1.7 Consensus in Judgments

Consensus was measured in two ways: (1) consensus among individual judgments -- the total distance between every two group members' post-discussion judgments; (2) consensus with group judgments -- the distance between the group judgments and its members’ post-discussion individual judgments.

7.1.7.1 Consensus Among Group Members' Individual Judgments

The absolute distance between the group members’ post-discussion individual judgments was computed and compared across the expert and novice groups. The hypothesis to be tested is:

\[ H_{ij} \text{: Experts' consensus among individual judgments is lower than novices' consensus among individual judgments.} \]

Evaluation of Assumptions of Parametric Model

Since the distribution for the novice group violates the normality assumption\(^{13}\), only the Mann-Whitney U test was used for the analysis.

Comparison of Consensus Among Individual Judgments Across Experimental Conditions

The descriptive statistics of the total absolute distance between the three group members’ post-discussion individual judgments are presented in Table 7-23.

---

\(^{13}\) The Shapiro-Wilk statistic for the novice group \((= .78; p = .02)\) indicates that the distribution \textit{violates the normality assumption}. The skewness \((= 2.05; p = .00)\) and kurtosis \((= 5.08; p = .00)\) measures also indicate non-normality of the distribution.
Table 7-23: Descriptive Statistics of Total Absolute Distance between Group Members’ Post-discussion Individual Judgments Analyzed at the Group Level

Table 7-24 shows the results of the Mann-Whitney $U$ test and the mean ranks of the absolute distance between the three group members’ post-discussion individual judgments. The results indicate that the level of consensus of the expert and novice groups differ significantly. The novices achieved a much higher level of consensus in their final individual judgments than the experts.

Table 7-24: Results of Mann-Whitney U Test – Total Absolute Distance between Group Members’ Post-discussion Individual Judgments

7.1.7.2 Consensus with Group Judgments

The absolute distance of the group members’ post-discussion individual judgments from group judgments was computed and compared across the expert and novice groups.

$H_{g4}$: Experts’ consensus with group judgments is lower than novices’ consensus with group judgments.
Evaluation of Assumptions of ANOVA Model

Since the distribution for the novice group violates the normality assumption\(^\text{14}\), only the Mann-Whitney \(U\) test was used for the analysis.\(^\text{15}\)

Comparison of Consensus with Group Judgments Across Experimental Conditions

The descriptive statistics of the total absolute distance of the three group members' post-discussion individual judgments from the group judgments are presented in Table 7-25.

<table>
<thead>
<tr>
<th>Source</th>
<th>N</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Std. Error</th>
<th>Min</th>
<th>Max</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experts/Professionals</td>
<td>6</td>
<td>9.67</td>
<td>1.63</td>
<td>.67</td>
<td>8</td>
<td>12</td>
<td>4</td>
</tr>
<tr>
<td>Novices/Students</td>
<td>9</td>
<td>4.89</td>
<td>4.11</td>
<td>1.37</td>
<td>1</td>
<td>15</td>
<td>14</td>
</tr>
<tr>
<td>Total</td>
<td>15</td>
<td>6.80</td>
<td>4.06</td>
<td>1.05</td>
<td>1</td>
<td>15</td>
<td>15</td>
</tr>
</tbody>
</table>

Table 7-25: Descriptive Statistics of Total Absolute Distance of Group Members' Post-discussion Individual Judgments from Group Judgments Analyzed at the Group Level

Table 7-26 shows the results of the Mann-Whitney \(U\) test and the mean ranks of the absolute distance of the three group members' post-discussion individual judgments from group judgments.

\(^{14}\) The Shapiro-Wilk statistic for the novice group (=.78; \(p=.02\)) indicates that the distribution \textit{violates the normality assumption}. The skewness (=2.05; \(p=.00\)) and kurtosis (=5.08; \(p=.00\)) measures also indicate non-normality of the distribution.

\(^{15}\) The Kolmogorov-Smirnov (Lillefors) (=.28; \(p=.04\)) and the Shapiro-Wilk (=.76; \(p=.01\)) statistics indicate that the distribution of the total absolute distance of the group members' post-discussion individual judgments from group judgments for the novice group \textit{violates the normality assumption}.
Table 7-26: Results of Mann-Whitney U Test — Total Absolute Distance of Group Members’ Post-discussion Individual Judgments from Group Judgments

<table>
<thead>
<tr>
<th>Source</th>
<th>N</th>
<th>Mean Rank</th>
<th>Sum of Ranks</th>
<th>Mann-Whitney U</th>
<th>p-value (2-tailed)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experts/Professionals</td>
<td>6</td>
<td>11.50</td>
<td>69.00</td>
<td>6.00</td>
<td>.01</td>
</tr>
<tr>
<td>Novices/Students</td>
<td>9</td>
<td>5.67</td>
<td>51.00</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The results in Table 7-26 indicate that the expert and novice groups differ significantly in their level of consensus with their group judgments. The novices were at a much higher level of consensus with their group judgments than the experts.

7.1.8 Summary of Results

The novices in our study were final year undergraduate students majoring in Accounting, or graduate or final year undergraduate students who have taken the Financial Statement Analysis course. The experts in our study were professional financial analysts whose major responsibilities include making commercial loan decisions on a daily basis. The pre-tests indicate that both groups of subjects were comparable in the deviation of their initial individual judgments (i.e., made with no ESS analyses or explanations support) from the consensus judgments of the original experts. However, with ESS analyses and explanations support, the judgments of the novices were much closer to those of the original experts than the judgments of the experts were. This indicates that the novices were more likely to be influenced by ESS analyses and explanations support than the experts. Lastly, the novices achieved a higher level of consensus in their group and final individual judgments than the experts.
7.2 Comparison of Novices' versus Experts' Perceptions with Full ESS Support

The perceptions of the novice versus expert subjects that were provided with the complete ESS support (i.e., with explanations) were compared. The validity of the assumptions underlying the nested ANOVA model was tested before statistical tests were applied. If the assumptions of the nested ANOVA model are violated, these measures will be averaged across the group and analyzed at the group level.

7.2.1 Summary of Comparison of Novices' versus Experts' Perceptions

Table 7-27 summarizes the results of comparison of the novices' versus experts' perceptions where ESS and its explanations support were provided. Sections 7.2.2 to 7.2.5 present the detailed results.

<table>
<thead>
<tr>
<th>Section #</th>
<th>Perceptions of Novices</th>
<th>Hypothesis</th>
<th>Supported?</th>
</tr>
</thead>
<tbody>
<tr>
<td>7.2.2</td>
<td>Satisfaction with Group Process</td>
<td>Novice&gt;Expert</td>
<td>No</td>
</tr>
<tr>
<td>7.2.3</td>
<td>Satisfaction with Group Judgments</td>
<td>Novice&gt;Expert</td>
<td>No</td>
</tr>
<tr>
<td>7.2.4</td>
<td>Perceived Usefulness of ESS</td>
<td>Novice&gt;Expert</td>
<td>Yes</td>
</tr>
<tr>
<td>7.2.5</td>
<td>Trust in ESS</td>
<td>Novice&gt;Expert</td>
<td>No</td>
</tr>
</tbody>
</table>

Table 7-27: Summary of Results of Novices' versus Experts' Perceptions

Novices and experts who used the ESS analyses and explanations support for group decision making did not differ in their level of satisfaction with the group process and the group judgments. Although novices and experts did not differ in their level of trust in the ESS, the novices found the ESS and its explanations support to be more useful than the experts did.
7.2.2 Perception Measure — Satisfaction with Group Process

The perceived satisfaction with group process of the expert and novice groups that were provided with the complete ESS support (i.e., with explanations) was compared. The hypothesis to be tested is:

\[ H_{p6}: \text{The experts are less satisfied with the group decision making process than the novices.} \]

Evaluation of Assumptions of ANOVA Model

Testing Assumptions of Nested ANOVA Design with Individual as Level of Analysis. The distributions of the residuals do not satisfy the normality assumption. Although the Levene statistic of 1.46 (\( p=.23 \)) indicates that the homogeneity of variances assumption is not violated, the Kolmogorov-Smirnov (Lillefors) and the Shapiro-Wilk statistics for both groups suggest that the distributions do not satisfy the normality assumption. As such, the use of \( t \) and Mann-Whitney \( U \) tests were considered next.

Testing Assumptions of \( t \) Test with Group as Level of Analysis. The novice and expert subjects' perceived satisfaction with the group process was assessed at the group level by averaging their individual measures across the three members in a group. In this case, the normality, homogeneity of variances, and independence assumptions are satisfied. The Kolmogorov-Smirnov (Lillefors) and the Shapiro-Wilks' statistics support the normality assumption, and the Levene statistic (\( =.56; p=.47 \)) supports the homogeneity of variances assumption. Therefore, both the \( t \) and the Mann-Whitney \( U \) tests were used to assess the subjects' perceived satisfaction with group process.
Comparison of Novice Subjects’ Satisfaction with Group Process Across Treatments

The descriptive statistics of the novice and expert subjects’ satisfaction with the group process are presented in Table 7-28 (with individual as the unit of analysis), and Table 7-29 (with group as the unit of analysis).

<table>
<thead>
<tr>
<th>Source</th>
<th>N</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Std. Error</th>
<th>Min</th>
<th>Max</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experts/Professionals</td>
<td>18</td>
<td>19.89</td>
<td>5.33</td>
<td>1.26</td>
<td>7</td>
<td>27</td>
<td>20</td>
</tr>
<tr>
<td>Novices/Students</td>
<td>27</td>
<td>21.11</td>
<td>4.09</td>
<td>.79</td>
<td>12</td>
<td>28</td>
<td>16</td>
</tr>
<tr>
<td>Total</td>
<td>45</td>
<td>20.62</td>
<td>4.61</td>
<td>.69</td>
<td>7</td>
<td>28</td>
<td>21</td>
</tr>
</tbody>
</table>

*Table 7-28: Descriptive Statistics of Perceived Satisfaction with Group Process (Individual Level)*

<table>
<thead>
<tr>
<th>Source</th>
<th>N</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Std. Error</th>
<th>Min</th>
<th>Max</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experts/Professionals</td>
<td>6</td>
<td>19.89</td>
<td>3.34</td>
<td>1.37</td>
<td>15.33</td>
<td>24.33</td>
<td>9.00</td>
</tr>
<tr>
<td>Novices/Students</td>
<td>9</td>
<td>21.11</td>
<td>2.64</td>
<td>.88</td>
<td>17.33</td>
<td>26.67</td>
<td>9.33</td>
</tr>
<tr>
<td>Total</td>
<td>15</td>
<td>20.62</td>
<td>2.89</td>
<td>.75</td>
<td>15.33</td>
<td>26.67</td>
<td>11.33</td>
</tr>
</tbody>
</table>

*Table 7-29: Descriptive Statistics of Perceived Satisfaction with Group Process (Group Level)*

Table 7-30 summarizes the results of the analysis, which are presented in greater detail in Tables 7-31 and 7-32.

<table>
<thead>
<tr>
<th>Test \ Level of Analysis</th>
<th>Group</th>
</tr>
</thead>
<tbody>
<tr>
<td>Difference between Means (t)</td>
<td>p=.44</td>
</tr>
<tr>
<td>Mann-Whitney (U)</td>
<td>p=.52</td>
</tr>
</tbody>
</table>

*Table 7-30: Results of Analysis of Novices’ versus Experts’ Satisfaction with Group Process*
Table 7-31: Results of t Test – Satisfaction with Group Process

<table>
<thead>
<tr>
<th>Source</th>
<th>t</th>
<th>DF</th>
<th>p-value (2-tailed)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Expertise</td>
<td>.79</td>
<td>13</td>
<td>.44</td>
</tr>
</tbody>
</table>

Table 7-32: Results of Mann-Whitney U Test – Satisfaction with Group Process

The novice and expert subjects’ perception of satisfaction with group process did not differ (t test: $p=.44$; Mann-Whitney U test: $p=.52$).

7.2.3 Perception Measure – Satisfaction with Group Judgments

The perceived satisfaction with group judgments of the expert and novice groups that were provided with the complete ESS support (i.e., with explanations) was compared. The hypothesis to be tested is:

$H_{p5}$: The experts are less satisfied with the group judgments than the novices.

Testing Assumptions of Nested ANOVA Design with Individual as Level of Analysis. The distributions of the residuals satisfy the normality, homogeneity of variances, and independence assumptions. The Kolmogorov-Smirnov (Lillefors) and the Shapiro-Wilk statistics support the normality assumption, and the Levene statistic ($=.15; p=.70$) supports the homogeneity of variances assumption. As such, it is not necessary to carry out the analysis at the group level.
Comparison of Novice versus Expert Subjects' Satisfaction with Group Judgments

The descriptive statistics of the novice versus expert subjects' satisfaction with group judgments are presented in Table 7-33 (with individual as the level of analysis), and the results of the analysis are presented in Table 7-34.

<table>
<thead>
<tr>
<th>Source</th>
<th>N</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Std. Error</th>
<th>Min</th>
<th>Max</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experts/Professionals</td>
<td>18</td>
<td>45.39</td>
<td>3.71</td>
<td>.87</td>
<td>36</td>
<td>54</td>
<td>18</td>
</tr>
<tr>
<td>Novices/Students</td>
<td>27</td>
<td>44.26</td>
<td>4.82</td>
<td>.93</td>
<td>32</td>
<td>52</td>
<td>20</td>
</tr>
<tr>
<td>Total</td>
<td>45</td>
<td>44.71</td>
<td>4.40</td>
<td>.66</td>
<td>32</td>
<td>54</td>
<td>22</td>
</tr>
</tbody>
</table>

Table 7-33: Descriptive Statistics of Satisfaction with Group Judgments (Individual Level)

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>DF</th>
<th>MS</th>
<th>F</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Expertise</td>
<td>13.78</td>
<td>1</td>
<td>13.78</td>
<td>.43</td>
<td>.52</td>
</tr>
<tr>
<td>Group within Treatment</td>
<td>413.46</td>
<td>13</td>
<td>31.80</td>
<td>2.24</td>
<td>.03</td>
</tr>
<tr>
<td>Error</td>
<td>426.00</td>
<td>30</td>
<td>14.20</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 7-34: Results of Nested ANOVA – Satisfaction with Group Judgments

The results show no difference in the level of satisfaction with group judgments between the novice and expert subjects (p=.52). The results also indicate that the level of satisfaction with group differs among groups.

7.2.4 Perception Measure – Perceived Usefulness of ESS

The expert and novice groups that were provided with the complete ESS support (i.e., with explanations) was compared in terms of their perceived usefulness of ESS. The hypothesis to be tested is:
Hₚ7: The experts' perceived usefulness of ESS is lower than that of the novices.

Evaluation of Assumptions of ANOVA Model

Testing Assumptions of Nested ANOVA Design with Individual as Level of Analysis. The distributions of the residuals satisfy the normality, homogeneity of variances, and independence assumptions\(^{16}\). As such, the nested design was used for the analysis.

Comparison of Novice versus Expert Subjects' Perceived Usefulness of ESS

The descriptive statistics of the novice and expert subjects' perceived usefulness of ESS are presented in Table 7-35 (with individual as the unit of analysis), and the results of the analysis are presented in Table 7-36.

<table>
<thead>
<tr>
<th>Source</th>
<th>N</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Std. Error</th>
<th>Min</th>
<th>Max</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experts/Professionals</td>
<td>18</td>
<td>42.50</td>
<td>10.19</td>
<td>2.40</td>
<td>18</td>
<td>55</td>
<td>37</td>
</tr>
<tr>
<td>Novices/Students</td>
<td>27</td>
<td>49.30</td>
<td>6.52</td>
<td>1.25</td>
<td>39</td>
<td>62</td>
<td>23</td>
</tr>
<tr>
<td>Total</td>
<td>45</td>
<td>46.58</td>
<td>8.75</td>
<td>1.30</td>
<td>18</td>
<td>62</td>
<td>44</td>
</tr>
</tbody>
</table>

Table 7-35: Descriptive Statistics of Perceived Usefulness of ESS (Individual Level)

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>DF</th>
<th>MS</th>
<th>F</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Expertise</td>
<td>498.85</td>
<td>1</td>
<td>498.85</td>
<td>5.24</td>
<td>.04</td>
</tr>
<tr>
<td>Group within Treatment</td>
<td>1236.80</td>
<td>13</td>
<td>95.14</td>
<td>1.75</td>
<td>.10</td>
</tr>
<tr>
<td>Error</td>
<td>1633.33</td>
<td>30</td>
<td>54.44</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 7-36: Results of Nested ANOVA – Perceived Usefulness of ESS

\(^{16}\) The Kolmogorov-Smirnov (Lillefors) and the Shapiro-Wilk statistics for the expert group are .20 (\(p=.07\)) and .89 (\(p=.05\)) respectively, indicating a minor violation of the normality assumption. In fact, these statistics are not considered violations of the normality assumption if Stevens' (1996) recommendation to adopt an alpha level of .01 is accepted. The Levene statistic of 1.71 (\(p=.14\)) indicates that the homogeneity of variances assumption is satisfied.
The results indicate that the novices found the complete ESS support to be more useful than the experts did (p=.04). This finding is consistent with what we expected as novices lack experience in carrying out the task and therefore would find the ESS support to be more helpful.

7.2.5 Perception Measure — Trust in ESS

The expert and novice groups that were provided with the complete ESS support (i.e., with explanations) was compared in terms of their trust in the ESS. The hypothesis to be tested is:

\[ H_{p8} : \text{The experts'} trust in ESS is lower than that of the novices. \]

Evaluation of Assumptions of ANOVA Model

Testing Assumptions of Nested ANOVA Design with Individual as Level of Analysis. The distributions of the residuals satisfy the normality, homogeneity of variances, and independence assumptions. The Kolmogorov-Smirnov (Lillefors) and the Shapiro-Wilk statistics support the normality assumption, and the Levene statistic (=.12; p=.28) supports the homogeneity of variances assumption. As such, the nested design was used for the analysis.

Comparison of Novice versus Expert Subjects’ Trust in ESS

The descriptive statistics of the novice and expert subjects’ trust in ESS are presented in Table 7-37 (with individual as the unit of analysis), and the results of the analysis are presented in Table 7-38.
Table 7-37: Descriptive Statistics of Trust in ESS (Individual Level)

<table>
<thead>
<tr>
<th>Source</th>
<th>N</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Std. Error</th>
<th>Min</th>
<th>Max</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experts/Professionals</td>
<td>15</td>
<td>28.87</td>
<td>3.25</td>
<td>.84</td>
<td>22</td>
<td>34</td>
<td>12</td>
</tr>
<tr>
<td>Novices/Students</td>
<td>27</td>
<td>31.00</td>
<td>4.11</td>
<td>.79</td>
<td>24</td>
<td>39</td>
<td>15</td>
</tr>
<tr>
<td>Total</td>
<td>45</td>
<td>30.24</td>
<td>3.92</td>
<td>.61</td>
<td>22</td>
<td>39</td>
<td>17</td>
</tr>
</tbody>
</table>

Table 7-38: Results of Nested ANOVA – Trust in ESS

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>DF</th>
<th>MS</th>
<th>F</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Expertise</td>
<td>43.89</td>
<td>1</td>
<td>43.89</td>
<td>2.00</td>
<td>.18</td>
</tr>
<tr>
<td>Group within Treatment</td>
<td>285.73</td>
<td>13</td>
<td>21.98</td>
<td>1.97</td>
<td>.07</td>
</tr>
<tr>
<td>Error</td>
<td>302.00</td>
<td>27</td>
<td>11.19</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The results show no difference in the level of trust in ESS between the novice and expert subjects (p=.18). This indicates that novices and experts have similar level of trust in the system. However, as discussed earlier (see Section 7.2.4), novices found the ESS support to be more helpful than the experts did. The results also indicate that the level of trust in the ESS differs among groups (p=.07).

7.3 Summary of Chapter 7

Novices are influenced by the ESS analyses and explanations support more than experts. This indicates that experts do not accept the advice and explanations given by the ESS as much as novices. In addition, with ESS analyses and explanations support, the novices achieve a higher level of consensus in their group judgments than the experts. Thus, experts tend to disagree, with both the ESS and among themselves, more than novices do.

17 Three expert subjects did not provide complete ratings of trust.
There is no difference in the level of satisfaction with group process and group judgments between the experts and novices when ESS analyses and explanations support are provided. The experts and novices also do not differ in their level of trust in the ESS. As expected, the novices find the ESS to be more useful than the experts find it.
CHAPTER 8: RESULTS OF QUALITATIVE ANALYSIS

Chapters 6 and 7 present the results of statistical (or quantitative) analysis while Chapter 8 presents the qualitative results to support the quantitative analysis. Using a combination of qualitative and quantitative methodologies, we are able to draw upon the strengths of both. The quantitative approach facilitates comparison and statistical aggregation of the data, thus allowing the findings to be presented succinctly and parsimoniously. On the other hand, the qualitative approach produces a wealth of detailed information and depth on, and individual meaning to, the phenomena studied, thus increasing our understanding of the group processes.

In this research, a qualitative analysis of the group decision making processes and interactions was carried out to identify differences across treatments, groups, and user characteristics (i.e. domain expertise). Video recordings and transcripts of the group decision making processes as well as computer logs captured during the experiment aid the process analysis. The voluminous qualitative data collected was then organized into readable narrative description with major themes, categories, and illustrative case examples extracted through content analysis. These qualitative findings are used to support, complement, and explain the quantitative findings described in Chapters 6 and 7.

Patton (1990) and Becker and Geer (1970) argue that participant observation is the most comprehensive and complete of all types of research strategies. It allows the researcher to understand fully the complexities of the situations. In this research, direct quotations from subjects were presented whenever appropriate to support the analysis. Section 8.1 identifies anecdotal evidence and supporting observations from the protocol to supplement the results of quantitative analysis.
8.1 Qualitative Analysis to Support Quantitative Analysis

There are three main findings in the quantitative analysis:

1. Knowledge transfer to novice users increases with increasing level of ESS support;
2. Novices are more influenced by ESS analyses and explanations support than experts;
3. With ESS analyses and explanations support, novices achieve a higher level of consensus in their judgments than experts.

8.1.1 Knowledge Transfer to Users Increases with Level of ESS Support

The quantitative results show that both the ESS analyses and explanations support are responsible for knowledge transfer from the ESS to the users. The following three subsections explain how the ESS conclusions, feedforward explanations, and feedback explanations lead to knowledge transfer.

8.1.1.1 How Do Conclusions Given by the ESS Increase Knowledge Transfer?

From the analysis of conversations and interactions, three factors were identified as contributing to knowledge transfer:

1. ESS conclusions highlight important points and issues,
2. ESS conclusions help users identify and correct errors in their inferences,
3. ESS conclusions are used by members of the group to help them reinforce, support, and convince others of their points.

Highlight Important Points and Issues

The ESS conclusions brought up important points that were taken into account by the groups, providing direct evidence of knowledge transfer from the ESS to the groups. This contributed to the process gains of the groups. Some examples are given below.
Example 1:
S3: "(reading from conclusion) Company is earning sufficient funds to cover all fixed charges such as interest, lease, and rent payments. This position has been improving steadily suggesting little danger of the company defaulting on its fixed obligations. However, there is a concern as regards to having much too large a safety cushion for these charges. (end of conclusion) That's right."
S1: "So that's not bad, I mean, for long term solvency, it's quite good."
S3: "Oh yeah, it's good for us (commercial lenders)."
(and the discussion continued...)

Example 2:
"But how about the funds flow adequacy. From the computer, it's not very optimistic right? Computer indicates that."

Example 3:
"... the last (ESS) recommendation was saying that they might not be reinvesting enough, right?"

Example 4:
"Hasn't the equity been increasing though? Didn't it (ESS) say before that..."

Identify and Correct Errors in Inferences

When the ESS conclusions were inconsistent with the inferences made by a group, they prompted the group to reason whether the ESS or the group was right. In most cases, novices accepted the conclusions based on the credibility of the source. This indicates the expert power of ESS. For instance, one of the members in a novice group made the following remark after looking at a conclusion provided by the ESS,

"Computers can't be wrong."

Another example illustrates the case where the ESS conclusion prompted a member of a novice group to his error, which was then corrected. The member mistakenly thought the
proportion of accounts receivable of the company had decreased over the years while the conclusion indicated the reverse. The conversation of the group was as follows:

SI: 
"(reading conclusion) There is a trend toward having a high proportion of sales on credit. While the proportion has tripled in the last 5 years, it still remains within a range well below industry levels. (end of conclusion) Let's see the balance sheet."

S3: “So ... this (accounts receivable) has increased.”
S2: “Decreased. That one decreased from last year. Right?”
S3: “What?”
S2: “It’s lower than last year.”
S3: “So then it’s the overall trend...”
S2: “Right. Increasing.” (mistake noted)

Reinforce One’s Points that are Consistent with ESS

The ESS conclusions were used by some members to support, emphasize, and help them bring the attention of other group members to their own points. This indicates the use of the expert power of ESS by individuals in the group. Two examples follow:

Example 1:
(After reading conclusion) "That’s what I said. They use equity to finance their improvements rather than debt."

Example 2:
“Point number (conclusion) 2. I used that for financial management, that’s why I gave such a low mark.”

Discussion

The ESS was being viewed as an additional member in the group. With ESS analyses support, added knowledge and expertise become available for decision making, more
perspectives and issues are taken into account, and errors and inconsistencies are sometimes identified by the decision makers when they compare their inferences against those provided by the ESS. These process gains are likely to result in a more thorough and complete analysis being carried out.

8.1.1.2 How Do Feedforward Explanations Increase Knowledge Transfer?

Feedforward explanations constitute non case-specific, generalized information pertaining to the input cues of an analysis provided to users prior to an analysis. The feedforward explanations provided by the ESS are also responsible for knowledge transfer from the ESS to the users. The following four factors explain how feedforward explanations increase knowledge transfer from the ESS to the users:

1. learn new concepts and their implications,
2. confirm or compare with one’s own knowledge,
3. understand or resolve disagreement with ESS conclusions,
4. seek or browse for additional or missing information.

Learn New Concepts and their Implications

Majority of the times, the feedforward explanations were accessed to learn new or unfamiliar domain concepts. Examples of such cases are as follows:

Example 1:

S2: “I’d like to have a look at internal growth rate. Do you know what it is?”
S1: “No.”
S3: “No, let’s take a look.”

Example 2:

“Can I see financial leverage? I don’t really understand what it means.”
Example 3:
“What is asset turnover... how do they compute that?”

Example 4:
“Could you go to the liquidity index and see how it is calculated?”

The feedforward explanations were useful in helping users learn new domain concepts. By providing feedforward explanations, users were given the opportunity to find out what these concepts were whenever the need arose.

There were indications in the protocol that call for the need for feedforward explanations. When feedforward explanations were not available, the group members ignored the unfamiliar ratios most of the time. This put a limit on process gains. The following illustrates some examples.

Example 1:
“I can’t remember acid test. Do you know that?” (group ignored the ratio)

Example 2:
“I didn’t know how to read working capital.” (no discussion on working capital follows)

Example 3:
S1: “What is liquidity index? Who knows what that is?”
S3: “I have no idea what liquidity index is.”

Example 4:
S1: “What is this, liquidity index? I don’t know.”
S2: “I don’t know that either. But it’s pretty high.”
S1: “What does that mean?”
S3: “I don’t know.”
Confirm or Compare One’s Own Knowledge

Feedforward explanations may also be accessed when users already have some idea about a domain concept. The use of feedforward explanations benefit the groups in one of two ways: 1) confirmed that their knowledge of the domain concept is correct, or 2) corrected users’ misconceptions of the domain concept if the knowledge presented by the ESS was inconsistent with that of the users. The following examples illustrate the cases:

Example 1 (confirmation):
One group made use of the feedforward explanation to confirm that their definition of fixed asset turnover was consistent with that of the ESS.

S1: “What was that (fixed asset turnover)?”
S2: “Let’s look at it.”
S3: “Just click, can you click on fixed asset turnover and go to how?”
(View explanation) ...
S2: “That’s what we were all thinking of.”
S3: “Yeah.”

Example 2 (mistake identified and corrected):
The feedforward explanation was used by one group to verify their definition of “conversion period”. The explanation helped them to identify and correct their misconceptions.

S2: “... The conversion period... that’s a term I’ve not seen.”
S3: “I think it just means converting the current asset to cash.. how many days does it take to turn.” (error in subject’s definition)
S2: “How is that different from days in inventory?”
S3: “Well, it’s essentially the same test, right?”
S2: “Is it?”
Example 3 (misconception identified and corrected):
In the following example, the feedforward explanation corrected one member's misconception and allowed other members to learn a new concept.

S2: “Liquidity index. I don’t know what...”
S3: “What is that? Yeah, lets take a look.”
(looked at explanations)
...
S1: “So the lower the better?”
S2: “Yeah.”
S1: “I thought the other way around.” (misconception corrected)

The feedforward explanations provided users the opportunity to compare and confirm their definition and idea of domain concepts. The ESS helped them to either confirm their knowledge or to identify and correct their misconceptions of the domain concepts.

Understand or Resolve Disagreement with ESS Conclusions

One interesting feature of FINALYZER, which was discussed in Chapter 5, is the provision of hypertext linkages to feedforward explanations from the ESS conclusion screens. In other words, if an ESS conclusion includes a domain concept, deep knowledge on that domain concept can be directly accessed from the conclusion screen instead of having to traverse back several screens to access them. (Refer to the conclusion screen provided in Appendix B to see an example of hypertext linkages to deep explanations.) This feature provided convenient access to feedforward explanations which further improved understanding of the ESS conclusions. The following are examples of accesses to feedforward explanations which
were carried out by the users to increase their understanding of the ESS conclusions or to resolve disagreement with the conclusions.

Example 1 (increase understanding):
One subject remarked that she did not understand the conclusion on funds flow adequacy. The group then requested the contextualized (i.e., hypertext-linked) feedforward explanation on funds flow adequacy provided within the conclusion to help them understand the conclusion.

SI: "I don't get the second one (conclusion)."
S2: "Can we have an explanation on the funds flow adequacy please."

Example 2 (increase understanding):
One group, in trying to understand an ESS conclusion, went back several screens in the system to access two feedforward explanations related to the analysis. The feedforward explanation on profit margin was accessed in an attempt to understand its relationship with asset turnover.

"The net profit is high, profit margin's high, but our asset turnover is low, so it's offsetting. Can I see the net profit margin? Click on that little circle and see why."

---

1 The conclusion reads: "The funds flow adequacy of the company is low. It is not generating sufficient cash from operations to cover capital expenditures and net investments in inventories, etc. There is a need to secure additional financing for operations."

2 The conclusion reads: "Canacom's management is following a policy of accepting a lower asset turnover for higher profit margins. This is paying off currently in the form of better-than-industry return on assets as the increase in the return on sales more than offsets the expense of the lower asset utilization rates. However, there is a concern about Canacom's ability to continue this policy in the future, especially in the face of the increased competition of the electrical and electronic products marketplace."
Example 3 (resolve disagreement):

After looking at the ESS conclusion which indicates that the funds flow adequacy of the company is low, and therefore it is not generating sufficient cash from operations to cover capital expenditures and net investments in inventories, the group disagreed with the conclusion and sought the feedforward Why and How explanations on funds flow adequacy in an attempt to resolve the disagreement.

S1: “Let’s go back to funds flow adequacy and see what it is.”
S2: “We want to look at everything on it because we’re in disagreement.”

Example 4 (clarify surprise):

One expert group sought the feedforward explanation to help them understand an ESS conclusion.

S2: “I’m not sure how they (ESS) come to that conclusion.”
S1: “Do number two, the funds flow adequacy on how please.”

Example 5 (resolve disagreement):

The ESS came to a different conclusion from the expert group, which then requested the feedforward How explanation on funds flow adequacy to resolve the disagreement.

---

3 The conclusion reads: “The funds flow adequacy of the company is low. It is not generating sufficient cash from operations to cover capital expenditures and net investments in inventories, etc. There is a need to secure additional financing for operations.”
4 The conclusion reads: “The funds flow adequacy of the company is low. It is not generating sufficient cash from operations to cover capital expenditures and net investments in inventories, etc. There is a need to secure additional financing for operations.”
5 The conclusion reads: “The funds flow adequacy of the company is low. It is not generating sufficient cash from operations to cover capital expenditures and net investments in inventories, etc. There is a need to secure additional financing for operations.”
"It came to a different conclusion entirely from myself. I was looking at this thing as being generating cash. And this is saying that the cash flow is insufficient to cover the total of capital investments, inventory and cash equivalent paid out. Or it didn’t pay any dividends." (requested explanation)

Thus, feedforward explanations were also used to increase the users’ understanding of the ESS conclusions and to resolve disagreements with the conclusions. Such requests were facilitated by the hypertext linkages to feedforward explanations from the ESS conclusions in FINALYZER.

Seek Additional Information

This is a general category where feedforward explanations were accessed to find additional information about domain concepts that users had some knowledge about. Examples of such uses include:

Example 1:
One group browsed the feedforward explanation on equity to total debt just to see if they could gather additional information.

"Let’s go to equity to total debt. We’ll see what they talk about. Why (explanation). Is this firm highly leveraged?"

Example 2:
One group browsed the feedforward explanation on days to sell inventory just to see what it says.

S2: “Well, what does it say under days to sell inventory? Is it the same type of information?”
S3: “Let’s try it. Days to sell inventory.”
In short, feedforward explanations were frequently used to browse for additional information on domain concepts in the attempt to improve users' understanding of the concepts.

Discussion

The feedforward explanations provide users the opportunity to learn new concepts and their implications. They allow users to verify their knowledge of domain concepts as well as to identify and correct their misconceptions of domain concepts. The feedforward explanations are also used to increase users' understanding of the domain and ESS conclusions, and to resolve disagreement with the conclusions.

Occasionally, feedforward explanations were not used when they might have been useful. Such situations occurred when one or more group members offered an explanation or thought there was no need for explanations. Sometimes, the right explanation was given by the group member(s) who volunteered the explanation, while at other times, the reasoning or explanation given were wrong. The following are some examples from the transcripts:

Correct or acceptable explanations given by group member(s):

Example 1:
S3: "What is acid test?"
S1: "I think acid test is without the inventory basically (as compared to current ratio)"
S3: "Wasn't that quick ratio or is acid test the same as quick?"
S1: "Yup."
S3: "Okay."

(Subject 1 answered Subject 3's query before any ESS explanation was accessed.)
Example 2:
S2: "Earning price ratio. What is that?"
S3: "Earning price ratio is just the opposite, the inverse (of price earning ratio)."

(Subject 3 gave an acceptable answer to earning price ratio, as price earning ratio is more common and better known.)

Wrong explanations given by group member(s):

Example 1:
S2: "Can we see internal growth rate? ... Can I get into how? No, no, why, why, why."
S1: "Increase in total equity." (volunteered an explanation before ESS explanation was accessed)
S2: "Oh, okay."

(Subject 1 gave the wrong definition for internal growth rate. ESS explanation was not accessed after one member volunteered an explanation.)

Example 2:
S2: "Conversion period. What's that mean? ..."
S1: "Conversion period is the number of days to convert raw materials into finished goods."

(Subject 1 gave the wrong definition for conversion period.)

The above examples illustrate cases where explanations were not used when the situation called for their use. Unlike the case in individual decision making, other members in the group may provide the explanations. If the explanation provided by the group member is correct, then time is saved in not accessing the system (e.g., time taken to carry out system request is eliminated, listening is quickly than reading), resulting in a process gain. On the
other hand, if the group member gives the wrong explanation, then the use of the ESS explanation facilities would have benefited the group.

8.1.1.3 How Do Feedback Explanations Increase Knowledge Transfer?

Feedback explanations constitute case-specific information provided to users at the end of an analysis. The feedback explanations provided by the ESS are also responsible for the knowledge transfer from the ESS to the users. From the analysis of conversations and group interactions, five factors were identified as contributing to knowledge transfer:

(1) understand the significance or process of ESS analyses,
(2) confirm or compare with one's own reasoning process,
(3) clarify surprises or resolve disagreement with ESS conclusions,
(4) provide overall assessments of analyses,
(5) seek or browse for additional or missing information.

Understand the Significance or Process of ESS Analyses

Some of the ESS conclusions were not understood by the users or their significance in relation to the task was not obvious to the users. The feedback explanations provided users the opportunity to understand the significance and process of the ESS analyses. The following illustrates examples of such uses:

Example 1:

One subject didn’t understand how an ESS conclusion was derived and requested a feedback How explanation of the conclusion.

"I don't really understand how they get that. ... How (requesting explanation)."
The feedback explanation not only increased the users’ understanding of the conclusion but also convinced them to accept the conclusion. From their understanding of the conclusion, the group proceeded to make further inferences about the task.

Example 2:
After reading an ESS conclusion\(^6\) that the growth rate of the company has peaked, one subject requested the feedback Why explanation of the conclusion while making the following remark,

"Why does it make a conclusion that its growth rate is peaking? Where did you see that?"

In short, the feedback explanations were used to understand the reasoning and significance of the ESS conclusions, thus they enhanced users’ understanding of the ESS analyses.

Confirm or Compare with One’s Own Reasoning Process
Feedback explanations were sought by the subjects to compare the reasoning process of the ESS with their own. This occurred when the subjects had idea more or less consistent with the conclusions given by the ESS and, hence, wanted to compare and make sure that their reasoning process was consistent with that of the ESS. The following shows examples of such uses of feedback explanations:

\(^6\) The conclusion reads: “The company’s growth rate has peaked and the over-optimistic stock price can be expected to fall more in line with the market in the future. It suggests management may have to diversify into newer markets and/or products to sustain the rapid growth of the last 5 years.”
Example 1:
Having identified the inventory problem in an earlier discussion, one subject, after looking at an ESS conclusion* that indicates the unfavorable inventory situation, said,

"Looks like this inventory problem is pretty big. Can we see a why on the (conclusion on) inventory?"

The group proceeded to compare their reasoning on inventory with that of the ESS.

Example 2:
The following conversation took place after a novice group read an ESS conclusion* which raised the concern that the firm may not be investing sufficiently to take advantage of its growth opportunities.

S3: "Is that because they have a lot of cash somehow?"
S2: "Maybe."

The group then checked both the feedback Why and the How explanations of that conclusion and accepted the conclusion after that:

S3: "Oh, so it's saying it needs to reinvest quite a bit of money to be able to stay competitive. And it's falling slightly short of that because it's making an aggressive move."

---

* The conclusion reads: "Inventory is very unfavorable in comparison with major competitors and industry composites. The decreasing trend also suggests the need to extend tighter control over inventory management."

* The conclusion reads: "The firm is reinvesting adequately to maintain its operating capacity. However, there is a concern if it is investing sufficiently to take advantage of its growth opportunities."
Example 3:

After reading the ESS conclusion\(^9\) that the company's growth rate has peaked and checking its feedback Why explanation, one member of the group was still not convinced that the growth rate had peaked.

*S3:* "Why does it make a conclusion that it's growth rate is peaking? Where did you see that?"

*S2:* "Well, if you look at how much it was growing per annum, the growth rate has slowed down over the last year."

*S3:* "Why is it peaking, when it still has the capacity to grow assuming there's a continuing market for its products. In its limited form, it still has the capacity to increase its sales, I mean, it even said that it wasn't using its receivable level that optimally and I just don't see that immediately."

*S2:* "Well, I think maybe what it's doing is gathering a bunch of information, if sales, growth rate has slowed down, if its asset utilization's down, its inventories are not turning, it can't split its market mix, that's telling you that maybe there's a lot of competition out there. ..."

*S1:* "Yes, I agree with you. It's not positioned itself well to keep the growth growing."

*S2:* "Yeah."

*S3:* "And it's in a very limited market too, it's only in one city essentially."

*S1:* "Yeah, it was 75% or 90% in Vancouver, Lower Mainland."

*S2:* "That's right. So how much more can it grow here, I guess."

The feedback How explanation was then requested to confirm their reasoning process.

---

\(^9\) The conclusion reads: "The company's growth rate has peaked and the over-optimistic stock price can be expected to fall more in line with the market in the future. It suggests management may have to diversify into newer markets and/or products to sustain the rapid growth of the last 5 years."
In short, the feedback explanations were used by the subjects to confirm their reasoning process with that of the ESS, which helped to correct any errors in their reasoning process and further increased the group members’ understanding of the ESS analyses.

**Clarify Surprises or Resolve Disagreement with ESS Conclusions**

A substantial portion of the uses of feedback explanations was for clarifying surprises or resolving disagreement with the ESS conclusions. Examples of such cases are described next.

**Example 1:**

Having decided earlier that the *stock price of the company was undervalued*, the group could not agree with the ESS conclusion\(^6\) that the stock price of the company is slightly “overvalued”.

*S3: “Why is it overvalued? That’s where I don’t agree.”*
*S1: “Let’s go to why. No, let’s go to how.”*

They requested the feedback How explanation of the conclusion to help them resolve the disagreement and were later convinced by the system that the stock price was overvalued.

*S3: “P/E ratio is increasing while the growth rate is declining. Oh, that’s true, that’s true.” (convinced by the explanation)*

---

\(^6\) The conclusion reads: “Considering its fundamentals, the stock price of Canacom is high and slightly “overvalued”. This suggests that it could be a good time for management to raise equity capital. As stocks are expected to level off in the medium-term, it is not recommended that a convertible stock be accepted for the purposes of securing a loan.”
Example 2:

Upon seeing the ESS conclusion on receivable management and the strategic explanation indicating that the assessment of receivable management of the company is negative, one subject remarked,

"It (the ESS conclusion) says it's still well below the industry standards. ... Well, how can you get a negative, if it's well below (the industry standard)?"

The group proceeded to view the feedback Why and How explanations of the conclusion to resolve the disagreement. The group accepted the conclusion after looking at the explanations.

Example 3:

After reading the ESS conclusion that the stock price of the company is slightly "overvalued", one subject disagreed,

"How could it be slightly overvalued?"

The group requested a justification using the feedback Why explanation and accepted it without any further question.

Example 4:

One subject, after reading the ESS conclusion that there was an increasing trend towards having a high proportion of sales on credit, curiously asked,

---

11 The conclusion reads: "There is an increasing trend toward having a high proportion of sales on credit. While the proportion has tripled in the last 5 years, it still remains within a range well below industry levels."
12 The conclusion reads: "Considering its fundamentals, the stock price of Canacom is high and slightly "overvalued". This suggests that it could be a good time for management to raise equity capital. As stocks are expected to level off in the medium-term, it is not recommended that a convertible stock be accepted for the purposes of securing a loan."
"Where's that trend?"

The group requested the feedback Why and How explanations and still disagreed with the conclusion after viewing the explanations.

"Yeah, right. There's an error there. At 3 or 7% of total sales, there's no trend to finance sales. ..."

One subject insisted there must be an error in the system because the amount of sales on credit was too small to justify the trend. (This example is also used to demonstrate the heightened criticality hypothesis of experts in Section 8.1.2.)

Example 5:
One group could not agree with the ESS conclusion that the company was not generating sufficient cash from operations to cover capital expenditures and net investment in inventories, etc. One of the group members said,

"I thought they had a surplus cash flow."

The member thought the company had a surplus cash flow, indicating that he was surprised with the ESS conclusion. The group proceeded to look at the feedback Why and How explanations for the conclusion to try to figure out where exactly the disagreement lied.

---

13 The conclusion reads: "There is an increasing trend toward having a high proportion of sales on credit. While the proportion has tripled in the last 5 years, it still remains within a range well below industry levels."

14 The conclusion reads: "The funds flow adequacy of the company is low. It is not generating sufficient cash from operations to cover capital expenditures and net investments in inventories, etc. There is a need to secure additional financing for operations."
One of the major roles of feedback explanations in this study was to resolve disagreement or surprises that the users had with the ESS conclusions. The feedback explanations provided users the opportunity to understand the significance and reasoning of the ESS conclusions, thus increasing the chances for the ESS conclusions to be accepted by the users.

View Overall Assessments

The feedback Strategic explanations provide an overall assessment for each analysis. The subjects seemed to find the overall assessments useful as the groups that started using it tended to continue to use it in the other analyses that followed. For example, novice groups 2 and 5 each requested 4 out of 5 feedback Strategic explanations, while novice group 7 and expert group 3 requested all 5 feedback Strategic explanations. Most subjects who requested this explanation indicated they need some kind of a summary for the analysis. Since the assessments were presented in "pluses and minuses" form which correspond to positive and negative assessments, most subjects verbalized their intention to view the assessments in ways similar to the following:

"If we go to strategic, would it tell us pluses and minuses?"

"... the strategic shows the pluses and minuses"

Seek Additional Information

This is a general category in which feedback explanations were used to find details that might be useful to back up or justify an ESS conclusion. The following are examples of such cases.
Example 1:
One ESS conclusion highlighted the inventory problem which was a concern in the group.

S2: "That's what you're concerned."
S1: "So they (ESS) don't like inventory either. Maybe we should look into (the conclusion on) inventory. See what they talk about it."

The feedback Why explanation was requested to look for additional information on the significance of the problem.

Example 2:
After reading an ESS conclusion on the declining trend of asset utilization ratios, the group requested the feedback Why explanation to gather additional information.

"I wouldn't mind seeing an explanation of this one, the second one."

In short, feedback explanations were frequently used by the subjects to gather additional or missing information on the ESS conclusions. Sometimes, the explanations provided the answers to their questions, but sometimes they did not.

Discussion
The feedback explanations provided users the opportunity to understand the significance or process of ESS analyses, compare their reasoning with that of the ESS to identify any

---

15 The conclusion reads: "Inventory is very unfavorable in comparison with major competitors and industry composites. The decreasing trend also suggests the need to extend tighter control over inventory management."

16 The conclusion reads: "Asset utilization ratios have been on a declining trend, with the exception of net property, plant, and equipment. Management should more seriously consider the trade-offs between the benefit of holding back each particular class of assets and the cost of tying up funds in that class of assets. The declining trend suggests that management is having a low level of success in coping with the changing business environment."
discrepancies and correct any errors, clarify surprises and resolve disagreement with the ESS conclusions, compare their overall assessments with those of the ESS, and increase users’ understanding of the ESS conclusions.

There were circumstances where one or more members interrupted the explanation request of another member to provide his/her explanation. At times, the explanation given could be erroneous, as illustrated by the following example:

**Example 3:**

S3: “Why if you have a lower asset turnover, you have higher profit margin. I don’t understand.”...

S2: “So maybe we should go to why instead of...” (request interrupted)

S3: “Oh, okay, I get it. I get it. Okay when you use less assets, your depreciation expense is much less, right?”

S1: “Yeah. So you have higher income?”

S3: “So you have higher net income. And you have higher profit margin.”

(Subject 3 interrupted request for explanation by giving the wrong explanation.)

On the other hand, for ESS conclusions that were not evident to the novices, the novices tended to make up some justifications for these conclusions when feedback explanations were not available. However, such justifications were not always correct.

“Do you guys understand number (conclusion) 2 because I don’t... I don’t understand it at all.” (The group proceeded to figure out explanation by themselves)

“What is this trying to say, number (conclusion) 3?” (group generated explanation for the conclusion)
8.1.1.4 Summary
ESS analyses and explanations support contribute a number of process gains to group
decision making. First, it provides conclusions which 1) highlight important points and
issues for consideration, 2) help users identify and correct errors in their inferences, and 3)
are used by members of the group to reinforce, support, and convince others of their points.
Second, the feedforward and feedback explanations also contribute to the process gains. The
feedforward explanations allow users to 1) learn new concepts and their implications, 2)
check and confirm that their knowledge is consistent with that of the ESS, 3) correct any
misconceptions or errors in their knowledge, 4) increase their understanding of ESS
conclusions, 5) resolve disagreement with ESS conclusions, and 6) seek additional
information about domain concepts. The feedback explanations 1) increase users’
understanding of the significance of ESS conclusions and the process of ESS analyses, 2)
allow users to check if their reasoning process is consistent with that of the ESS, and if not,
correct any errors identified, 3) clarify surprises and resolve disagreement with ESS
conclusions, 4) provide overall assessments of analyses, and 5) allow users to seek additional
information about the conclusions. These factors contribute as process gains to the group
decision making process.

8.1.2 ESS Analyses and Explanations Support Influence Novices More than Experts
Two main differences between the experts and novices were apparent from the analysis of
the protocols: criticality with ESS analyses, and level of reliance on the ESS. The experts
tended to be very critical toward the analyses given by the ESS, while the novices tended to
rely more heavily on the ESS conclusions than the experts. No significant difference existed
in the total number of explanations accessed between the experts and novices, indicating that
the greater influence of the ESS on the novices was not a result of seeking more explanations.
8.1.2.1 Differences between Experts and Novices

The initial judgments (i.e., before ESS was introduced) of the experts and novices did not differ in their amount of deviation from the judgments of the original experts. However, after using the ESS analyses and explanations support in the group setting, the judgments of the novices became closer to those of the original experts than the judgments of the experts were. The quantitative results indicate that the experts were less influenced by the ESS than the novices, but the quantitative approach is not able to explain the reason for such an occurrence. As such, we drew upon the qualitative approach, which involved the analysis of conversations and group interactions, to explain the reasons why and how it occurred. The two differences identified — criticality characteristic of experts and greater reliance of novices on the ESS — were used to explain for the difference in judgments between the experts and novices.

Criticality Characteristic of Experts

Some of the experts were very critical of the conclusions and explanations given by FINALYZER, and they challenged these conclusions and explanations. On the other hand, the novices were less critical with the ESS conclusions and explanations and were less likely to challenge them. This is probably because the experts have a higher level of experience, understanding, and knowledge to critic the conclusions of FINALYZER and hence, are less likely to accept them. This finding is consistent with the heightened criticality hypothesis found in experts (Biek and Wood, 1996), which explains why the experts were less consistent with the conclusions given by the ESS.
Example 1 (Expert Group 2):

One expert subject, after looking at the ESS conclusion which stated that the company was following a policy of accepting a lower asset turnover for higher profit margin, said it was too “dangerous” for the ESS to make that conclusion. His response to the conclusion was as follows:

“It’s just that I think that’s a bit of a dangerous conclusion to say that one (higher profit margin) follows from the other (lower asset turnover). They’re accepting a lower asset turnover, there’s no question. But that doesn’t mean that’s why they’re generating a higher profit margin. They could be generating a higher profit margin because they have a better computer. And the fact that they have low turnover is just bad management. So, you know, I appreciate that may be the case, but it’s certainly not an easy conclusion to draw. It’s like saying your friend at Volkswagen has got profitable business this year because he’s a good manager.”

Example 2 (Expert Group 2):

One expert subject had some problem accepting one ESS conclusion, even after viewing the feedforward and feedback explanations associated with it.

“Okay, well I just think that it’s worth noting that very few companies finance their growth in inventory from cash flow. I think if you’re constantly depleting your working capital because you’re not generating enough cash flow to maintain your working capital position that could be a big problem. But generally speaking, inventory can be financed through credit and other current liabilities. So I don’t think if you’ve got a high growth company you can expect to finance inventory growth from cash flow. But I’m not sure that I agree with that conclusion although I take what they’re saying. So I’m fine with that now.”

17 The conclusion reads: “Canacom’s management is following a policy of accepting a lower asset turnover for higher profit margins. This is paying off currently in the form of better-than-industry return on assets as the increase in the return on sales more than offsets the expense of the lower asset utilization rates. However, there is a concern about Canacom’s ability to continue this policy in the future, especially in the face of the increased competition of the electrical and electronic products marketplace.”

18 The conclusion reads: “The funds flow adequacy of the company is low. It is not generating sufficient cash from operations to cover capital expenditures and net investments in inventories, etc. There is a need to secure additional financing for operations.”
Example 3 (Expert Group 4):

After reading the conclusion that there was an increasing trend towards having a high proportion of sales on credit, and viewing the feedback Why and How explanations to find out the trend identified by the system, the group disagreed critically with the conclusion. The following are the conversations that took place after the explanations were accessed:

S2: “How come the proportion of receivables are so small? If you look at them, in comparison to everything else, its really not too big.”

S1: “Yeah, right. There’s an error there. At 3 or 7% of total sales, there’s no trend to finance sales. Total sales of how many millions of dollars? Two and a half million bucks to have $80,000 outstanding in receivables is peanuts. Are we missing something there?”

S2: “I don’t know that’s relevant.”

The experts could not accept the explanations given by the system as they felt that an increase from 3 to 7% of accounts receivable was too small to justify “a trend”.

None of the novice groups disagreed with the ESS conclusions as strongly as the experts. In fact, the novices tended to be satisfied with the explanations given and went along with the ESS conclusions after reading the explanations. In cases where the subjects did not seem to be completely convinced by the explanations, the novices gave the benefit of doubt to the ESS by accepting or not disagreeing with the conclusions whereas the experts continued to exert and voice their disagreement with the conclusions.

19 The conclusion reads: “There is an increasing trend toward having a high proportion of sales on credit. While the proportion has tripled in the last 5 years, it still remains within a range well below industry levels.”
Reliance on ESS

It was evident from the group interactions that the novices tended to trust and rely more heavily on the ESS in deriving their group judgments. This point supplements the previous point on the criticality characteristic of experts. While rehearsing and elaborating the ESS conclusions and their supporting arguments, the novices were more likely than the experts to respond in a positive manner to the message. For instance, there is the tendency for novices to respond positively to, and be persuaded by, the ESS explanations. As an example, one novice group initially disagreed with the ESS that the stock was overvalued but later agreed with it after viewing the explanations. On the other hand, the experts were more critical with the ESS conclusions and generated more negative responses to the conclusions (as discussed earlier), leading them to be less persuaded by the conclusions.

From the conversations and interactions, we observed that the majority of the novices treated the ESS as a knowledgeable and respected "member" due to its perceived source credibility. For instance, one subject even made the remark, "Computers can't be wrong." On the other hand, the experts, due to their greater processing capability, knowledge, and experience, treated the ESS more as an equal. Considering themselves experts in the financial analysis and commercial lending area, they were more likely to reject the ESS conclusions due to their ego-involvement, and by nature of being experts, they were more critical of the ESS conclusions. Thus, the observations from the group processes are consistent with the social judgment theory and the heightened criticality hypothesis.

Another observation from the group processes is that the use of ESS support induced a higher level of cognitive loafing among the novices than the experts. For instance, some of the novice groups had the tendency to think that the ESS was right all the time, and they spent the majority of their time figuring out "what the system says".
8.1.2.2 Similarities between Experts and Novices

The group processes of the experts and novices were similar in many ways. For instance, there was no difference in the number of explanation types that the experts and novices accessed. These numbers were captured through the computer logs. Both experts and novices followed the sequence suggested by the question set in their problem solving processes and used the same set of financial information and ratios for making the judgments. The problem solving processes of the experts and novices were similar except for the two differences identified and discussed in the previous section. These similarities were ruled out as possible reasons for the difference in judgments between the experts and novices.

Comparison in Number of Explanations Used by Novices and Experts

Table 8-1 and Figure 8-1 show the number of feedforward and feedback explanations used by the experts and novices. There is no significant difference in the average number of feedforward and feedback explanations used by the experts and novices (feedforward: \( t = .40, p = .69 \); feedback: \( t = .67, p = .51 \); total: \( t = .61, p = .56 \)). Overall, more feedforward than feedback explanations were used.

<table>
<thead>
<tr>
<th>Average Number of Explanations Used</th>
<th>Experts</th>
<th>Novices</th>
</tr>
</thead>
<tbody>
<tr>
<td>Feedforward Explanations</td>
<td>9 (54/9)</td>
<td>10.6 (95/9)</td>
</tr>
<tr>
<td>Feedback Explanations</td>
<td>5 (30/6)</td>
<td>7 (63/9)</td>
</tr>
<tr>
<td>Total</td>
<td>14</td>
<td>17.6</td>
</tr>
</tbody>
</table>

*Table 8-1: Average Number of Feedforward and Feedback Explanations Used by Experts and Novices*
Structure of Problem Solving

In the experimental task, the experimental materials and procedures imposed a certain structure on the problem solving process. This was done to control for structure in the control and treatment groups. By structuring the problem into a number of sub-analyses and having the judgment sheets structured in the same manner for all groups, structure was controlled in the experiment and can be ruled out as a potential contributing factor. This may have limited the amount of flexibility in the group processes, which further limits the differences that may be found in the group processes of experts and novices.

Identification of Cues

The literature on expert-novice differences indicates that experts can better identify relevant information and they use fewer cues than novices in making judgments (Davis, 1996; Etterson, Shanteau and Krogstad, 1987). In this experimental task, all of the relevant cues were identified and presented in an organized manner to the subjects in the control and treatment groups. By specifying and limiting the information search process, the problem solving processes of the experts and novices became more controlled.
8.1.2.3 Summary

The ESS analyses and explanations support influenced the novices more than the experts. The novices, in contrast to the experts, tended to trust, rely on, and value what the ESS said. The experts, on the other hand, were more critical of the ESS conclusions and less willing to accept the conclusions. No difference existed between the experts and novices in terms of the total number of explanations sought, so the greater influence on the novices was not a result of seeking more explanations. The structure of the problem solving processes and the ability to identify relevant cues were also ruled out as possible reasons for the difference in the level of ESS influence on the experts and novices.

8.1.3 Higher Consensus Among Novices than Experts

With ESS analyses and explanations support, the novices achieved a higher level of consensus in their judgments than the experts. This finding can best be explained by the heightened criticality hypothesis where experts tend to be critical not only with the ESS conclusions but also with each other. However, it was not apparent from the group interactions that the disagreement among group members was higher in the expert groups than the novice groups, perhaps because of people’s unwillingness to show their disagreement with others openly. The quantitative results indicate that true consensus is more difficult to achieve among a group of experts than a group of novices while the qualitative analysis provides no indication to this result. This difference provides support for the importance of triangulation in empirical research.

8.2 Summary of Chapter 8

This chapter complements the earlier two chapters on the quantitative analysis of results by bringing in the qualitative perspectives. One of the findings in the quantitative analysis is that the knowledge transfer to the novice users increased with the level of ESS support. The
ESS conclusions highlighted important points and issues, helped users identify and correct errors in their inferences, and were used by individuals to reinforce, support, and convince others of their points. On the other hand, the explanations provided by ESS were helpful to decision makers in learning new concepts and their implications, confirming or comparing with one's own knowledge or reasoning process, understanding and resolving disagreements with ESS conclusions, and providing overall assessments of analyses. The ESS, therefore, increased informational influence in the group. However, there was no apparent evidence from the group interactions that normative influence was affected by the use of ESS. In general, the group processes tended to be dominated almost entirely by informational influence in carrying out the financial statement analysis task, regardless of whether the ESS was available as a support tool or not.

Two observations that were apparent from the group interactions support the findings from the quantitative analysis 1) ESS analyses and explanations support influenced novices more than experts, and 2) the level of consensus among the novices was higher than among the experts. The group interactions indicate that the experts were more critical with the ESS conclusions and explanations than the novices were and, therefore, were less willing to accept the conclusions. This explains why the experts were less influenced by the ESS analyses and explanations support than the novices were. On the other hand, cognitive loafing was more evident among the novices than the experts, resulting in the novices relying more heavily on the ESS conclusions than the experts did. In this way, a higher level of consensus was achieved among the novices than the experts.
CHAPTER 9: DISCUSSION AND CONCLUSIONS

This chapter concludes the dissertation with a summary of the research findings and their implications, the contributions of this research, a discussion about the limitations of this research, the problems and challenges encountered in carrying out this research, and future research directions.

9.1 Summary and Implications of Research Findings

The following discussion of the research findings and their implications addresses two central research questions. The first sub-section evaluates the extent to which the use of ESS analyses and explanations support benefits group decision making among novice financial analysis users. The second sub-section evaluates the impact of the use of ESS analyses and explanations support on group decision making by experts versus novices. The third sub-section discusses some of the supplementary findings in this research and the next two sub-sections discuss the implications of the findings for research and practice.

9.1.1 Effects of ESS Support on Group Decision Making by Novices

The amount of knowledge transfer from the ESS to the users increases with the level of ESS support while the level of ESS support has no effect on consensus of judgments. The following sub-sections will discuss the findings in more detailed.

Level of ESS Support and Knowledge Transfer

ESS analyses and explanations support are helpful in increasing the consistency of novices’ judgments with those of the original experts who were involved in developing the system. ESS analyses support is helpful to novices because it provides advice that are useful in making judgments or decisions. ESS explanations support is useful because it increases
novices’ comprehension and understanding of the ESS conclusions. The findings can be explained using the elaboration likelihood model and the literature on lens model.

According to the elaboration likelihood model, persuasion can take place in one of two ways: a central route or systematic perspective where argument-based thinking is being emphasized, and a peripheral route or heuristic perspective where specific psychological mechanisms that do not implicate argument processing, such as positive/negative affect and attractive/expert sources, is being emphasized. In practice, these two routes are not mutually exclusive. In the case of ESS support, the ESS analyses support corresponds to the peripheral route while the ESS explanations support corresponds to the central route. When only ESS analyses support is provided, users may yield to the ESS conclusions because the ESS is perceived to have expert power. Thus, the ESS may have persuaded those who take the peripheral route, but not those who prefer the central route to persuasion. However, when ESS explanations support is provided, users who take the central route to persuasion may now be persuaded by the ESS. In this way, persuasion or knowledge transfer increases with the level of ESS support. This finding is also consistent with the findings from the lens model or “multiple-cue probability learning” literature where it was found that knowledge transfer increases with task support (Balzer, Doherty, and O'Connor, 1989)

Level of ESS Support and Consensus

One of the main findings of this research is that knowledge transfer increases with increased levels of ESS support. With greater knowledge transfer from the ESS to the users, we expect consensus to increase with increased levels of ESS support. However, the results were not so, indicating that the level of ESS support has no effect on consensus. As indicated by some of the subjects in the post-study questionnaire, the ESS support did not help them to reach consensus; consensus was achieved through group discussion.
Level of ESS Support and Perceptions

Consistent with the results of meta-analyses on GDSS research, the use of ESS support for group decision making decreases the users’ satisfaction with the group process. Introducing the ESS into the group process may have created confusion in the way the system was being appropriated. However, the problem is expected to subside as norms on the use of the system emerge over time. Thus, the decreased satisfaction with the group process which arises from uncertainties in the appropriation of the system is expected to improve over time.

The novices did not perceive increased satisfaction with group judgments as the level of ESS support increased even though the novices’ judgments were closer to those of the original experts as the level of support increased. This indicates that decision makers are poor judges of the value of a support system, which is consistent with the finding by Davis and Kottemann (1994). We believe the decreased satisfaction with the group process might have impeded satisfaction with group judgments as the former is a mediator of the latter.

Consistent with our predictions, ESS explanations support increases users’ trust in the ESS. This finding is consistent with the finding of Ye and Johnson (1995). ESS explanations increase novices’ comprehension and understanding of the ESS conclusions, which further lead to increased trust and confidence in the ESS. However, ESS explanations support does not increase users’ perceived usefulness of the ESS. These results differ from the findings of Dhaliwal (1993) and Mao (1995) in which ESS explanations support was evaluated in the individual decision making context. We, therefore, deduce that explanations support is not perceived to be useful in the group decision making context because of the perception that other members in the group may provide the explanations.
Summary

Increased levels of ESS support increase knowledge transfer from the ESS to the users. This finding is consistent with our prediction derived from the elaboration likelihood model and the lens model literature. Assuming that a high quality ESS is available to support group decision making, the use of ESS analyses and explanations support would improve the quality of group decisions. With regard to consensus, the level of ESS support has no effect on consensus in decision making. As indicated by the subjects, group discussion is the main factor leading to consensus in decision making. Consistent with the GDSS literature, the use of ESS support for group decision making decreases users' satisfaction with the group process. However, the level of ESS support has no impact on users' satisfaction with group judgments and perceived usefulness of the ESS, thus providing support for the findings of Davis and Kottemann (1994) that users are poor judges of the value of a support system. Consistent with the findings of Ye and Johnson (1995), ESS explanations support increases users' trust in the ESS.

9.1.2 Effects of ESS Support on Group Decision Making of Novices versus Experts

Knowledge transfer from the ESS to novices is greater than knowledge transfer to experts when ESS analyses and explanations support are provided. This indicates that novices are more likely than experts to buy-in to the analyses and advice given by the ESS. Consensus is also lower among experts than novices, indicating that experts are more likely than novices to disagree with one another. The following sub-sections will discuss the findings in more detailed.

Expertise and Knowledge Transfer

The quantitative results indicate that, with ESS analyses and explanations support, there was greater knowledge transfer from the ESS to the novices than to the experts. Comparing the novices' and experts' judgments with those of the original experts involved in developing
FINALYZER, the novices' judgments were closer to those of the original experts. This indicates that the novices were more willing than the experts to accept the ESS conclusions.

This phenomenon can be explained using persuasion theories. Novices are more willing than experts to accept conclusions from an ESS because they perceive the ESS to be of higher credibility or, more specifically, higher expertise than themselves. According to the elaboration likelihood model, a recipient with less prior knowledge, experience, and ability to process and comprehend a message is more likely to take the peripheral route to persuasion than one who is more knowledgeable and experienced. Hence, novices are more likely to take the peripheral route to persuasion than experts. By taking the peripheral route, novices accept the messages based on peripheral cues, such as the perceived expertise of the ESS.

On the other hand, experts are more likely to take the central route to persuasion because of their greater ability to process messages. However, because they are experts, they also tend to be more critical toward the ESS conclusions and explanations. This is in line with the heightened criticality hypothesis which states that experts process messages in a critical, objective manner (Biek and Wood, 1996). With greater knowledge and experience, experts are capable of critically evaluating the ESS conclusions and, thus, are more likely to generate opposing opinions from the ESS conclusions, which in turn reduce their acceptance of the conclusions. These explain why experts tend to be less influenced by the ESS conclusions and explanations than novices. The phenomenon can also be explained by the social judgment theory. As experts, they tend to be more ego-involved. The social judgment theory predicts that because experts are more ego-involved, they have a fairly large latitude of rejection, indicating that they are more likely to strongly reject arguments that are different from theirs. This prediction is evident in both our quantitative and qualitative results. In the quantitative analysis, the experts' judgments were found to be
further away from those of the original experts than the novices' judgments were, indicating the higher level of rejection of ESS conclusions by the experts. In the qualitative analysis, the experts were found to be particularly critical of the ESS conclusions and were unwilling to accept some of the ESS conclusions even after viewing the explanations. In short, experts are superior to novices in their information processing capability and their critical attitude toward processing messages. On the other hand, their tendency to be ego-involved may work at a disadvantage by diverting them away from taking the information processing approach.

Expertise and Consensus
With ESS analyses and explanations support, novices achieve a higher level of consensus in their judgments than experts. This finding can best be explained by the heightened criticality and ego-involvement of experts where experts tend to be critical not only with the ESS but also among themselves. This results indicate that true consensus is more difficult to achieve among a group of experts than a group of novices.

Expertise and Perceptions
Consistent with our hypothesis, the novices perceived the ESS to be more useful than the experts did. Since novices have less experience and prior knowledge, the ESS analyses and explanations support are more helpful to them than it is to experts. However, it is surprising that the experts and novices did not differ in their perceived level of trust in the ESS. This finding is, nevertheless, consistent with that by Mao (1995). The elaboration likelihood model can be used to explain the insignificant difference in the level of trust between experts and novices. According to the elaboration likelihood model, novices tend to trust an ESS based on peripheral cues, such as the perceived expertise of the ESS. In contrast, peripheral cues have little effect on experts. Experts, being interested in taking the central route to persuasion, would trust an ESS that is equipped with explanation facilities. In this case, since
both ESS analyses and explanations support were provided, the experts would trust the ESS as much as novices because of the availability of the explanation facilities.

Summary

Experts are more critical and objective in processing messages but they also tend to be ego-involved in their area of specialization. Thus, experts are less likely than novices to accept the analyses and advice given by the ESS. With ESS analyses and explanations support, complete consensus is more difficult to achieve among a group of experts than novices. Experts and novices do not differ in their level of satisfaction with the group process and with the group judgments. Although novices and experts do not differ in their level of trust in the ESS, novices perceive the ESS to be more useful than experts do.

9.1.3 Supplementary Findings

In analyzing the use of explanation types by experts and novices, we found that preferences for the use of explanation types varied greatly among different groups. Interestingly, no difference existed in the preferences for use of explanation types by experts and novices (as discussed in Chapter 8), but the differences were large across the different groups. Figures 9-1 and 9-2 show the comparative number of explanation types used by the expert groups and the novice groups. For instance, one novice group did not access any explanations at all while 4 out of 9 novice groups accessed more than 20 explanations. One expert group was so involved in the process that the members had to compete for air time to voice their views, while the other expert groups were better coordinated.
Another related finding is that different groups appropriated the use of the ESS technology differently. More importantly, these differences had no observable relationship with treatment or user expertise. For instance, the degree to which groups were able to comfortably assimilate the ESS into their group processes and the sequence or timing in
which the ESS’s analyses and recommendations were consulted had no direct relationship with the level of user expertise or whether explanations were provided. This is in line with the Adaptive Structuration Theory (AST) proposed by Poole and DeSanctis (1990, 1991). AST helps in understanding and explaining why the outcomes from the introduction of new technology or systems are not uniform across groups. The basic assumption is that groups draw on social structures as rules and resources for interaction; examples of social structure are the rules governing turn-taking in speech, the norm of deference to a leader or other influential person, or specific decision-making procedures that have been learned in the past or are currently available. When an ESS is used, another source of social structures is introduced; for example, individual and group preferences for explanation types, and sequence or timing of accessing ESS analyses and explanations are all methods for structuring group process. The study of group adoption of ESS involves an analysis of the structures available in the technology followed by an analysis of how these are brought into group interaction, how certain structures are preferred over others, and how ESS structures are blended with other structures. The study of the group adoption of an ESS is an important topic; but it is not within the scope of this research. As such, it is not explored further but will be left as future research and extension of this work.

9.1.4 Implications for Research

This research shows that ESS analyses and explanations support have a high potential of supporting group decision making; but its use in supporting group decision making has not been adequately studied by researchers. Future research is needed to evaluate its use in organizations and in dispersed settings, and to incorporate the ESS technology into GDSS.

Is ESS for group support a form of GDSS? A number of researchers have asked this question, especially with regard to this research. ESS, when used to support group decision making, can be considered as a level 2 GDSS. Benbasat and Lim (1993) compared the
moderating effect of level 2 versus level 1 GDSS. Level 2 GDSS generally lead to greater improvement in performance, satisfaction, and consensus than level 1 GDSS; these results highlight the importance of the modeling and structuring capabilities of GDSS. In this research, increased levels of ESS support are found to 1) contribute to greater knowledge transfer, 2) have no effect on consensus and satisfaction with group judgments, and 3) decrease satisfaction with group process. However, the explanations support provided by ESS increases the decision makers’ trust in the system. In terms of performance measure, both ESS and level 2 GDSS benefit group decision making. The effects of ESS and GDSS are also consistent in that these systems decrease satisfaction with group process. However, ESS does not perform as well as other forms of level 2 GDSS in helping groups to reach consensus in decision making. One possible explanation for the difference is that most level 2 GDSS provide cognitive support whereas ESS provides task support. In other words, most level 2 GDSS attempt to make explicit the cognitive information of the decision makers, that is, relations perceived by the decision makers about the task environment. In this way, it helps decision makers to recognize differences among themselves, thus providing greater opportunity for these differences to be resolved. On the other hand, the task support provided by ESS corresponds to the original experts’ knowledge about the relations in the task environment; it does not help to reveal the decision makers’ cognitive information. The “multiple cue probability learning” literature suggests that cognitive support aids decision makers in reaching consensus while task support is responsible for improvement in decision accuracy.

The elaboration likelihood model predicts that explanations support is necessary to convince and persuade experts to accept conclusions provided by ESS. According to the elaboration likelihood model, novices can be persuaded by the ESS conclusions solely based on the perceived expertise of the ESS, or the so called expert power. In contrast,
experts require argument-based reasoning to be persuaded. Empirical evaluation of such a hypothesized difference between experts and novices is warranted.

9.1.5 Implications for Practice

The findings in this study are not only of interest to researchers in group decision making but also valuable to professionals in the financial sector and companies that are considering the use of ESS to complement their group decision making processes.

This research identifies the benefits of using ESS analyses and explanations support for group decision making. Increased levels of ESS support lead to increased knowledge transfer. Assuming that a high quality ESS is available to support group decision making, the use of ESS analyses and explanations support significantly improves the quality of group decisions. Thus, organizations are advised and encouraged to look into the possibility of using ESS to support their group decision making processes.

However, one concern with the use of ESS support for group decision making is that it decreases users’ satisfaction with the group process. Such a finding is consistent with those found in GDSS research. The decreased satisfaction with the group process can be explained by the confusion created when an ESS is introduced into group settings. As pointed out by Zigurs and Kozar (1994), “Computers still encourage extreme reactions in people, probably because we humans find it difficult to place a computer system in an appropriate role in our social world. Such confusion is bound to be exacerbated in a GDSS environment, where group members are interacting with one another and with the system.”

As a result, the group may find it frustrating and difficult to appropriate the system into their group discussion. This phenomenon was also reported by Sviokla (1989) in his study on the use of PlanPower, an ESS that supports financial planning at the Financial Collaborative. PlanPower was later removed from the group meetings and placed in the
back-room for consultation only after the group decisions were made. The reasons given for doing so were that PlanPower stifled their creativity, made the meetings less interesting, and caused the financial planners to feel less important. Interestingly and in contrast, the use of an ESS at Imperial Chemical Industries was proving very successful and effective in supporting the group sessions. We believe that, as norms on the use of ESS develop over time and if a high level of interactivity and flexibility are incorporated into a high quality ESS, the general unsatisfaction with group process can be overcome. For instance, the incorporation of hypertext capabilities into ESS increases its interactivity (Mao, 1995), and the incorporation of ESS into GDSS could increase its overall flexibility. Some development work on ESS and the use of longitudinal studies are needed to assess our claim and other long-term effects of ESS usage.

This research also highlights the importance of explanations support. As indicated by the results, the combination of the ESS analyses and explanations support is responsible for the transfer of knowledge from the ESS to the users. Thus, explanations support is important for groups to buy-in to the ESS conclusions and to increase their trust and confidence in the ESS.

9.2 Contributions of Research

Satzinger and Olfman (1995) pose an unanswered (or poorly answered) question in GDSS research: “Of all the types of support that might be provided under the umbrella of groupware, which approaches are perceived to be most useful?” This research attempts to partially answer this question by investigating the usefulness of ESS and its explanation facilities in supporting group decision making. If ESS support is found to be useful in supporting group decision making, organizations may wish to incorporate its capabilities into GDSS.
The contributions of this research can be viewed from an empirical and theoretical perspective. From an empirical perspective, this research extends the use of the ESS technology to the group context. To the best of our knowledge, it is the first study to offer empirical evidence on the effects of using ESS analyses and explanations support for group decision making. In addition, it compares the effects of ESS usage on expert and novice groups, and discusses the similarities and differences in their group processes. To date, few studies have examined group decision making by experts versus novices, and this study is believed to be the first attempt to carry out such a comparison in the MIS literature.

From a theoretical perspective, it links ES/ESS research with group decision making and persuasion theories by using concepts from the literature to explain the phenomena observed in this study. Two phenomena were explained using theories from the persuasion and group decision making research – the effects of the use of different levels of ESS support on group decision making of novices, and the effects of the use of ESS analyses and explanations support on group decision making of novices versus experts. Not only has there been few research on decision making by groups of experts, the integration of persuasion theories with the empirical evaluation of decision making by groups of experts versus novices is an important contribution to the MIS literature. Another contribution is its use of persuasion theories to explain the impact of different levels of ESS support. This work also qualifies as interdisciplinary research that links computer science, or more specifically, AI, with psychology and sociology theories.
9.3 Limitations, Problems and Challenges

The limitations of this research and the problems and challenges encountered in carrying out this research are discussed under four main categories:

(1) Limited sample size;
(2) Cross-sectional study;
(3) Generalizability;
(4) Process tracing.

Limited Sample Size

The sample size was restricted by the number of qualified subjects. In this study, a financial analysis and commercial lending task was chosen for two reasons: 1) the expert system technology has a high potential in supporting the financial analysis and commercial lending domain, and 2) a prototype ESS is available. The domain-specific nature of the task severely restricted the number of qualified subjects for this study. In the business world, only individuals who have considerable knowledge of accounting and financial statement analysis are involved in carrying out such tasks. For the experiment to be realistic, the subjects participating in this study must also satisfy these criteria. Through an extensive recruitment process, we managed to recruit 75 student subjects for the study, representing about 33% of the pool of qualified student subjects — senior undergraduate students who were majoring in Accounting or students (undergraduate or graduate) who were taking the Financial Statement Analysis course. The recruitment of expert subjects was a bigger problem. Many of the professionals who qualified as experts for the study had taken part in earlier studies carried out by Dhaliwal (1993) and Mao (1995), thus severely restricting the number of expert subjects available for this study. Through direct contact with four major financial institutions, we managed to recruit 21 expert subjects. Despite the large amount of effort taken in the recruitment process, the sample size problem still exists.
This is a limitation in this research and in other studies involving groups because a relatively larger sample size is required for studies involving groups.

Due to the limited number of experts available for this study, we were unable to study the effects of different levels of ESS support on decision making by expert groups. Instead, we assigned all of the available expert subjects into the most important cell with respect to our research objective, thus creating an unbalanced design.

To alleviate the sample size problem, we also took appropriate steps in our statistical approaches. The sample size problem was overcome by the application of non-parametric statistics and nested ANOVA designs. To increase the power of the statistical tests, the dependent variables were analyzed at the individual level using a nested ANOVA design whenever appropriate, and non-parametric statistics were applied whenever the assumptions of the parametric model were not satisfied. The only assumption of non-parametric statistics is that the distribution is continuous, unlike the numerous assumptions required for parametric statistics.

Cross-sectional Study
This study was conducted at one point in time and the ESS was used only once by the subjects; the effect over time was not studied. Although cross-sectional study is the norm in most experimental settings, it poses a potential problem due to novelty effects. In other words, we examined the use of an ESS and its explanation facilities for first time users. The subsequent appropriation of the technology and effects were not investigated.
Generalizability

The experimental task in this research relates to financial analysis and commercial loan decision making. Although we would like to generalize the findings to other task domains, the generalizability of the findings will be limited until the study is replicated in other domains.

Process Tracing

Another problem encountered in this research is the coding of the protocols. To supplement the quantitative analysis with a qualitative perspective so that richer explanations and supporting evidence could be provided, we carried out systematic coding of the protocols to identify the reasons for the use of different explanation types. However, a major problem was encountered during the coding process. Many of the times, the subjects in the study did not verbalize or make explicit their reasons for accessing explanations. This problem is similar in concept to that encountered in survey research where response rate is typically low. With low response rate, the results would be unreliable as information involving non-respondents is unknown. In this case, only about 50% (123/244) of the total number of accesses to explanations were identifiable with a reason, giving us a known rate (or "response rate") of only 50%. This finding is both disappointing and "frustrating" to the researcher given the large amount of time and effort invested in the coding process, not to mention the tediousness of the process.

Because of the problems encountered in using the systematic coding method to analyze the use of ESS explanations, conversations and interactions of the group processes were analyzed to identify cues and anecdotal evidence to supplement the quantitative results. The analysis of conversations and interactions was carried out to identify the different ways in which the ESS analyses and explanations support were used by the groups, the problems encountered in the absence of explanations support, the benefits of ESS analyses
support, and the similarities and differences in the group processes of experts and novices. The computer logs of the experimental sessions, which recorded the commands carried out on the ESS, were also used for process tracing to analyze the groups' interaction with the ESS.

9.4 Directions for Future Research

The following are six directions to extend this research:

1. Comparing the use of ESS support for individual versus group decision making;
2. Analyzing the effects of ESS analyses and explanations support by domain experts;
3. Analyzing the effects of ESS analyses and explanations support longitudinally;
4. Analyzing the appropriation of the ESS technology; and
5. Replicating the study under different context and area of domain.

Comparing the Use of ESS Support for Individual versus Group Decision Making

Dhaliwal (1993) and Mao (1995) carried out similar studies in the individual decision making context, investigating various aspects of the ESS explanation facilities. A natural extension of this study would be to compare 1) the preferences for use of explanation types between the individual and group decision making context, and 2) the individual and group decision making processes involved in the use of the ESS technology. FINALYZER was also used in both of the earlier studies. However, one concern needs to be resolved before making the comparisons. Because Dhaliwal (1993) and Mao (1995) were interested in looking at the change in users' agreement with the ESS conclusions before and after viewing explanations, they captured the rating both before and after explanations were viewed. In this study, we did not capture the rating because it does not fall within the objectives of this research. We recognize that, by capturing (or not capturing) the agreement rating, the users' decision making processes and interaction with the ESS may change. Due to this limitation, we did not include the comparisons in this dissertation. We
will take caution in our future research to resolve the limitation and its related potential effects when making the comparisons.

Analyzing the Effects of ESS Analyses and Explanations Support by Domain Experts
The original intention of this study was to evaluate the effects of ESS analyses and explanations support by both novices and experts using a balanced 6-cell experimental design. However, due to the limited number of expert subjects available, the effects of these ESS features on the experts were not studied. Instead of assigning the expert subjects to the three levels of decision support conditions as originally intended, all of the available expert subjects were assigned to the full ESS support treatment. In our future research, we would like to extend this study by evaluating the impact of ESS analyses and explanations support on domain experts. Since the novices in our study were not naive, but were fairly knowledgeable, in the financial analysis domain, we would expect the results we have obtained with novice users on the effect of ESS analyses and explanations support to also hold for the experts.

Analyzing the Effects of ESS Analyses and Explanations Support on a Longitudinal Basis
As mentioned earlier, this study was carried out in a one-shot manner, which means it was the first time that experts and novices were exposed to the ESS. As such, the effect of novelty cannot be overlooked. For future research, we are interested in carrying out a longitudinal study to examine the long-term and on-going effects of the use of ESS by groups.

Analyzing the Appropriation of the ESS Technology
One of the findings in this study is that groups appropriated the use of the ESS technology in different ways. This supports the Adaptive Structuration Theory (AST) proposed by Poole and DeSanctis (1990, 1991). The AST may be able to help us in understanding and
explaining why the outcomes from the introduction of ESS technology or systems are not uniform across groups. A case and interpretive approach may be undertaken in the future to study the appropriation of the ESS technology as a supplement to this research.

Replicating the Study under Different Context and Domain

This study may be replicated under different contexts and in various domains to increase the generalizability of the results. An interesting extension of this study would be to evaluate the use of the ESS under different group decision making scenarios, such as providing each group member with access to the ESS as well as evaluating the use of the system in dispersed settings. The effect of ESS in these various scenarios warrants investigation. Although we believe that the findings in our study will hold in the dispersed setting, this, however, needs to be tested. Finally, we are also interested in generalizing our findings to different domains, to organizational decision making context, and to a real setting. To do so, we will need to replicate the study in different domain areas, in organizations where a larger number of people are involved in the decision making processes, and in the industry.
References


269


LIST OF APPENDICES

A. Experimental Materials
B. Sample Screens of FINALYZER
C. Materials for Recruitment of Subjects
APPENDIX A: EXPERIMENTAL MATERIALS
FINANCIAL ANALYSIS GROUP DECISION MAKING STUDY

CONTENTS:

- Information Sheet
- Consent Form
- Background Information Questionnaire
- Individual Evaluation of Canacom Corporation Loan Analysis Case
- Judgment Recording Sheets (Set 1)
- Group Evaluation of Canacom Corporation Loan Analysis Case
- A Short Note on Expert Systems
- Tutorial on the CREDIT-ADVISOR Expert System
- Using FINALYZER to Support your Group Decision Making Process
- Group Judgment Recording Sheets
- Final Evaluation of Canacom Corporation Loan Analysis Case
- Judgment Recording Sheets (Set 2)
- Post-Study Questionnaire
- Debriefing Protocol
FINANCIAL ANALYSIS GROUP DECISION MAKING STUDY

Information Sheet

We are currently studying issues related to the design of computer-based expert support systems for financial analysis group decision making. The objectives of the study are to measure the value of the systems to the users and to understand the manner in which they should be designed to be of greatest value to the users. In the initial stages of the study, an expert support system has been developed. Plans are currently being made to evaluate this system in terms of facilitating financial analysis for loan decision making.

Your participation will involve making decisions both individually and in a group. First, you will work individually to perform the financial statement analysis of a hypothetical company and to make a set of judgments regarding the financial health of the company. Next, you will be introduced to two other members who will carry out the financial statement analysis task with you as a group to derive a set of mutually agreeable group judgments. Your group will be provided with a financial analysis expert support system called FINALYZER which you and your group members will use to perform the task. Upon completing the task, you will be asked to re-evaluate the individual judgments you have made earlier by taking into account all the information and knowledge you have gathered from the FINALYZER system and your group members. Finally, you will be asked to provide evaluations of the usefulness of the FINALYZER system and your perceptions of the group process and outcome.

It is estimated that your participation will take between 2 to 3 hours. You will be paid an honorarium of $25 for participating in this study. In addition, each participant will have a one in five chance of winning a $30 cash award. That is, a $90 cash award will be given to each of the top 20% performing groups. Any information obtained in connection with your participation that can be identified with you will remain confidential. In any reports or publications, only aggregate data will be presented to preserve the anonymity of participants. Although part of the study will be videotaped, the video clips captured during the experiment will not be viewed by any other parties other than the researchers. Participation in this study is completely voluntary and you are free to discontinue participation at any time.

If you have any questions now, please approach the research assistant. If you have questions later, please contact Professor Izak Benbasat at 822-8396 or Ms. Fui Hoon Nah at 224-6264.

Thank you. Your participation is greatly appreciated.
CONSENT FORM

This form is to be completed after you have read and understood the contents of the Information Sheet that is attached.

Agreement to Participate:

I have read and understood the contents of the Information Sheet provided and have decided to participate in the study.

Agreement to Confidentiality:

I understand that some of my classmates and friends may also be participating in this study. I realize that my discussion of the details of this study with them may distort the results. Therefore, I agree not to discuss with any other participant any aspect of the study prior to their participation.

__________________________________________  ______________
Signature of Participant                          Date

Participant's Name: ___________________________

We thank you very much for your assistance in this study.
BACKGROUND INFORMATION QUESTIONNAIRE

1. Name (please print): ____________________________

2. Contact Address: ________________________________________

3. Contact Telephone Number: ____________________________

4. Educational and Professional Qualifications (College, Undergraduate/Graduate degree, etc.)
   and area of specialization: (please specify the number of years completed if in-progress)

   ______________________________________________________

   ______________________________________________________

   Professional Affiliations (e.g., CFA, ACIB, CA, CGA, CMA, etc.): __________

5. List all courses you have taken or are currently taking that dealt with aspects of financial
   ratios or financial statement analysis:

   ______________________________________________________

   ______________________________________________________

6. Have you ever held any full-time, part-time, or summer jobs that involved analyzing financial
   ratios? If yes, briefly describe the job(s) and the length of the period of each job, state
   NONE otherwise:

   ______________________________________________________

   ______________________________________________________

7. List the types of financial modelling and analysis software (e.g., Lotus 1-2-3, FISCAL,
   FSAP, etc.) that you have used or are currently using (state NONE otherwise):

   ______________________________________________________

   ______________________________________________________

8. Have you ever used expert systems in the past? __________

   If yes, briefly describe the expert system(s) and the circumstances of your use (state
   NONE otherwise):

   ______________________________________________________

   ______________________________________________________

281
9. I would rate my familiarity with expert systems in general as:

   Not Familiar at all: 1 - 2 - 3 - 4 - 5 - 6 - 7 : Very Familiar

10. I would rate my familiarity with decision support systems or financial modelling and analysis software in general as:

   Not Familiar at all: 1 - 2 - 3 - 4 - 5 - 6 - 7 : Very Familiar

11. I usually dominate during group discussion.

   Strongly Disagree: 1 - 2 - 3 - 4 - 5 - 6 - 7 : Strongly Agree

12. I prefer group problem solving to individual problem solving.

   Strongly Disagree: 1 - 2 - 3 - 4 - 5 - 6 - 7 : Strongly Agree

13. I like working in groups.

   Strongly Disagree: 1 - 2 - 3 - 4 - 5 - 6 - 7 : Strongly Agree

14. How do you rate yourself as a financial analyst (of corporate loan decision)?

   (1 -- Excellent, 2 -- Good, 3 -- Somewhat good, 4 -- Fair, 5 -- Somewhat poor, 6 -- Poor, 7 -- Bad)
CANACOM CORPORATION LOAN ANALYSIS CASE

Assume that you are a corporate loan evaluation officer working for a large financial institution in Western Canada. Your supervisor has asked you to evaluate the attached financial statements of Vancouver-based Canacom Corporation. These statements are the critical components of a more complete application for borrowings of $800 million that has been filed by the company for the purposes of streamlining operations. Your assessment of the company will form the basis of a more comprehensive loan evaluation to be undertaken by the Corporate Loans Committee, of which your supervisor is a member. Note that you have been told that the total repayment period will not exceed three years and that there is the possibility of the loan being convertible into stock at the end of that period.

Canacom Corporation, through its chain of Computron Corner outlets, is one of North America's leading distributors of technology to individual consumers. Close to 75 percent of its business is in the United States, with 15 percent being in Canada, and the balance is mostly in Europe. Through more than 5700 company-owned retail stores and 3000 dealer/franchise outlets, Canacom manufactures and distributes a broad product line that includes microcomputers and related software; televisions, radios, audio equipment, tape recorders, and related accessories; toys, antennas, security devices, timers and calculators; electronic parts, batteries, and test equipment among other products. While microcomputers, software, and peripheral equipment were not part of the company product line five years ago and represented only 2.4 percent of sales, they represented the largest component, 34.6 percent, of total sales in 1994.

The financial statements of Canacom for 1990-94 and additional relevant qualitative information are presented below. The auditor's opinions on the financial statements have been unqualified for the past five years. To help you with your analysis, common-size financial statements and a complete set of the relevant financial ratios have been prepared for you using a computerised financial analysis package. Comparative information of Hightech Computer Corporation, a Pittsburgh-based major competitor of Canacom is provided. Additionally, the industry composites of the electronic computing equipment manufacturing segment and the radios, televisions, and record players retail segment have been obtained for your use.
You have been specifically instructed to focus on all aspects of Canacom's valuation, liquidity, long-term solvency, asset utilization, and profitability. As part of your report, you will have to provide specific judgments for: 1) the exact amount of the $800 million loan requested that you recommend as being allowable assuming that no collateral or guarantees are provided, 2) a specific estimate of the expected total net earnings of the company in the coming year, 3) ratings of the quality of Canacom's financial management, operating management, liquidity position, long-term solvency position, asset utilization performance, and 4) a rating of the value of Canacom stock as loan collateral. You will also have to provide your subjective probabilities of the correctness of these four judgments. Please review the attached Judgement Recording Sheets now to understand the exact format in which these judgments are to be recorded.

== Financial Statements and Ratio Tables Attached ==
## TABLE 1

### CANACOM CORPORATION
Balance Sheets (In Thousands)
December 31, 1990-1994

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>ASSETS</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cash &amp; Equivalents</td>
<td>37,621</td>
<td>56,365</td>
<td>141,944</td>
<td>167,547</td>
<td>279,743</td>
</tr>
<tr>
<td>Accounts &amp; Notes Receivable (Net)</td>
<td>15,841</td>
<td>25,725</td>
<td>42,088</td>
<td>83,616</td>
<td>107,530</td>
</tr>
<tr>
<td>Inventories</td>
<td>381,649</td>
<td>435,160</td>
<td>513,709</td>
<td>670,568</td>
<td>844,097</td>
</tr>
<tr>
<td>Other Assets</td>
<td>12,590</td>
<td>13,809</td>
<td>11,416</td>
<td>27,000</td>
<td>31,928</td>
</tr>
<tr>
<td><strong>Total Current Assets</strong></td>
<td>447,701</td>
<td>531,059</td>
<td>709,157</td>
<td>948,731</td>
<td>1,263,298</td>
</tr>
<tr>
<td>Net Property and Equipment</td>
<td>156,670</td>
<td>165,140</td>
<td>190,429</td>
<td>224,995</td>
<td>257,620</td>
</tr>
<tr>
<td>Other Assets</td>
<td>5,218</td>
<td>14,099</td>
<td>36,909</td>
<td>53,918</td>
<td>60,990</td>
</tr>
<tr>
<td><strong>TOTAL ASSETS</strong></td>
<td>609,589</td>
<td>710,298</td>
<td>936,545</td>
<td>1,227,644</td>
<td>1,581,908</td>
</tr>
</tbody>
</table>

<p>| | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>LIABILITIES AND CAPITAL</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Notes Payable</td>
<td>37,189</td>
<td>25,918</td>
<td>34,862</td>
<td>24,942</td>
<td>55,737</td>
</tr>
<tr>
<td>Accounts Payable</td>
<td>34,390</td>
<td>58,926</td>
<td>54,560</td>
<td>63,641</td>
<td>64,640</td>
</tr>
<tr>
<td>Accrued Expenses</td>
<td>52,343</td>
<td>59,170</td>
<td>67,206</td>
<td>92,125</td>
<td>115,054</td>
</tr>
<tr>
<td>Income Tax Payable</td>
<td>13,931</td>
<td>24,703</td>
<td>47,152</td>
<td>52,160</td>
<td>50,668</td>
</tr>
<tr>
<td><strong>Total Current Liabilities</strong></td>
<td>137,853</td>
<td>168,717</td>
<td>203,780</td>
<td>232,868</td>
<td>286,099</td>
</tr>
<tr>
<td>Long-Term Notes Payable</td>
<td>8,688</td>
<td>6,523</td>
<td>3,903</td>
<td>20,642</td>
<td>15,482</td>
</tr>
<tr>
<td>Debentures (Net)</td>
<td>222,045</td>
<td>222,175</td>
<td>122,428</td>
<td>122,666</td>
<td>122,938</td>
</tr>
<tr>
<td>Store Managers Deposits</td>
<td>16,718</td>
<td>14,045</td>
<td>11,972</td>
<td>9,306</td>
<td>8,490</td>
</tr>
<tr>
<td>Deferred Income Taxes</td>
<td>10,978</td>
<td>8,902</td>
<td>12,069</td>
<td>18,886</td>
<td>17,682</td>
</tr>
<tr>
<td>Other Long-Term Liabilities</td>
<td>4,954</td>
<td>6,811</td>
<td>10,530</td>
<td>10,599</td>
<td>10,345</td>
</tr>
<tr>
<td><strong>Total Liabilities</strong></td>
<td>401,146</td>
<td>427,173</td>
<td>364,682</td>
<td>414,967</td>
<td>461,036</td>
</tr>
<tr>
<td>Shareholders' Equity</td>
<td>208,353</td>
<td>283,125</td>
<td>571,863</td>
<td>812,677</td>
<td>1,120,872</td>
</tr>
<tr>
<td><strong>TOTAL LIABILITIES AND CAPITAL</strong></td>
<td>609,589</td>
<td>710,298</td>
<td>936,545</td>
<td>1,227,644</td>
<td>1,581,908</td>
</tr>
</tbody>
</table>

<p>| <strong>285</strong> |</p>
<table>
<thead>
<tr>
<th>CANACOM CORPORATION</th>
</tr>
</thead>
<tbody>
<tr>
<td>Income Statements (In Thousands)</td>
</tr>
<tr>
<td>December 31, 1990-1994</td>
</tr>
<tr>
<td>--------------------------------------</td>
</tr>
<tr>
<td><strong>Net Sales</strong></td>
</tr>
<tr>
<td>1,215,483</td>
</tr>
<tr>
<td>1,384,637</td>
</tr>
<tr>
<td>1,691,373</td>
</tr>
<tr>
<td>2,032,555</td>
</tr>
<tr>
<td>2,475,188</td>
</tr>
<tr>
<td><strong>Other Revenue</strong></td>
</tr>
<tr>
<td>11,403</td>
</tr>
<tr>
<td>11,360</td>
</tr>
<tr>
<td>15,697</td>
</tr>
<tr>
<td>28,657</td>
</tr>
<tr>
<td>38,109</td>
</tr>
<tr>
<td><strong>Total Revenue</strong></td>
</tr>
<tr>
<td>1,226,886</td>
</tr>
<tr>
<td>1,395,997</td>
</tr>
<tr>
<td>1,707,070</td>
</tr>
<tr>
<td>2,061,212</td>
</tr>
<tr>
<td>2,513,297</td>
</tr>
<tr>
<td><strong>Cost of Goods Sold</strong></td>
</tr>
<tr>
<td>535,549</td>
</tr>
<tr>
<td>594,841</td>
</tr>
<tr>
<td>701,777</td>
</tr>
<tr>
<td>826,842</td>
</tr>
<tr>
<td>1,008,187</td>
</tr>
<tr>
<td><strong>Gross Income</strong></td>
</tr>
<tr>
<td>691,337</td>
</tr>
<tr>
<td>801,156</td>
</tr>
<tr>
<td>1,005,293</td>
</tr>
<tr>
<td>1,234,370</td>
</tr>
<tr>
<td>1,505,110</td>
</tr>
<tr>
<td><strong>Selling &amp; Administrative Expenses</strong></td>
</tr>
<tr>
<td>484,249</td>
</tr>
<tr>
<td>546,325</td>
</tr>
<tr>
<td>645,934</td>
</tr>
<tr>
<td>780,378</td>
</tr>
<tr>
<td>930,244</td>
</tr>
<tr>
<td><strong>Depreciation &amp; Amortization</strong></td>
</tr>
<tr>
<td>17,121</td>
</tr>
<tr>
<td>19,110</td>
</tr>
<tr>
<td>23,288</td>
</tr>
<tr>
<td>29,437</td>
</tr>
<tr>
<td>38,679</td>
</tr>
<tr>
<td><strong>Operating Income</strong></td>
</tr>
<tr>
<td>189,967</td>
</tr>
<tr>
<td>235,721</td>
</tr>
<tr>
<td>336,071</td>
</tr>
<tr>
<td>424,555</td>
</tr>
<tr>
<td>536,187</td>
</tr>
<tr>
<td><strong>Interest Expenses</strong>*</td>
</tr>
<tr>
<td>28,466</td>
</tr>
<tr>
<td>25,063</td>
</tr>
<tr>
<td>15,454</td>
</tr>
<tr>
<td>1,168</td>
</tr>
<tr>
<td>8,905</td>
</tr>
<tr>
<td><strong>Net Income Before Tax</strong></td>
</tr>
<tr>
<td>161,501</td>
</tr>
<tr>
<td>210,658</td>
</tr>
<tr>
<td>320,617</td>
</tr>
<tr>
<td>423,387</td>
</tr>
<tr>
<td>527,282</td>
</tr>
<tr>
<td><strong>Provision for Taxes</strong></td>
</tr>
<tr>
<td>78,272</td>
</tr>
<tr>
<td>98,423</td>
</tr>
<tr>
<td>151,015</td>
</tr>
<tr>
<td>199,302</td>
</tr>
<tr>
<td>248,761</td>
</tr>
<tr>
<td><strong>NET INCOME AFTER TAX</strong></td>
</tr>
<tr>
<td>83,229</td>
</tr>
<tr>
<td>112,235</td>
</tr>
<tr>
<td>169,602</td>
</tr>
<tr>
<td>224,085</td>
</tr>
<tr>
<td>278,521</td>
</tr>
<tr>
<td><strong>Common Shares Outstanding</strong></td>
</tr>
<tr>
<td>106,004</td>
</tr>
<tr>
<td>103,644</td>
</tr>
<tr>
<td>102,578</td>
</tr>
<tr>
<td>103,395</td>
</tr>
<tr>
<td>104,335</td>
</tr>
<tr>
<td><strong>Net Income per Share</strong></td>
</tr>
<tr>
<td>$0.79</td>
</tr>
<tr>
<td>$1.08</td>
</tr>
<tr>
<td>$1.65</td>
</tr>
<tr>
<td>$2.17</td>
</tr>
<tr>
<td>$2.67</td>
</tr>
<tr>
<td><strong>Stock Price</strong></td>
</tr>
<tr>
<td>$5.34</td>
</tr>
<tr>
<td>$10.38</td>
</tr>
<tr>
<td>$30.00</td>
</tr>
<tr>
<td>$27.50</td>
</tr>
<tr>
<td>$50.00</td>
</tr>
<tr>
<td><strong>Includes manufacturing payroll</strong></td>
</tr>
<tr>
<td>28,344</td>
</tr>
<tr>
<td>32,958</td>
</tr>
<tr>
<td>42,128</td>
</tr>
<tr>
<td>53,105</td>
</tr>
<tr>
<td>71,892</td>
</tr>
<tr>
<td><strong>Include:</strong></td>
</tr>
<tr>
<td>Nonmanufacturing payroll</td>
</tr>
<tr>
<td>206,507</td>
</tr>
<tr>
<td>232,569</td>
</tr>
<tr>
<td>286,494</td>
</tr>
<tr>
<td>339,559</td>
</tr>
<tr>
<td>395,135</td>
</tr>
<tr>
<td>Advertising expense</td>
</tr>
<tr>
<td>114,238</td>
</tr>
<tr>
<td>124,138</td>
</tr>
<tr>
<td>137,722</td>
</tr>
<tr>
<td>160,905</td>
</tr>
<tr>
<td>199,128</td>
</tr>
<tr>
<td>Rental Expense</td>
</tr>
<tr>
<td>54,606</td>
</tr>
<tr>
<td>61,491</td>
</tr>
<tr>
<td>73,857</td>
</tr>
<tr>
<td>89,732</td>
</tr>
<tr>
<td>106,970</td>
</tr>
<tr>
<td>Foreign currency translation</td>
</tr>
<tr>
<td>3,230</td>
</tr>
<tr>
<td>1,722</td>
</tr>
<tr>
<td>-5,295</td>
</tr>
<tr>
<td>3,216</td>
</tr>
<tr>
<td>590</td>
</tr>
<tr>
<td>*<strong>Net of interest income of:</strong></td>
</tr>
<tr>
<td>1,234</td>
</tr>
<tr>
<td>2,334</td>
</tr>
<tr>
<td>7,179</td>
</tr>
<tr>
<td>20,946</td>
</tr>
<tr>
<td>15,139</td>
</tr>
<tr>
<td>--------</td>
</tr>
<tr>
<td><strong>Cash Flows from Operations</strong></td>
</tr>
<tr>
<td>Net Income</td>
</tr>
<tr>
<td>Add (deduct) items not affecting cash</td>
</tr>
<tr>
<td>Depreciation Expense</td>
</tr>
<tr>
<td>Increase in Accounts Receivable</td>
</tr>
<tr>
<td>Increase in Inventories</td>
</tr>
<tr>
<td>Change in Accounts Payable</td>
</tr>
<tr>
<td>Other</td>
</tr>
<tr>
<td>Total of items not affecting cash</td>
</tr>
<tr>
<td><strong>Net Cash Flow from Operating Activities</strong></td>
</tr>
<tr>
<td><strong>Cash Flows used by Investing Activities</strong></td>
</tr>
<tr>
<td>Net Purchases of Land and Equipment</td>
</tr>
<tr>
<td>Net Reductions in Long-Term Debt</td>
</tr>
<tr>
<td>Purchase of Treasury Stock</td>
</tr>
<tr>
<td>Sale of Treasury Stock to Employees</td>
</tr>
<tr>
<td>Issue of Debentures</td>
</tr>
<tr>
<td>Foreign Currency Adjustments</td>
</tr>
<tr>
<td>Other</td>
</tr>
<tr>
<td><strong>Cash Provided by Financing Activities</strong></td>
</tr>
<tr>
<td>Cash at the Beginning of the Year</td>
</tr>
<tr>
<td><strong>Change in Cash During the Year</strong></td>
</tr>
<tr>
<td>Cash at the End of the Year</td>
</tr>
<tr>
<td>------------------------------</td>
</tr>
<tr>
<td><strong>ASSETS</strong></td>
</tr>
<tr>
<td>Cash &amp; Equivalents</td>
</tr>
<tr>
<td>Accounts &amp; Notes</td>
</tr>
<tr>
<td>Receivable (Net)</td>
</tr>
<tr>
<td>Inventories</td>
</tr>
<tr>
<td>Other Assets</td>
</tr>
<tr>
<td><strong>Total Current Assets</strong></td>
</tr>
<tr>
<td>Net Property and Equipment</td>
</tr>
<tr>
<td>Other Assets</td>
</tr>
<tr>
<td><strong>TOTAL ASSETS</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><strong>LIABILITIES AND CAPITAL</strong></td>
</tr>
<tr>
<td>Current Liabilities</td>
</tr>
<tr>
<td>Debentures (Net)</td>
</tr>
<tr>
<td>Store Managers Deposits</td>
</tr>
<tr>
<td>Deferred Income Taxes</td>
</tr>
<tr>
<td>Other Long-Term Liabilities</td>
</tr>
<tr>
<td><strong>Total Liabilities</strong></td>
</tr>
<tr>
<td>Shareholders' Equity</td>
</tr>
<tr>
<td><strong>TOTAL LIABILITIES AND CAPITAL</strong></td>
</tr>
<tr>
<td>----------------------</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Net Sales</td>
</tr>
<tr>
<td>Other Revenue</td>
</tr>
<tr>
<td>Total Revenue</td>
</tr>
<tr>
<td>Cost of Goods Sold*</td>
</tr>
<tr>
<td>Gross Income</td>
</tr>
<tr>
<td>Selling &amp; Administrative Expenses**</td>
</tr>
<tr>
<td>Depreciation &amp; Amortization</td>
</tr>
<tr>
<td>Operating Income</td>
</tr>
<tr>
<td>Interest Expenses***</td>
</tr>
<tr>
<td>Net Income Before Tax</td>
</tr>
<tr>
<td>Provision for Taxes</td>
</tr>
<tr>
<td>Net Income After Tax</td>
</tr>
</tbody>
</table>

*Includes manufacturing payroll

** Includes:
- Nonmanufacturing payroll
- Advertising expense
- Rental Expense
- Foreign currency translation

*** Net of interest income of:
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Current Ratio</strong></td>
<td>3.25</td>
<td>3.15</td>
<td>3.48</td>
<td>4.07</td>
<td>4.42</td>
<td>3.64</td>
<td>2.70</td>
<td>1.40</td>
</tr>
<tr>
<td><strong>Acid-Test Ratio</strong></td>
<td>0.39</td>
<td>0.49</td>
<td>0.90</td>
<td>1.08</td>
<td>1.35</td>
<td>2.17</td>
<td>1.10</td>
<td>0.30</td>
</tr>
<tr>
<td><strong>Accounts Receivable</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Turnover</td>
<td>76.73</td>
<td>53.82</td>
<td>40.19</td>
<td>24.31</td>
<td>23.02</td>
<td>7.20</td>
<td>5.60</td>
<td>45.50</td>
</tr>
<tr>
<td>Inventory Turnover</td>
<td>1.40</td>
<td>1.37</td>
<td>1.37</td>
<td>1.23</td>
<td>1.19</td>
<td>3.55</td>
<td>3.10</td>
<td>3.60</td>
</tr>
<tr>
<td>Day Sales in Receivables</td>
<td>4.80</td>
<td>6.80</td>
<td>9.10</td>
<td>15.00</td>
<td>15.90</td>
<td>50.70</td>
<td>65.20</td>
<td>8.00</td>
</tr>
<tr>
<td>Days to Sell Inventory</td>
<td>260.70</td>
<td>266.40</td>
<td>266.40</td>
<td>296.70</td>
<td>306.70</td>
<td>102.80</td>
<td>117.70</td>
<td>101.40</td>
</tr>
<tr>
<td>Conversion Period (days)</td>
<td>265.50</td>
<td>273.20</td>
<td>275.50</td>
<td>311.70</td>
<td>322.60</td>
<td>153.50</td>
<td>182.90</td>
<td>109.40</td>
</tr>
<tr>
<td>%Cash to Current Assets</td>
<td>8.40</td>
<td>10.61</td>
<td>20.02</td>
<td>17.66</td>
<td>22.14</td>
<td>30.55</td>
<td>14.80</td>
<td>11.60</td>
</tr>
<tr>
<td>%Cash to Current Liabilities</td>
<td>27.29</td>
<td>33.41</td>
<td>69.68</td>
<td>71.95</td>
<td>97.78</td>
<td>1112.57</td>
<td>29.60</td>
<td>16.50</td>
</tr>
<tr>
<td>Working Capital (millions$)</td>
<td>309.85</td>
<td>362.34</td>
<td>505.43</td>
<td>715.86</td>
<td>977.2</td>
<td>340.21</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>Liquidity Index</td>
<td>233.10</td>
<td>230.20</td>
<td>203.40</td>
<td>228.10</td>
<td>222.50</td>
<td>51.1</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>Operating Cash Flow to Total Current Liabilities</td>
<td>0.73</td>
<td>0.78</td>
<td>0.95</td>
<td>1.09</td>
<td>1.11</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>--------------------------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------------------------</td>
<td>-----------------------------</td>
<td>--------------------------</td>
</tr>
<tr>
<td>Equity to Total Liabilities</td>
<td>0.54</td>
<td>0.68</td>
<td>1.61</td>
<td>2.03</td>
<td>2.50</td>
<td>2.61</td>
<td>0.92</td>
<td>0.46</td>
</tr>
<tr>
<td>Equity to Long-Term Liabilities</td>
<td>0.83</td>
<td>1.13</td>
<td>3.73</td>
<td>4.76</td>
<td>6.80</td>
<td>15.71</td>
<td>3.89</td>
<td>2.19</td>
</tr>
<tr>
<td>Equity to Net Property, Plant, and Equipment</td>
<td>1.36</td>
<td>1.74</td>
<td>3.03</td>
<td>3.65</td>
<td>4.39</td>
<td>6.00</td>
<td>2.10</td>
<td>1.69</td>
</tr>
<tr>
<td>Times Interest Earned</td>
<td>6.67</td>
<td>9.41</td>
<td>21.75</td>
<td>363.49</td>
<td>60.21</td>
<td>N/A</td>
<td>3.10</td>
<td>2.40</td>
</tr>
<tr>
<td>Earnings Coverage of Fixed Charges</td>
<td>2.92</td>
<td>3.37</td>
<td>4.32</td>
<td>4.79</td>
<td>5.02</td>
<td>9.40</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>Cash Flow Coverage of Interest Expenses</td>
<td>2.22</td>
<td>4.67</td>
<td>7.55</td>
<td>62.24</td>
<td>15.38</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>CANACOM CORPORATION</td>
<td>Asset Utilization and Profitability Ratios</td>
<td>December 31, 1990-1994</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------------------------------------------</td>
<td>---------------------------------------------------------</td>
<td>------------------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Hightech Computer Composite</td>
<td>Manufacturing Composite</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Retailing Composite</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sales to Cash &amp; Equivalents</td>
<td>32.30</td>
<td>1994 6.90</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>24.60</td>
<td>1994 3.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>11.90</td>
<td>1994 4.70</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>12.10</td>
<td>1994 2.90</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>8.80</td>
<td>1994 13.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>6.90</td>
<td>1994 34.80</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sales to Receivables</td>
<td>76.70</td>
<td>1994 7.20</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>53.80</td>
<td>1994 5.60</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>40.20</td>
<td>1994 33.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>24.30</td>
<td>1994 6.90</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>23.00</td>
<td>1994 4.70</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>7.20</td>
<td>1994 5.70</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sales to Inventories</td>
<td>3.20</td>
<td>1994 2.90</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>3.20</td>
<td>1994 4.50</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>3.30</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>3.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sales to Working Capital</td>
<td>3.90</td>
<td>1994 2.90</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>3.80</td>
<td>1994 4.50</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>3.30</td>
<td>1994 5.70</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sales to Net Property, Plant, &amp; Equipment</td>
<td>7.80</td>
<td>1994 8.20</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>8.40</td>
<td>1994 27.10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>8.90</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>9.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>9.60</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>14.70</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asset Turnover</td>
<td>1.99</td>
<td>1994 1.50</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.95</td>
<td>1994 2.80</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.81</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.66</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.56</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.77</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Return on Sales (%)</td>
<td>6.90</td>
<td>1994 4.40</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>8.10</td>
<td>1994 1.30</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>10.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>11.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>11.30</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Return on Assets (%)</td>
<td>13.70</td>
<td>1994 6.60</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>15.80</td>
<td>1994 3.60</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>18.10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>18.30</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>17.60</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Financial Leverage</td>
<td>2.93</td>
<td>1994 N/A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2.51</td>
<td>1994 N/A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.64</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.51</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.41</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.47</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Return on Equity (%)</td>
<td>40.00</td>
<td>1994 N/A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>39.60</td>
<td>1994 N/A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>29.70</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>27.60</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>24.90</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>20.30</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Return on Long-Term Liabilities &amp; Equity (%)</td>
<td>16.20</td>
<td>1994 N/A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>17.90</td>
<td>1994 N/A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>19.40</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>19.20</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>18.40</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>13.80</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Before Tax Return on Total Assets (%)</td>
<td>26.50</td>
<td>1994 8.60</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>29.70</td>
<td>1994 4.10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>34.20</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>34.50</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>33.30</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>26.30</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>8.60</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>4.10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>--------------------------------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td></td>
</tr>
<tr>
<td>Price Earnings Ratio</td>
<td>6.28</td>
<td>9.27</td>
<td>18.18</td>
<td>12.67</td>
<td>18.73</td>
<td>27.01</td>
<td>13.55</td>
<td></td>
</tr>
<tr>
<td>Earnings Price Ratio</td>
<td>0.16</td>
<td>0.11</td>
<td>0.06</td>
<td>0.08</td>
<td>0.05</td>
<td>0.04</td>
<td>0.07</td>
<td></td>
</tr>
<tr>
<td>Price to Cash from Operations</td>
<td>5.44</td>
<td>8.22</td>
<td>15.35</td>
<td>10.90</td>
<td>16.28</td>
<td>20.90</td>
<td>N/A</td>
<td></td>
</tr>
<tr>
<td>Price to Book Value of Equity</td>
<td>2.72</td>
<td>3.80</td>
<td>5.38</td>
<td>3.50</td>
<td>4.65</td>
<td>5.48</td>
<td>N/A</td>
<td></td>
</tr>
<tr>
<td>Dividend Yield (%)</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.15</td>
<td></td>
</tr>
<tr>
<td>Dividend Payout Ratio (%)</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.1</td>
<td>N/A</td>
<td></td>
</tr>
<tr>
<td>Internal Growth Rate (%)</td>
<td>40.00</td>
<td>39.60</td>
<td>29.70</td>
<td>27.60</td>
<td>24.90</td>
<td>20.30</td>
<td>N/A</td>
<td></td>
</tr>
<tr>
<td>---------------------</td>
<td>----------------------------------------------------------</td>
<td>-----------------------------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Cash Flows from Operations</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Net Income</strong></td>
<td>$83,229</td>
<td>$112,235</td>
<td>$169,602</td>
<td>$224,085</td>
<td>$278,521</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Add (deduct) items not affecting cash</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Depreciation Expense</td>
<td>$17,121</td>
<td>$19,110</td>
<td>$23,228</td>
<td>$29,437</td>
<td>$38,679</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Increase in Accounts Receivable</td>
<td>$(5,206)</td>
<td>$(9,884)</td>
<td>$(16,363)</td>
<td>$(41,528)</td>
<td>$(23,914)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Increase in Inventories</td>
<td>$(48,584)</td>
<td>$(53,511)</td>
<td>$(78,549)</td>
<td>$(156,859)</td>
<td>$(173,529)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Change in Accounts Payable</td>
<td>$(15,543)</td>
<td>$24,626</td>
<td>$(4,366)</td>
<td>$9,081</td>
<td>$999</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>$3,772</td>
<td>$(536)</td>
<td>$7,637</td>
<td>$7,318</td>
<td>$3,272</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Total of items not affecting cash</strong></td>
<td>$(48,440)</td>
<td>$(20,195)</td>
<td>$(68,413)</td>
<td>$(152,551)</td>
<td>$(154,493)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Net Cash Flow from Operating Activities</strong></td>
<td>$34,789</td>
<td>$92,040</td>
<td>$101,189</td>
<td>$71,534</td>
<td>$124,028</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Cash Flows used by Investing Activities</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Net Purchases of Land and Equipment</strong></td>
<td>$(26,579)</td>
<td>$(31,063)</td>
<td>$(48,494)</td>
<td>$(67,678)</td>
<td>$(72,675)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Cash Flows from Financing Activities</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Net Reductions in Long-Term Debt</td>
<td>$(106,669)</td>
<td>$(2,165)</td>
<td>$(4,022)</td>
<td>$16,739</td>
<td>$(5,223)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Purchase of Treasury Stock</td>
<td>$(27,396)</td>
<td>$(53,342)</td>
<td>$0</td>
<td>$0</td>
<td>$0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sale of Treasury Stock to Employees</td>
<td>$12,954</td>
<td>$15,833</td>
<td>$21,077</td>
<td>$29,048</td>
<td>$33,654</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Issue of Debentures</td>
<td>$98,875</td>
<td>$0</td>
<td>$0</td>
<td>$0</td>
<td>$0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Foreign Currency Adjustments</td>
<td>$0</td>
<td>$0</td>
<td>$0</td>
<td>$(10,688)</td>
<td>$(6,928)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>$7,032</td>
<td>$(7,578)</td>
<td>$(26,003)</td>
<td>$(17,825)</td>
<td>$(8,027)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Cash Provided by Financing Activities</strong></td>
<td>$(15,204)</td>
<td>$(47,252)</td>
<td>$(8,948)</td>
<td>$17,274</td>
<td>$13,476</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Cash at the Beginning of the Year</strong></td>
<td>$35,778</td>
<td>$28,784</td>
<td>$42,509</td>
<td>$86,256</td>
<td>$107,386</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Change in Cash During the Year</strong></td>
<td>$(6,994)</td>
<td>$13,725</td>
<td>$43,747</td>
<td>$21,130</td>
<td>$64,829</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Cash at the End of the Year</strong></td>
<td>$28,784</td>
<td>$42,509</td>
<td>$86,256</td>
<td>$107,386</td>
<td>$172,215</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Funds Reinvestment Ratio (%)</td>
<td>7.38</td>
<td>16.99</td>
<td>13.81</td>
<td>7.19</td>
<td>9.57</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Funds Adequacy Ratio (5 Year: 1990-1994)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.56</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Please indicate a reference code for yourself: __________

JUDGEMENT RECORDING SHEETS (Set 1)
Please answer the following questions:

Question 1

Based on your analysis and under current economic and interest-rate conditions, rate Canacom's current liquidity position. Please circle the correct answer.

Very Weak Position: 1 - 2 - 3 - 4 - 5 - 6 - 7 - 8 - 9 - 10: Very Strong Position

How confident are you that this judgment is correct? Provide a number between 0% and 100%, with 0% meaning you are completely unsure and 100% meaning you are completely confident.

___%  

Question 2

Based on your analysis and under current economic and interest-rate conditions, rate Canacom's long-term solvency position. Please circle the correct answer.

Very Weak Position: 1 - 2 - 3 - 4 - 5 - 6 - 7 - 8 - 9 - 10: Very Strong Position

How confident are you that this judgment is correct? Provide a number between 0% and 100%, with 0% meaning you are completely unsure and 100% meaning you are completely confident.

___%
Question 3

Based on your analysis and under current economic and interest-rate conditions, rate Canacom's asset utilization performance. Please circle the correct answer.

Very Weak Performance: 1 - 2 - 3 - 4 - 5 - 6 - 7 - 8 - 9 - 10 : Very Strong Performance

How confident are you that this judgment is correct? Provide a number between 0% and 100%, with 0% meaning you are completely unsure and 100% meaning you are completely confident.

_%

Question 4

Based on your analysis and under current economic and interest-rate conditions, rate the value of Canacom stock as loan collateral. Please circle the correct answer.

Very Low Value: 1 - 2 - 3 - 4 - 5 - 6 - 7 - 8 - 9 - 10 : Very High Value

How confident are you that this judgment is correct? Provide a number between 0% and 100%, with 0% meaning you are completely unsure and 100% meaning you are completely confident.

_%
Question 5

Based on your analysis and under current economic and interest-rate conditions, rate the quality of Canacom's financial management. Please circle the correct answer.

Very Poor Quality: 1 - 2 - 3 - 4 - 5 - 6 - 7 - 8 - 9 - 10: Excellent Quality

How confident are you that this judgment is correct? Provide a number between 0% and 100%, with 0% meaning you are completely unsure and 100% meaning you are completely confident.

___ %

Question 6

Based on your analysis and under current economic and interest-rate conditions, rate the quality of Canacom's operating management. Please circle the correct answer.

Very Poor Quality: 1 - 2 - 3 - 4 - 5 - 6 - 7 - 8 - 9 - 10: Excellent Quality

How confident are you that this judgment is correct? Provide a number between 0% and 100%, with 0% meaning you are completely unsure and 100% meaning you are completely confident.

___ %
**Question 7**

Based on your analysis and under current economic and interest-rate conditions, what is your estimate of Canacom's expected net income (after-tax) in the coming year?

My estimate of Canacom's expected net income is

$\_

million.

How confident are you that this judgment is correct? Provide a number between 0% and 100%, with 0% meaning you are completely unsure and 100% meaning you are completely confident.

___

**Question 8**

Based on your analysis and under current economic and interest-rate conditions, how much of the $800 million loan being requested would you recommend as being allowable to Canacom for the purposes of streamlining operations, assuming that it is unsecured?

I estimate that $\_

million should be allowable to Canacom

for the purposes of streamlining operations.

How confident are you that this judgment is correct? Provide a number between 0% and 100%, with 0% meaning you are completely unsure and 100% meaning you are completely confident.

___
GROUP EVALUATION OF CANACOM CORPORATION LOAN ANALYSIS

Assume that your superior has acquired a financial analysis expert system, called FINALYZER (short for FINancial AnaLYZER), and has requested you and two of your colleagues to JOINTLY evaluate the financial statements of Canacom Corporation by utilizing the system provided. The assessment made jointly by you and your colleagues will form the basis of a more complete loan evaluation to be undertaken by the Corporate Loans Committee of which your superior is a member. The analysis to be performed by your group has exactly the same requirements as that performed by you earlier. That is, you are specifically instructed to focus on all aspects of Canacom's valuation, liquidity, long-term solvency, asset utilization, and profitability. As a group, you will also provide the same set of specific judgments. Please perform as thorough and complete an analysis as possible.

The accuracy of your group judgements (questions 1-8) will be evaluated by a group of experts. Groups with scores placing in the top 20 percent of the total number of groups participating under similar conditions will be awarded a prize of $90 per group or $30 per individual in the group. Note that the prizes will be awarded strictly based on the quality of your group judgments and not the amount of time you spend. You have approximately two hours to complete the task.

When working in your group, you are required to employ the CONSENSUS rule. Consensus decision making requires that all members be involved in making the decision and accept the final decision. It need not be an unanimous agreement, but it is one that all members accept as probably the best decision they could reach under consensus decision making. The following guidelines will provide you with ground rules for conducting decision making by consensus:
1. Avoid arguing for your own position. Present your views and reasons as logically as possible. Listen to the other members' reactions and consider them carefully before pressing your point. Approach the decision on the basis of logic.

2. Do not assume that someone must win or lose. Look for the next most acceptable option for ALL members.

3. Do not change your mind simply to avoid conflict or to reach agreement and harmony. When agreement seems to come too quickly and easily, be suspicious. Explore the reasons and be sure everyone accepts the solution for basically similar or complementary reasons. Yield only to positions that have objective and logically sound foundations.

4. Avoid conflict-reducing techniques such as flipping a coin, voting, compromising, and averaging. These techniques require someone to lose and do not ensure reasoned decisions. Instead, strive to resolve any conflict that arises.

5. View differences of opinions as helpful rather than a hindrance in decision making. Differences of opinion are natural and the most valuable resource for critical decision making. Seek out the differences and try to involve everyone in the decision process. Disagreements can help the group's decision because the wider the range of information and opinions considered, the greater the chance of high-quality decisions.

6. If open discussion seems to leave out certain members, be sure to ask their opinion. You may wish to use the round-robin to be sure all members have a final say if you doubt all views have been aired.

Before you proceed to use the FINALYZER expert system, you will be introduced to the various features provided by the system.
A SHORT NOTE ON EXPERT SYSTEMS

Expert systems are computer systems that make the specialized knowledge and expertise of a particular domain available to decision makers. They are generally developed by closely modelling the expertise and mode of operation used by human experts.

Similar to human experts interacting with their clients, expert systems attempt to provide users with relevant information about the inputs they use in their analysis, as well as the conclusions that they reach. For example, human medical experts commonly inform patients of the inputs they are using for their diagnosis (e.g., why they are requiring a blood test or blood pressure reading) as well as their diagnostic conclusions. They provide, usually at the patients' request, explanations relating to the inputs used for the diagnosis or the conclusions reached. Similarly, expert systems go beyond the capability of conventional computer systems to provide users with various types of explanations.

More specifically, the expert system you will be using will provide you with the following information: 1) the ratios and other inputs that it will use for each analysis, 2) the specific conclusions arising from each analysis that it has performed, and 3) explanations relating to both (1) and (2). For each input used or conclusion presented, three different types of explanations will be provided. These are the WHY, HOW and STRATEGIC explanations:

**WHY explanations** either justify *why a particular input or ratio is needed* for an analysis, or rationalize *why a particular conclusion that has been reached is important* for the task.

**HOW explanations** either detail *how a particular input or ratio is defined and computed*, or reveal *how a particular conclusion has been reached* by presenting a trace of the evaluations.

**STRATEGIC explanations** either provide the *overall structure* in which all the relevant input information is organized, or the *overall problem solving strategies* in which a particular conclusion fits.
CREDIT-ADVISOR SYSTEM TUTORIAL

The CREDIT-ADVISOR tutorial system has features that are identical to FINALYZER which you will use later. The objective of using this tutorial system is to familiarize you with the types and sequences of the screens that you will encounter in the use of FINALYZER, and the way you interact with FINALYZER. Now, you should focus on the system features of CREDIT-ADVISOR, rather than its contents. It is important to try ALL the features of the tutorial system, to know what to expect when clicking on any buttons, and to become comfortable with using the system. Note, your use of CREDIT-ADVISOR is NOT part of the evaluation of your performance.

The three types of screens that you will see are DATA SCREENs, INFORMATION SCREENs, and RECOMMENDATION SCREENs. On ALL of the three types of screens, there are WHY, HOW, and STRATEGIC explanations for each data item, analysis, or conclusion, respectively (see previous page for definitions of these explanations). The system structure is illustrated on the page after next.

By the end of this tutorial, you should know exactly what to expect on the DATA SCREENs, INFORMATION SCREENs, and RECOMMENDATION SCREENs, and know what will happen when you click on the NEXT-SCREEN, PREVIOUS-SCREEN, BACK-TO-ANALYSIS, WHY, HOW, and STRATEGIC buttons.

Assume that you are a credit officer at a local bank responsible for the evaluation of credit-card applications. The bank has received the application of Mr. Robert Mortenstein, and the relevant data from his application (given on next page) has been input into a computerised data file. You now wish to use the CREDIT-ADVISOR expert system to evaluate the merit of the application.

Please practice using CREDIT-ADVISOR in your group. You will interact with CREDIT-ADVISOR by requesting the facilitator to carry out the operations.
**CREDIT INFORMATION**

<table>
<thead>
<tr>
<th>Name:</th>
<th>Robert Mortenstein</th>
</tr>
</thead>
<tbody>
<tr>
<td>Income (gross):</td>
<td>$23,000 /yr</td>
</tr>
<tr>
<td>Occupation:</td>
<td>Supervisor (City of Vancouver)</td>
</tr>
</tbody>
</table>

**Assets**

<table>
<thead>
<tr>
<th>Item</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Car (Sundance 1989)</td>
<td>$7,000</td>
</tr>
<tr>
<td>RRSP Deposits</td>
<td>$3,000</td>
</tr>
<tr>
<td>Bank Accounts</td>
<td>$2,300</td>
</tr>
</tbody>
</table>

**Liabilities**

<table>
<thead>
<tr>
<th>Item</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Car Loan Remaining</td>
<td>$5,600</td>
</tr>
<tr>
<td>Credit-Card Balance Outstanding</td>
<td>$1,000</td>
</tr>
</tbody>
</table>

**Monthly Payments:**

<table>
<thead>
<tr>
<th>Item</th>
<th>Amount</th>
</tr>
</thead>
<tbody>
<tr>
<td>Car Loan</td>
<td>$400</td>
</tr>
<tr>
<td>Housing (Rental)</td>
<td>$700</td>
</tr>
<tr>
<td>Credit-Card Payments</td>
<td>$150</td>
</tr>
<tr>
<td>Total</td>
<td>$1,250</td>
</tr>
</tbody>
</table>

**Credit Status**

- VISA: $800
- Canadian Tire: $300

**Age:** 30 Years

**Marital Status:** Married / No Children

**Other Info:**
- Chequing Account set up in 1988
- Living At Present Address since 1992
- Home Ownership: Renting
- At present job since May 1993
- Nature of job is full-time and continuing till April 1996

**Credit History:** R1 Status
The System Flow Chart of CREDIT-ADVISOR
USING FINALYZER TO SUPPORT YOUR GROUP DECISION MAKING PROCESS

The FINALYZER expert system was developed by modelling the expertise and knowledge of several experts in the field. It conducts five specific analyses, and for each analysis informs you of the types of ratios to be used as part of that analysis, on an INFORMATION SCREEN. It also computes and presents tables of ratios identical to the ones you were provided with earlier, on a DATA SCREEN. After applying its expertise to these ratios, it generates a set of recommendations relevant to the particular analysis, on a RECOMMENDATION SCREEN. The system structure is shown on next page.

The interface and format of FINALYZER are identical to that of the CREDIT-ADVISOR expert system that you have used earlier. It also makes available to you the WHY, HOW and STRATEGIC explanations on various types of screens.

You are now requested to use the FINALYZER expert system to support your group evaluation of the CANACOM CORPORATION LOAN ANALYSIS. Please apply the CONSENSUS DECISION MAKING RULE in your group discussion. PLEASE GO THROUGH ALL THE ANALYSES IN FINALYZER and INDICATE YOUR GROUP EVALUATION ON THE GROUP JUDGMENT RECORDING SHEETS.
INTRODUCTION

FINANCIAL STATEMENT PRESENTATION SCREENS

ANALYSIS SELECTION SCREEN

OVERALL SUMMARY SCREEN

EXIT SCREEN

BALANCE SHEET ANALYSIS

INCOME STATEMENT ANALYSIS

FUNDS FLOW ANALYSIS

LIQUIDITY ANALYSIS

CAPITAL STRUCTURE ANALYSIS

PROFITABILITY ANALYSIS

MARKET VALUE ANALYSIS

The System Flow Chart of FINALYZER
GROUP JUDGEMENT RECORDING SHEETS

Please answer the following questions:

**Question 1**

Based on your analysis and under current economic and interest-rate conditions, rate Canacom's current liquidity position. Please circle the correct answer.

Very Weak Position: 1 - 2 - 3 - 4 - 5 - 6 - 7 - 8 - 9 - 10 : Very Strong Position

How confident are you that this judgment is correct? Provide a number between 0% and 100%, with 0% meaning you are completely unsure and 100% meaning you are completely confident.

___ %

**Question 2**

Based on your analysis and under current economic and interest-rate conditions, rate Canacom's long-term solvency position. Please circle the correct answer.

Very Weak Position: 1 - 2 - 3 - 4 - 5 - 6 - 7 - 8 - 9 - 10 : Very Strong Position

How confident are you that this judgment is correct? Provide a number between 0% and 100%, with 0% meaning you are completely unsure and 100% meaning you are completely confident.

___ %
Question 3

Based on your analysis and under current economic and interest-rate conditions, rate Canacom's asset utilization performance. Please circle the correct answer.

Very Weak Performance: 1 - 2 - 3 - 4 - 5 - 6 - 7 - 8 - 9 - 10 : Very Strong Performance

How confident are you that this judgment is correct? Provide a number between 0% and 100%, with 0% meaning you are completely unsure and 100% meaning you are completely confident.

% 

Question 4

Based on your analysis and under current economic and interest-rate conditions, rate the value of Canacom stock as loan collateral. Please circle the correct answer.

Very Low Value: 1 - 2 - 3 - 4 - 5 - 6 - 7 - 8 - 9 - 10 : Very High Value

How confident are you that this judgment is correct? Provide a number between 0% and 100%, with 0% meaning you are completely unsure and 100% meaning you are completely confident.

% 

308
**Question 5**

Based on your analysis and under current economic and interest-rate conditions, rate the quality of Canacom's financial management. Please circle the correct answer.

Very Poor Quality: 1 - 2 - 3 - 4 - 5 - 6 - 7 - 8 - 9 - 10 : Excellent Quality

How confident are you that this judgment is correct? Provide a number between 0% and 100%, with 0% meaning you are completely unsure and 100% meaning you are completely confident.

___ %

---

**Question 6**

Based on your analysis and under current economic and interest-rate conditions, rate the quality of Canacom's operating management. Please circle the correct answer.

Very Poor Quality: 1 - 2 - 3 - 4 - 5 - 6 - 7 - 8 - 9 - 10 : Excellent Quality

How confident are you that this judgment is correct? Provide a number between 0% and 100%, with 0% meaning you are completely unsure and 100% meaning you are completely confident.

___ %
**Question 7**

Based on your analysis and under current economic and interest-rate conditions, what is your estimate of Canacom's expected net income (after tax) in the coming year?

My estimate of Canacom's expected net income is $\_\_\_\_\_ million.

How confident are you that this judgment is correct? Provide a number between 0% and 100%, with 0% meaning you are completely unsure and 100% meaning you are completely confident. 

\_\_\_\_\_%

**Question 8**

Based on your analysis and under current economic and interest-rate conditions, how much of the $800 million loan being requested would you recommend as being allowable to Canacom for the purposes of streamlining operations, assuming that it is unsecured?

I estimate that $\_\_\_\_\_ million should be allowable to Canacom for the purposes of streamlining operations.

How confident are you that this judgment is correct? Provide a number between 0% and 100%, with 0% meaning you are completely unsure and 100% meaning you are completely confident.

\_\_\_\_\_%
FINAL EVALUATION OF CANACOM CORPORATION LOAN ANALYSIS

Assume that your supervisor has requested you to INDIVIDUALLY perform a final evaluation of the CANACOM CORPORATION LOAN ANALYSIS based on the amount of information and knowledge you have gathered from your group discussion and the FINALYZER system. This activity is to be carried out because of the importance of this transaction which involves a very large loan amount. Any error in judgments could have significant consequences -- in cases where the amount of loan granted is too high or too low compared to the ideal. Due to its importance, your supervisor has requested that you PERSONALLY take a final look at the case and work out what you believe are the best judgments. Judgments for all of questions 1 through 8 are of equal importance as the Corporate Loans Committee feels that the rationale for determining the amount of loan to be granted (as indicated by your answers to questions 1 through 7) is just as important as the amount itself (as indicated by question 8). Furthermore, you have been given indications that the Corporate Loans Committee is likely to seek clarifications from you regarding your final individual judgments. As such, it is important for you to perform well in this final evaluation.

Please proceed with the final evaluation.
JUDGEMENT RECORDING SHEETS (Set 2)
Please answer the following questions:

**Question 1**

Based on your analysis and under current economic and interest-rate conditions, rate Canacom's **current liquidity position**. Please circle the correct answer.

Very Weak Position: 1 - 2 - 3 - 4 - 5 - 6 - 7 - 8 - 9 - 10 : Very Strong Position

How confident are you that this judgment is correct? Provide a number between 0% and 100%, with 0% meaning you are **completely unsure** and 100% meaning you are **completely confident**.

___%

**Question 2**

Based on your analysis and under current economic and interest-rate conditions, rate Canacom's **long-term solvency position**. Please circle the correct answer.

Very Weak Position: 1 - 2 - 3 - 4 - 5 - 6 - 7 - 8 - 9 - 10 : Very Strong Position

How confident are you that this judgment is correct? Provide a number between 0% and 100%, with 0% meaning you are **completely unsure** and 100% meaning you are **completely confident**.

___%
**Question 3**

Based on your analysis and under current economic and interest-rate conditions, rate Canacom's asset utilization performance. Please circle the correct answer.

Very Weak Performance: 1 - 2 - 3 - 4 - 5 - 6 - 7 - 8 - 9 - 10 : Very Strong Performance

How confident are you that this judgment is correct? Provide a number between 0% and 100%, with 0% meaning you are completely unsure and 100% meaning you are completely confident.

___ %

**Question 4**

Based on your analysis and under current economic and interest-rate conditions, rate the value of Canacom stock as loan collateral. Please circle the correct answer.

Very Low Value: 1 - 2 - 3 - 4 - 5 - 6 - 7 - 8 - 9 - 10 : Very High Value

How confident are you that this judgment is correct? Provide a number between 0% and 100%, with 0% meaning you are completely unsure and 100% meaning you are completely confident.

___ %
**Question 5**

Based on your analysis and under current economic and interest-rate conditions, rate the quality of Canacom's financial management. Please circle the correct answer.

Very Poor Quality: 1 - 2 - 3 - 4 - 5 - 6 - 7 - 8 - 9 - 10 : Excellent Quality

How confident are you that this judgment is correct? Provide a number between 0% and 100%, with 0% meaning you are completely unsure and 100% meaning you are completely confident.

__%__

**Question 6**

Based on your analysis and under current economic and interest-rate conditions, rate the quality of Canacom's operating management. Please circle the correct answer.

Very Poor Quality: 1 - 2 - 3 - 4 - 5 - 6 - 7 - 8 - 9 - 10 : Excellent Quality

How confident are you that this judgment is correct? Provide a number between 0% and 100%, with 0% meaning you are completely unsure and 100% meaning you are completely confident.

__%__
Question 7

Based on your analysis and under current economic and interest-rate conditions, what is your estimate of Canacom's expected net income (after-tax) in the coming year?

My estimate of Canacom's expected net income is

$____ million.

How confident are you that this judgment is correct? Provide a number between 0% and 100%, with 0% meaning you are completely unsure and 100% meaning you are completely confident.

____%  

Question 8

Based on your analysis and under current economic and interest-rate conditions, how much of the $800 million loan being requested would you recommend as being allowable to Canacom for the purposes of streamlining operations, assuming that it is unsecured?

I estimate that $____ million should be allowable to Canacom

for the purposes of streamlining operations.

How confident are you that this judgment is correct? Provide a number between 0% and 100%, with 0% meaning you are completely unsure and 100% meaning you are completely confident.

____%
POST STUDY QUESTIONNAIRE

INSTRUCTIONS

To what extent do the following statements reflect your views of the FINALYZER system in supporting the group decisions you have made? For each statement CIRCLE the one number that best represents your agreement or disagreement with the statement (1 - Strongly Disagree, 2 - Disagree, 3 - Slightly Disagree, 4 - Neutral, 5 - Slightly Agree, 6 - Agree, 7 - Strongly Agree).

PLEASE ANSWER ALL THE FOLLOWING ITEMS.

1. The use of FINALYZER greatly enhanced the quality of my group's judgements.
   Strongly Disagree: 1 - 2 - 3 - 4 - 5 - 6 - 7 : Strongly Agree

2. Using FINALYZER gave my group more control over the financial analysis task.
   Strongly Disagree: 1 - 2 - 3 - 4 - 5 - 6 - 7 : Strongly Agree

3. FINALYZER provided good advice across different situations.
   Strongly Disagree: 1 - 2 - 3 - 4 - 5 - 6 - 7 : Strongly Agree

4. FINALYZER is dependable in important decisions.
   Strongly Disagree: 1 - 2 - 3 - 4 - 5 - 6 - 7 : Strongly Agree

5. When FINALYZER gave unexpected advice, my group is confident that the advice is correct.
   Strongly Disagree: 1 - 2 - 3 - 4 - 5 - 6 - 7 : Strongly Agree
6. Using FINALYZER made the financial analysis task carried out by my group easier to perform.

   Strongly Disagree: 1 - 2 - 3 - 4 - 5 - 6 - 7 : Strongly Agree

7. Using FINALYZER enabled my group to accomplish the financial analysis task more quickly.

   Strongly Disagree: 1 - 2 - 3 - 4 - 5 - 6 - 7 : Strongly Agree

8. FINALYZER is a reliable source of knowledge for financial analysis.

   Strongly Disagree: 1 - 2 - 3 - 4 - 5 - 6 - 7 : Strongly Agree

9. Using FINALYZER improved the quality of the analysis my group performed.

   Strongly Disagree: 1 - 2 - 3 - 4 - 5 - 6 - 7 : Strongly Agree

10. FINALYZER supported all types of analysis needed by my group to make its decisions.

    Strongly Disagree: 1 - 2 - 3 - 4 - 5 - 6 - 7 : Strongly Agree

11. Using FINALYZER increased my group's productivity.

    Strongly Disagree: 1 - 2 - 3 - 4 - 5 - 6 - 7 : Strongly Agree

12. I think users with little expertise would trust the advice given by FINALYZER.

    Strongly Disagree: 1 - 2 - 3 - 4 - 5 - 6 - 7 : Strongly Agree

13. Overall, I found FINALYZER useful in helping my group analyze the financial statements.

    Strongly Disagree: 1 - 2 - 3 - 4 - 5 - 6 - 7 : Strongly Agree
14. Using FINALYZER enhanced my group's effectiveness in completing the financial analysis task.

   Strongly Disagree: 1 - 2 - 3 - 4 - 5 - 6 - 7 : Strongly Agree

15. FINALYZER gave the same advice for the same situation over time.

   Strongly Disagree: 1 - 2 - 3 - 4 - 5 - 6 - 7 : Strongly Agree

16. Using FINALYZER allowed my group to accomplish more analysis than would otherwise have been possible.

   Strongly Disagree: 1 - 2 - 3 - 4 - 5 - 6 - 7 : Strongly Agree

17. FINALYZER behaved in a very consistent manner.

   Strongly Disagree: 1 - 2 - 3 - 4 - 5 - 6 - 7 : Strongly Agree

18. FINALYZER helped my group make good decisions.

   Strongly Disagree: 1 - 2 - 3 - 4 - 5 - 6 - 7 : Strongly Agree
INSTRUCTIONS

To what extent do the following statements reflect your views of the group decision making process you have just experienced and the decisions reached by your group? For each statement CIRCLE the one number (1 - 7) that best represents your view with the statement. PLEASE ANSWER ALL THE FOLLOWING ITEMS.

1. How would you describe your group's problem solving process?

   Efficient: 1 - 2 - 3 - 4 - 5 - 6 - 7: Inefficient

   Coordinated: 1 - 2 - 3 - 4 - 5 - 6 - 7: Uncoordinated

   Fair: 1 - 2 - 3 - 4 - 5 - 6 - 7: Unfair

   Confusing: 1 - 2 - 3 - 4 - 5 - 6 - 7: Understandable

   Satisfying: 1 - 2 - 3 - 4 - 5 - 6 - 7: Dissatisfying

2. How different or similar are your final individual decisions from your group's decisions?

   Very different: 1 - 2 - 3 - 4 - 5 - 6 - 7: Very similar

3. To what extent do you feel personally responsible for the correctness of your group's solution?

   Not at all: 1 - 2 - 3 - 4 - 5 - 6 - 7: To a very great extent

4. Do you disagree or agree with your group's solution?

   Strongly disagree: 1 - 2 - 3 - 4 - 5 - 6 - 7: Strongly agree
5. To what extent do you feel committed to your group's solution?

Not at all: 1 - 2 - 3 - 4 - 5 - 6 - 7 : To a very great extent

6. How satisfied or dissatisfied are you with the quality of your group's solution?

Very dissatisfied: 1 - 2 - 3 - 4 - 5 - 6 - 7 : Very satisfied

7. To what extent are you confident that your group's solution is correct?

Not at all: 1 - 2 - 3 - 4 - 5 - 6 - 7 : To a very great extent

8. To what extent does the group solution reflect your contributions?

Not at all: 1 - 2 - 3 - 4 - 5 - 6 - 7 : To a very great extent

9. To what extent do you support your group's solution?

Strongly oppose: 1 - 2 - 3 - 4 - 5 - 6 - 7 : Strongly support
ESSAY QUESTIONS:

In the space provided, please address the following issues as completely and specifically as possible.

1. Do you think the FINALYZER system helped your group resolve conflicts in any way?

2. Do you think the FINALYZER system helped your group in reaching consensus?

3. What are the major strengths of the FINALYZER system?

4. What are the major weaknesses of the FINALYZER system?

5. More specifically, what changes would you like to see in FINALYZER?

Thank you very much for your participation. Your contribution to this research is very much appreciated. If you have additional comments to add concerning this study, please use the space at the back of this page.
DEBRIEFING PROTOCOL

Thank you again for your participation. It is our hope that you have enjoyed working with FINALYZER in completing the case. Your judgments and evaluations about Canacom and the system, the data captured by the system of your interaction with it, and the videotape capture of your group interaction process will be analyzed using a variety of computer and statistical analyses. We hope to learn more about the optimal manner in which expert systems for financial analysis should be designed to maximize their potential to users such as yourself. Studies like this, that focus on users' behaviour in using computer support tools, provide valuable feedback that helps us design more user friendly and easy-to-learn expert systems.

The primary focus of this study is in examining whether and how expert systems can support group decision making. We are particularly interested in how you have interacted with the system and your group members. Our goal is to learn more about how expert systems should provide its features effectively.

The FINALYZER system that you have used is not a complete and fully functional financial analysis expert system yet. Developing a complete and fully functional system requires a vast amount of time and resources, and is a difficult task beyond the scope of the current study. However, this study contributes to that goal by tackling some aspects of the complete problem.

The above information was not revealed to you at the start of the session in the interest of what is termed experimental control and validity. For a study like this to be successful, it is of critical importance that you, as the user of the system, behave as objectively and as naturally as you would in a real situation requiring the use of such a system. Because it was felt that your knowing this information would distort or bias your behaviour in using FINALYZER, it was decided that this information would only be revealed to you at the end of your participation.

If you have any additional comments or questions, please approach the research assistant now.
"I hereby confirm that I have received $25.00 as honorarium for participating in the financial analysis group decision making study."

Name (in full): ____________________________

Social Insurance Number: _____________________

Signature: __________________

Date: __________________
Research assistant name:

Participant name:

Time of arrival:

Time to start problem solving (manually):

Time to start tutorials:

Time to start using FINALYZER (including judgement questionnaire):

Time to start final evaluation and post-study-questionnaire:

Time of finishing:

Special observations and comments:
APPENDIX B: SAMPLE SCREENS OF FINALYZER
Press the NEXT-SCREEN button to continue
Select Type of Financial Analysis

- Liquidity Analysis
- Capital Structure Analysis
- Asset Utilization and Profitability Analysis
- Funds Statement Analysis
- Market - Value Analysis
- Overall Summary and EXIT THE SYSTEM

Click on one of the above radio buttons in order to proceed to the next screen.
Liquidity Analysis

The following information will be used by the system as part of its analysis. Should you wish to receive explanations about any of these, click on the appropriate radio-button followed by a click on the relevant explanation button. At end, please click NEXT-SCREEN to proceed.

- **Current Ratio**
- **Days to Sell Inventory**
- **Acid-Test Ratio**
- **Conversion Period (days)**
- **Accounts Receivable Turnover**
- **% Cash to Current Assets**
- **Inventory Turnover**
- **% Cash to Current Liabilities**
- **Day Sales in Receivables**
- **Working Capital (million $)**
- **Oper. Cash Flow to Total Curr. Liab.**
- **Liquidity Index**

Explanations:  
[WHY] [HOW] [STRATEGIC]

Screen Control:  
[PREVIOUS-SCREEN] [NEXT-SCREEN]
The current ratio indicates the company's ability to meet short term obligations. A value that is greater than 1 is desirable, but high values could indicate excessive receivables, inventory, and/or cash.
Current Assets
Current Ratio = ____________

Current Liabilities

Note, the numerator of this ratio consists of inventories and accounts receivables, which are not as liquid as cash. Such a limitation in this measure of liquidity is partly overcome by the 'acid-test ratio' which excludes from the numerator inventories [the least liquid component of current assets] and prepaid expenses.

The 'percent cash to current liabilities' ratio further exclude from the numerator receivables [more liquid than inventories but still one step removed from cash].

Moreover, since it takes time to convert receivables and inventories into cash, when using the current ratio, one should take into account the nature and turnover rates of the assets and liabilities involved in the numerator and denominator. These turnover rates include 'days sales in receivables', 'days to sell inventory', and the amount of time allowed for the payment of current liabilities.
STRATEGIC EXPLANATION - (FFST6)

LIQUIDITY ANALYSIS

○ Liquidity Index

Current Condition Review

○ Working Capital

Review Composition

○ Current Ratio

C. Assets Items

C. Liabilities Items

○ Acid-Test Ratio

○ Cash to Current Assets

○ Cash to Current Liabilities

Operational Impact Review

Inventory

○ Inventory Turnover

○ Days to Sell Inventory

Cash Flow

○ Operating Cash Flow to Total Current Liabilities

Receivables

○ Accounts Receivable Turnover

○ Day Sales in Receivables

Explanations

WHY

HOW

Screen Control

PREVIOUS-SCREEN

BACK-TO-ANALYSIS
# TABLE 6

**CANACOM CORPORATION**  
*Short-Term Liquidity Ratios*  
**December 31, 1990-1994**

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Current Ratio</td>
<td>3.25</td>
<td>3.15</td>
<td>3.48</td>
<td>4.07</td>
<td>4.42</td>
<td>3.64</td>
<td>2.70</td>
<td>1.40</td>
</tr>
<tr>
<td>Acid-Test Ratio</td>
<td>0.39</td>
<td>0.49</td>
<td>0.90</td>
<td>1.08</td>
<td>1.35</td>
<td>2.17</td>
<td>1.10</td>
<td>0.30</td>
</tr>
<tr>
<td>Accounts Receivable Turnover</td>
<td>76.73</td>
<td>53.82</td>
<td>40.19</td>
<td>24.31</td>
<td>23.02</td>
<td>7.20</td>
<td>5.60</td>
<td>45.50</td>
</tr>
<tr>
<td>Inventory Turnover</td>
<td>1.40</td>
<td>1.37</td>
<td>1.37</td>
<td>1.23</td>
<td>1.19</td>
<td>3.55</td>
<td>3.10</td>
<td>3.60</td>
</tr>
<tr>
<td>Day Sales in Receivables</td>
<td>4.80</td>
<td>6.80</td>
<td>9.10</td>
<td>15.00</td>
<td>15.90</td>
<td>50.70</td>
<td>65.20</td>
<td>8.00</td>
</tr>
<tr>
<td>Days to Sell Inventory</td>
<td>260.70</td>
<td>266.40</td>
<td>266.40</td>
<td>296.70</td>
<td>306.70</td>
<td>102.80</td>
<td>117.70</td>
<td>101.40</td>
</tr>
<tr>
<td>Conversion Period (days)</td>
<td>265.50</td>
<td>273.20</td>
<td>275.50</td>
<td>311.70</td>
<td>322.60</td>
<td>153.50</td>
<td>182.90</td>
<td>109.40</td>
</tr>
<tr>
<td>%Cash to Current Assets</td>
<td>8.40</td>
<td>10.61</td>
<td>20.02</td>
<td>17.66</td>
<td>22.14</td>
<td>30.55</td>
<td>14.80</td>
<td>11.60</td>
</tr>
<tr>
<td>%Cash to Current Liabilities</td>
<td>27.29</td>
<td>33.41</td>
<td>69.68</td>
<td>71.95</td>
<td>97.78</td>
<td>1112.57</td>
<td>29.60</td>
<td>16.50</td>
</tr>
<tr>
<td>Working Capital (millions$)</td>
<td>309.85</td>
<td>362.34</td>
<td>505.43</td>
<td>715.86</td>
<td>977.2</td>
<td>340.21</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>Liquidiy Index</td>
<td>233.10</td>
<td>230.20</td>
<td>203.40</td>
<td>228.10</td>
<td>222.50</td>
<td>51.1</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>Operating Cash Flow to Total Current Liabilities</td>
<td>0.73</td>
<td>0.78</td>
<td>0.95</td>
<td>1.09</td>
<td>1.11</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
</tbody>
</table>
Short-Term Liquidity Conclusions

The following conclusions have been reached by the system. Should you wish to receive explanations about the conclusions or the terms, click on the appropriate radio-button that precedes each of them, followed by a click on the relevant explanation button. At end, please click NEXT-SCREEN to proceed.

- Canacom is in a very favourable 'working capital' position indicating little risk in the short term of financial disaster. It would be an optimal client for a short-term loan.

- There is an increasing trend toward having a high proportion of sales on credit. While the proportion has tripled in the last five years, it still remains within a range well below industry levels.
As a rule of thumb, the more liquid a company's assets in relation to its current liabilities, the better it is able to settle its debts in the near future and the lesser the short-term risk of bankruptcy. Canacom is in good standing in this regard.
The conclusion was reached based on the following evaluations:

1. the 'current ratio' and 'acid-test ratio' are both above 1,
2. both these values are superior to the competitor and industry indices,
3. there is a healthy increasing trend over the last five years in these ratios and in the amount of 'working capital', and
4. the 'cash to current liabilities' ratio being at 97.78 percent suggests that the enterprise has the potential to settle virtually all its current liabilities immediately in cash.
LIQUIDITY ASSESSMENT

- Assessment of Working Capital [Positive]
- Assessment of Receivables Management [Negative]
- Assessment of Inventory Management [Negative]
- Assessment of Cash & Equivalents Management [Positive]
The following conclusions have been reached by the system. Should you wish to receive explanations about the conclusions or the terms, click on the appropriate radio-button that precedes each of them, followed by a click on the relevant explanation button. At end, please click NEXT-SCREEN to proceed.

- Canacom is in a very favourable working capital position indicating little risk in the short term of financial disaster. It would be an optimal client for a short-term loan.

- There is an increasing trend toward having a high proportion of sales on credit. While the proportion has tripled in the last five years, it still remains within a range well below industry levels.
The amount of working capital indicates the ability of a company to meet short term obligations with assets that would normally be consumed during a single operating cycle. It measures the net investment in short term operating assets.

Higher numbers indicate greater liquidity, but accounts receivable and inventory, which are part of current assets, should be at reasonable levels.

The absolute amount of working capital has significance only when related to other variables such as sales, total assets, and so forth. Therefore, ratios such as 'sales to working capital' and/or working capital as a percentage of total assets could also be calculated.
Working Capital = Current Assets - Current Liabilities

Working capital is often referred to as net working capital, as it represents the excess of current assets over current liabilities.

Working capital is closely related to the 'current ratio', but it is an absolute number. When comparing companies of different sizes, absolute amounts can be misleading. Therefore, working capital is generally used in conjunction with the current ratio.
The following conclusions have been reached by the system. Should you wish to receive explanations about the conclusions or the terms, click on the appropriate radio-button that precedes each of them, followed by a click on the relevant explanation button. At end, please click NEXT-SCREEN to proceed.

- Inventory is very unfavourable in comparison with major competitors and industry composites. The decreasing trend also suggests the need to extend tighter control over inventory management.

- The cash and cash equivalents position of Canacom has been improving from a below industry-average level to a better than industry-average level. This has to be carefully monitored to assure that cash levels are not excessive.
Select Type of Financial Analysis

- Liquidity Analysis
- Capital Structure Analysis
- Asset Utilization and Profitability Analysis
- Market - Value Analysis
- Funds Statement Analysis

[Overall Summary and EXIT THE SYSTEM]

Click on one of the above radio buttons in order to proceed to the next screen.

[Next Screen]  [Previous Screen]
OVERALL SUMMARY

A. POSITIVE FACTORS
1. Liquidity position is strong and improving. Little risk of default.
2. Market value is high considering the fundamentals.
3. Capital structure is highly equity based with scope to take on debt.
4. Profitability is above-average in industry and remains strong.
5. Operating revenues have been increasing steadily in the last five years.
6. Cash inflows and outflows have been managed well.
7. Reinvestment of funds is adequate with no dividend payout.
8. Has successfully used a policy of accepting lower "asset turnover" for higher margins.
9. "Price to earning" value is strong and improving consistently.

B. NEGATIVE FACTORS
1. "Funds flow adequacy" to cover capital expenditures and inventory purchase is low.
2. Too much investment is tied up in cash and its equivalents.
3. Conservative capital management may be holding up strategic potential.
4. May not be reinvesting adequately to maintain competitive position.
5. Selling and administrative costs are increasing, suggesting increased competition.
6. Asset utilization and "return on equity" have been declining consistently.
7. Growth rate has peaked and stock price may decline in the medium term.
8. Inventory and receivables management has been weak and worsening.
9. Too little debt in the capital structure is adversely affecting return on equity.
APPENDIX C: MATERIALS FOR RECRUITMENT OF SUBJECTS
FINANCIAL ANALYSIS GROUP DECISION MAKING

Researchers: Dr. Izak Benbasat, Ms. Fui-Hoon Nah

Objective: To advance knowledge in financial analysis decision making by groups

Task: Solve a financial analysis case individually, then solve it in a group of 3

Date: January 1996

Time Taken: Between 2 to 3 hours

Reward: HONORARIUM OF $25 for participation;
ONE IN FIVE chance of winning a GROUP AWARD OF $90,
that is, $30 CASH AWARD for each individual in the top 20% performing groups.

Benefit: Exposure to a user-friendly FINANCIAL EXPERT DECISION SUPPORT SYSTEM

Participation is voluntary. Any information obtained during the study that can be identified with individuals will remain confidential. Only aggregate data will be presented in academic reports or publications to preserve the anonymity of participants.

We would appreciate it very much if you could participate. To participate, please fill in the attached form and leave it in the box provided in the class or email Ms. Fui-Hoon Nah at fhnah@unixg.ubc.ca. The date for the experiment will be scheduled at your convenience. If you have further questions about the study, please feel free to contact Dr. Izak Benbasat at 822-8396 or Ms. Fui-Hoon Nah at telephone number 224-6264 or email fhnah@unixg.ubc.ca.

THANK YOU!
CONFIRMATION FORM

I would like to participate in the Financial Analysis Group Decision Making Study. Please call me to schedule my participation.

Name: ___________________________________________________________

Telephone Number: ________________________________________________

Preferred Dates of Participation: ____________________________________

We thank you very much for your interest in this study.

Please return this form to:

Ms. Fui Hoon Nah or Dr. Izak Benbasat
Office: 4A (directly facing office 456), 452
Financial Analysis Group Decision Making Project
Faculty of Commerce and Business Administration
University of British Columbia
2053 Main Mall, Vancouver, B.C., V6T 1Z2

PHONE (604) 224-6264, FAX (604) 822-9574