INFORMATION TO USERS

The most advanced technology has been used to photograph and reproduce this manuscript from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.
The impact of cognitive feedback on group decision-making

Sengupta, Kishore, Ph.D.
Case Western Reserve University, 1990

Copyright ©1989 by Sengupta, Kishore. All rights reserved.
THE IMPACT OF COGNITIVE FEEDBACK
ON GROUP DECISION MAKING

by

KISHORE SENGUPTA

Submitted in partial fulfillment of the requirements for the Degree of Doctor of Philosophy

Thesis Advisors: Michael J. Ginzberg and Dov Te’eni

Department of Management Information and Decision Systems
CASE WESTERN RESERVE UNIVERSITY
May, 1990.
Copyright (c) 1989 by KISHORE SENGUPTA
CASE WESTERN RESERVE UNIVERSITY

GRADUATE STUDIES

We hereby approve the thesis of

Kishore Sen Gupta

candidate for the Ph.D.
degree.*

Signed: [Signatures]

(Chairman)

Chancellor

Date: 1/31/1990

*We also certify that written approval has been obtained for any proprietary material contained therein.
I grant to Case Western Reserve University the right to use this work, irrespective of any copyright, for the University's own purposes without cost to the University or to its students, agents and employees. I further agree that the University may reproduce and provide single copies of the work, in any format other than in or from microforms, to the public for the cost of reproduction.

[Signature]

Kishore Sen Gupta
THE IMPACT OF COGNITIVE FEEDBACK ON GROUP DECISION MAKING

Abstract

by

KISHORE SENGUPTA

This study applied a three-system expansion of the lens model to analyze the impact of cognitive feedback at the interpersonal and group levels of decision making. The experimental was conducted in a computer-supported decision room with 90 subjects (split into 30 three-member groups), over four blocks. Subjects were given a candidate screening task involving three cues and were required to make estimates, individually and then in groups. Immediate cognitive feedback was provided on policy structures, consistency and the matching of members' policy structures. Outcome feedback was not provided. Decision quality was operationalized as consistency, matching index and agreement. The data was analyzed using a repeated-measures model.

Subjects in the feedback condition demonstrated significantly higher levels of consistency at the individual and group levels than subjects not receiving feedback. Subjects receiving feedback recorded a higher matching index than subjects not receiving feedback. This difference was significant at the group level, but not at the individual level. Recipients of cognitive feedback also registered higher
agreement than subjects in the no-feedback condition. Group behavior was analyzed from notes recorded by the experimenter. Groups in the feedback condition were more structured in their decision processes than groups in the no-feedback condition.

The study demonstrates the viability of cognitive feedback as a decision aid in systems for computer supported collaborative work. The results on interpersonal information processing also have specific implications in formulating design principles for multiuser interfaces.
DEDICATION

I dedicate this dissertation to my parents, Chira Ranjan and Pratima Sengupta.
ACKNOWLEDGEMENTS

I would like to thank the dissertation committee co-chairs Drs. Michael Ginzberg and Dov Te'eni, and members Drs. Patricia Brennan and Ananth Srinivasan, for serving on the committee, providing useful guidance at every stage of the project, and accommodating a rather constrained time schedule. I particularly appreciate the contribution Dr. Te'eni made by being closely involved in shaping every aspect of the thesis. Dr. Patricia Brennan gave extremely valuable guidance in clarifying the nuances of experimental design and data analysis.

Drs. Robert Colson, Michael Manning, Robert Mason, Gary Previts and Paul Salipante provided access to their students as experimental subjects. Dr. Jagdip Singh helped with data analysis and experimental design. Dr. Miles Kennedy, Chairman, and Robert Knight, Director, WSOM Computer Support Group, helped considerably by providing access to the facilities and equipment of the computer lab for running the experiment. Financial support provided by the department and a research initiation grant from the Ohio Regents Foundation made the dissertation possible.

Dr. Michael Doherty of the Department of Psychology, Bowling Green State University, helped with the experimental design, and provided valuable insights into the lens model. Marjorie Devany of the Naval Postgraduate School helped with the figures.

My personal debts are many and varied. Timothy Boyd-Wilson gave me access to his office and computer, and cheerfully acquiesced to my monopolizing the facilities. Hayagreeva Rao read the proposal
and offered several suggestions. Suzanne Robertson generously lent her computer at a particularly crucial juncture. At an equally critical moment, Patrick Lavelle resuscitated the hard disk of the computer (and with it, my dissertation) from the dead. Ronaldus Peeters did his bit by occasionally inquiring as to when the "masterpiece" would be ready for him to read. Special thanks to Aarti Badami for her friendship and support.

My siblings, Dr. Ajanta Poovaiah, Dr. Sumita and Sangeeta Sengupta contributed in very many ways, especially by not wondering aloud as to whether I would ever cease to be a student. My parents, through many years of support, sacrifice and encouragement, have made all of this possible. I appreciate the help given by N.K. Bose, A.C. Chakraborti and S.K. Gupta, who have been extremely generous and supportive over the years. Last, but certainly not the least, my thanks to Mariam, who has given me everything and much more.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>CHAPTER 1.</th>
<th>INTRODUCTION</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1 Research Context</td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>1.1.1 Group Decision Making</td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>1.1.2 Computer Supported Cooperative Work</td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>1.1.3 The Research Proposed Here</td>
<td></td>
<td>4</td>
</tr>
<tr>
<td>1.1.4 Relevance to Information Systems Research</td>
<td></td>
<td>4</td>
</tr>
<tr>
<td>1.2 Problem Domain</td>
<td></td>
<td>6</td>
</tr>
<tr>
<td>1.2.1 Cognitive feedback</td>
<td></td>
<td>6</td>
</tr>
<tr>
<td>1.2.2 Cognitive conflict</td>
<td></td>
<td>7</td>
</tr>
<tr>
<td>1.2.3 Research question</td>
<td></td>
<td>8</td>
</tr>
<tr>
<td>1.3 Contribution</td>
<td></td>
<td>8</td>
</tr>
<tr>
<td>1.3.1 Research in Information Systems</td>
<td></td>
<td>8</td>
</tr>
<tr>
<td>1.3.2 Research in Decision Feedback</td>
<td></td>
<td>9</td>
</tr>
<tr>
<td>1.4 Organization of the Thesis</td>
<td></td>
<td>10</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>CHAPTER 2.</th>
<th>THEORETICAL PREMISE</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1 The Research Question - A Conceptual Framework</td>
<td></td>
<td>11</td>
</tr>
<tr>
<td>2.2 Social Judgment Theory</td>
<td></td>
<td>13</td>
</tr>
<tr>
<td>2.2.1 Probabilistic Functionalism</td>
<td></td>
<td>13</td>
</tr>
<tr>
<td>Section</td>
<td>Title</td>
<td>Page</td>
</tr>
<tr>
<td>---------</td>
<td>----------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>2.2.2</td>
<td>Description of Human Judgment</td>
<td>15</td>
</tr>
<tr>
<td>2.2.3</td>
<td>A Taxonomy of Tasks</td>
<td>21</td>
</tr>
<tr>
<td>2.3</td>
<td><strong>Cognitive Conflict</strong></td>
<td>22</td>
</tr>
<tr>
<td>2.3.1</td>
<td>Defining Cognitive Conflict - a Triple-System Expansion of the Lens Model</td>
<td>22</td>
</tr>
<tr>
<td>2.3.2</td>
<td>The Cognitive Conflict Paradigm</td>
<td>26</td>
</tr>
<tr>
<td>2.3.3</td>
<td>Experimental Studies of Conflict</td>
<td>27</td>
</tr>
<tr>
<td>2.3.4</td>
<td>Cognitive Conflict Research: Findings</td>
<td>31</td>
</tr>
<tr>
<td>2.4</td>
<td><strong>Decision Feedback</strong></td>
<td>35</td>
</tr>
<tr>
<td>2.4.1</td>
<td>Feedback as a general concept</td>
<td>35</td>
</tr>
<tr>
<td>2.4.2</td>
<td>Individual versus Group feedback: the Referent Level</td>
<td>36</td>
</tr>
<tr>
<td>2.4.3</td>
<td>Two Conceptions of Feedback - Cognitive and Outcome</td>
<td>38</td>
</tr>
<tr>
<td>2.5</td>
<td><strong>Cognitive Feedback</strong></td>
<td>39</td>
</tr>
<tr>
<td>2.5.1</td>
<td>Defining Cognitive Feedback within the Lens Model Framework</td>
<td>39</td>
</tr>
<tr>
<td>2.5.2</td>
<td>Operationalizations of Cognitive Feedback</td>
<td>42</td>
</tr>
<tr>
<td>2.5.3</td>
<td>The Effectiveness of Cognitive Feedback</td>
<td>43</td>
</tr>
<tr>
<td>2.5.4</td>
<td>Cognitive feedback research: Issues</td>
<td>46</td>
</tr>
<tr>
<td>2.6</td>
<td><strong>Cognitive Feedback and the underlying Referent Level</strong></td>
<td>50</td>
</tr>
</tbody>
</table>
2.6.1 The plausibility of extending findings from Individual Literature 50
2.6.2 Cognitive Feedback at the Group Level 51
2.6.3 Cognitive Feedback at the Interpersonal Level 54

2.7 Framing the Research Question 56

CHAPTER 3. METHOD 58

3.1 Scope 58

3.2 Hypotheses 58
3.2.1 General Framework for Setting Hypotheses 58
3.2.2 H₁: Consistency 59
3.2.3 H₂: Matching Index 60
3.2.4 H₃: Agreement 61
3.2.5 H₄: No. of Iterations 62

3.3. Experimental Design 63
3.3.1 Description of Design 63
3.3.2 Phases of the Experiment 65
3.3.3 Experimental Setting 67

3.4 Subjects 68
3.4.1 Choice of Subjects 68
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.4.2</td>
<td>The use of Students as Subjects</td>
<td>72</td>
</tr>
<tr>
<td>3.5</td>
<td>Task</td>
<td>74</td>
</tr>
<tr>
<td>3.5.1</td>
<td>Description</td>
<td>74</td>
</tr>
<tr>
<td>3.5.2</td>
<td>Task Characteristics</td>
<td>75</td>
</tr>
<tr>
<td>3.6</td>
<td>Cognitive Feedback</td>
<td>76</td>
</tr>
<tr>
<td>3.6.1</td>
<td>Policy feedback</td>
<td>76</td>
</tr>
<tr>
<td>3.6.2</td>
<td>Consistency feedback</td>
<td>79</td>
</tr>
<tr>
<td>3.6.3</td>
<td>Interpersonal Feedback.</td>
<td>82</td>
</tr>
<tr>
<td>3.6.4</td>
<td>Screen Design</td>
<td>84</td>
</tr>
<tr>
<td>3.7</td>
<td>The Training Phase</td>
<td>86</td>
</tr>
<tr>
<td>3.7.1</td>
<td>Basis for Training</td>
<td>86</td>
</tr>
<tr>
<td>3.7.2</td>
<td>Training Procedure</td>
<td>86</td>
</tr>
<tr>
<td>3.7.3</td>
<td>Verifying the Effectiveness of Training</td>
<td>87</td>
</tr>
<tr>
<td>3.8</td>
<td>Experimental Phase</td>
<td>94</td>
</tr>
<tr>
<td>3.8.1</td>
<td>Individual Task</td>
<td>94</td>
</tr>
<tr>
<td>3.8.2</td>
<td>Group Task</td>
<td>95</td>
</tr>
<tr>
<td>3.8.3</td>
<td>Manipulation Check</td>
<td>95</td>
</tr>
<tr>
<td>3.9</td>
<td>Debriefing</td>
<td>97</td>
</tr>
<tr>
<td>3.10</td>
<td>Summary</td>
<td>98</td>
</tr>
</tbody>
</table>
CHAPTER 4. RESULTS

4.1 Scope

4.2 Framework for the Analysis of Data
   4.2.1 Model for Analysis
   4.2.2 Dependent Variables - Individual Data
   4.2.3 Dependent Variables - Group Data

4.3 Experimental Results: Individual Data
   4.3.1 Consistency:
   4.3.2 Matching among Members' Models
   4.3.3 Agreement among Members' Estimates
   4.3.4 Number of Iterations

4.4 Experimental Results: Group Data
   4.4.1 Consistency
   4.4.2 Matching with Members' Individual Models
   4.4.3 Agreement with Members' Individual Estimates
   4.4.4 Groups' use of Feedback
   4.4.5 Comparison with Cognitive Conflict Studies

4.5 Summary of Results
CHAPTER 5. CONCLUSIONS AND LIMITATIONS 137

5.1 Scope 137

5.2 Conclusions of the Study 137
  5.2.1 A Recapitulation of Past Research 138
  5.2.2 Research Question 139
  5.2.3 Findings of the Study 140
  5.2.4 Consistency in Decision Making 141
  5.2.5 Feedback and Group Behavior 143

5.3 Applications of Cognitive Feedback 144
  5.3.1 Cognitive Feedback in CSCW 144
  5.3.2 Knowledge Acquisition in Expert Systems 148

5.4 Limitations and Further Research 150
  5.4.1 Cognitive Conflict: the Need for a Richer Perspective 150
  5.4.2 Issues in Cognitive Feedback Research 151

5.5 Summary 156

APPENDIX 157

REFERENCES 169
# LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>19</td>
</tr>
<tr>
<td>Lens model measures</td>
<td></td>
</tr>
<tr>
<td>2.2</td>
<td>41</td>
</tr>
<tr>
<td>Components of Cognitive Feedback</td>
<td></td>
</tr>
<tr>
<td>2.3</td>
<td>44</td>
</tr>
<tr>
<td>Effectiveness of Cognitive Feedback</td>
<td></td>
</tr>
<tr>
<td>Knowledge</td>
<td></td>
</tr>
<tr>
<td>2.4</td>
<td>44</td>
</tr>
<tr>
<td>Effectiveness of Cognitive Feedback</td>
<td></td>
</tr>
<tr>
<td>Achievement</td>
<td></td>
</tr>
<tr>
<td>2.5</td>
<td>45</td>
</tr>
<tr>
<td>Effectiveness of Cognitive Feedback</td>
<td></td>
</tr>
<tr>
<td>Control</td>
<td></td>
</tr>
<tr>
<td>3.1</td>
<td>65</td>
</tr>
<tr>
<td>Research Design</td>
<td></td>
</tr>
<tr>
<td>3.2 (a)</td>
<td>70</td>
</tr>
<tr>
<td>Participant Profiles - Demographics</td>
<td></td>
</tr>
<tr>
<td>3.2 (b)</td>
<td>72</td>
</tr>
<tr>
<td>Participant Profiles - Task-specific Factor</td>
<td></td>
</tr>
<tr>
<td>3.3</td>
<td>88</td>
</tr>
<tr>
<td>Subjects' Matching Index (G) for Training</td>
<td></td>
</tr>
<tr>
<td>Means and Standard Deviations</td>
<td></td>
</tr>
<tr>
<td>3.4</td>
<td>89</td>
</tr>
<tr>
<td>Subjects' Achievement (ra) in Training</td>
<td></td>
</tr>
<tr>
<td>Means and Standard Deviations</td>
<td></td>
</tr>
<tr>
<td>3.5</td>
<td>96</td>
</tr>
<tr>
<td>Sequence of Tasks within each Block of the Experimental Phase</td>
<td></td>
</tr>
<tr>
<td>3.6</td>
<td>98</td>
</tr>
<tr>
<td>Subject Attitudes toward Task</td>
<td></td>
</tr>
<tr>
<td>4.1</td>
<td>102</td>
</tr>
<tr>
<td>Individual Data</td>
<td></td>
</tr>
<tr>
<td>Sample Sizes for Dependent Variables</td>
<td></td>
</tr>
<tr>
<td>4.2</td>
<td>104</td>
</tr>
<tr>
<td>Group Data</td>
<td></td>
</tr>
<tr>
<td>Sample Sizes for Dependent Variables</td>
<td></td>
</tr>
<tr>
<td>4.3 (a)</td>
<td>105</td>
</tr>
<tr>
<td>Individual Data: Between Subjects Analysis</td>
<td></td>
</tr>
<tr>
<td>Dependent Variable: Consistency</td>
<td></td>
</tr>
<tr>
<td>4.3 (b)</td>
<td>108</td>
</tr>
<tr>
<td>Individual Data: Within Subjects Analysis</td>
<td></td>
</tr>
<tr>
<td>Dependent Variable: Consistency</td>
<td></td>
</tr>
<tr>
<td>4.3 (c)</td>
<td>109</td>
</tr>
<tr>
<td>Individual Data: Trend Analysis for Blocks</td>
<td></td>
</tr>
<tr>
<td>Dependent Variable: Consistency</td>
<td></td>
</tr>
<tr>
<td>Table</td>
<td>Page</td>
</tr>
<tr>
<td>-------</td>
<td>------</td>
</tr>
<tr>
<td>4.4 (a) Individual Data: Between Subjects Analysis&lt;br&gt;Dependent Variable: Matching Index</td>
<td>109</td>
</tr>
<tr>
<td>4.4 (b) Individual Data: Within Subjects Analysis&lt;br&gt;Dependent Variable: Matching Index</td>
<td>111</td>
</tr>
<tr>
<td>4.4 (c) Individual Data: Trend Analysis for Blocks&lt;br&gt;Dependent Variable: Matching Index</td>
<td>111</td>
</tr>
<tr>
<td>4.5 (a) Individual Data: Between Subjects Analysis&lt;br&gt;Dependent Variable: Agreement</td>
<td>112</td>
</tr>
<tr>
<td>4.5 (b) Individual Data: Within Subjects Analysis&lt;br&gt;Dependent Variable: Agreement</td>
<td>112</td>
</tr>
<tr>
<td>4.5 (c) Individual Data: Trend Analysis for Blocks&lt;br&gt;Dependent Variable: Agreement</td>
<td>112</td>
</tr>
<tr>
<td>4.6 (a) Individual Data: Between Subjects Analysis&lt;br&gt;Dependent Variable: No. of Iterations</td>
<td>114</td>
</tr>
<tr>
<td>4.6 (b) Individual Data: Within Subjects Analysis&lt;br&gt;Dependent Variable: No. of Iterations</td>
<td>115</td>
</tr>
<tr>
<td>4.6 (c) Individual Data: Trend Analysis for Blocks&lt;br&gt;Dependent Variable: No. of Iterations</td>
<td>115</td>
</tr>
<tr>
<td>4.7 (a) Group Data: Between Subjects Analysis&lt;br&gt;Dependent Variable: Consistency</td>
<td>118</td>
</tr>
<tr>
<td>4.7 (b) Group Data: Within Subjects Analysis&lt;br&gt;Dependent Variable: Consistency</td>
<td>118</td>
</tr>
<tr>
<td>4.7 (c) Group Data: Trend Analysis for Blocks&lt;br&gt;Dependent Variable: Consistency</td>
<td>121</td>
</tr>
<tr>
<td>4.8 (a) Group Data: Between Subjects Analysis&lt;br&gt;Dependent Variable: Matching Index</td>
<td>121</td>
</tr>
<tr>
<td>4.8 (b) Group Data: Within Subjects Analysis&lt;br&gt;Dependent Variable: Matching Index</td>
<td>121</td>
</tr>
<tr>
<td>4.8 (c) Group Data: Trend Analysis for Blocks&lt;br&gt;Dependent Variable: Matching Index</td>
<td>122</td>
</tr>
<tr>
<td>4.9 (a) Group Data: Between Subjects Analysis&lt;br&gt;Dependent Variable: Agreement</td>
<td>124</td>
</tr>
<tr>
<td>Table</td>
<td>Page</td>
</tr>
<tr>
<td>-------</td>
<td>------</td>
</tr>
<tr>
<td>4.9 (b)</td>
<td>124</td>
</tr>
</tbody>
</table>
| Group Data: Within Subjects Analysis  
   *Dependent Variable: Agreement* | |
| 4.9 (c) | 126 |
| Group Data: Trend Analysis for Blocks  
   *Dependent Variable: Agreement* | |
| 4.10 (a) | 127 |
| Group Data: Analysis of Notes  
   *Formulation of Decision Rule* | |
| 4.10 (b) | 127 |
| Group Data: Analysis of Notes  
   *Time of Formation of Decision Rule* | |
| 4.10 (c) | 129 |
| Group Data: Analysis of Notes  
   *Use of Feedback by Groups* | |
| 4.10 (d) | 130 |
| Group Data: Analysis of Notes  
   *How Consistency was used as Feedback* | |
| 4.10 (e) | 131 |
| Group Data: Analysis of Notes  
   *How Policy was used as Feedback* | |
| 4.10 (f) | 132 |
| Group Data: Analysis of Notes  
   *How Matching was used as Feedback* | |
| 4.11 (a) | 135 |
| Summary of Results  
   *Individual Data* | |
| 4.11 (b) | 136 |
| Summary of Results  
   *Group Data* | |
| 5.1 | 155 |
| A Revised Experimental Design | |
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Lens Model - Regression Formulation</td>
<td>18</td>
</tr>
<tr>
<td>2.2</td>
<td>Lens Model - Triple-system Expansion</td>
<td>24</td>
</tr>
<tr>
<td>3.1</td>
<td>Experimental Sequence</td>
<td>66</td>
</tr>
<tr>
<td>3.2</td>
<td>Layout of Experiment Room</td>
<td>67</td>
</tr>
<tr>
<td>3.3</td>
<td>Policy Feedback</td>
<td>78</td>
</tr>
<tr>
<td>3.4</td>
<td>Consistency Feedback</td>
<td>81</td>
</tr>
<tr>
<td>3.5</td>
<td>Interpersonal Feedback</td>
<td>83</td>
</tr>
<tr>
<td>3.6</td>
<td>Full Screen for Experiment</td>
<td>85</td>
</tr>
<tr>
<td>3.7</td>
<td>Training Screen</td>
<td>87</td>
</tr>
<tr>
<td>3.8 (a)</td>
<td>Training - Work Experience</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Subjective Weights versus Actual Weights</td>
<td>91</td>
</tr>
<tr>
<td>3.8 (b)</td>
<td>Training - Test Scores</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Subjective Weights versus Actual Weights</td>
<td>92</td>
</tr>
<tr>
<td>3.8 (a)</td>
<td>Training - Education</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Subjective Weights versus Actual Weights</td>
<td>93</td>
</tr>
<tr>
<td>4.1</td>
<td>Individual Data: Consistency across Blocks</td>
<td>106</td>
</tr>
<tr>
<td>4.2</td>
<td>Individual Data: Matching Index across Blocks</td>
<td>110</td>
</tr>
<tr>
<td>4.3</td>
<td>Individual Data: Agreement across Blocks</td>
<td>113</td>
</tr>
<tr>
<td>4.4</td>
<td>Individual Data: No. of Iterations across Blocks</td>
<td>116</td>
</tr>
<tr>
<td>4.5</td>
<td>Group Data: Consistency across Blocks</td>
<td>119</td>
</tr>
<tr>
<td>Figure</td>
<td>Group Data:</td>
<td>Page</td>
</tr>
<tr>
<td>--------</td>
<td>-------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>4.6</td>
<td>Matching Index across blocks</td>
<td>123</td>
</tr>
<tr>
<td>4.7</td>
<td>Agreement across blocks</td>
<td>125</td>
</tr>
</tbody>
</table>
CHAPTER 1
INTRODUCTION

1.1 Research Context

1.1.1 Group Decision Making

Group decision making has been a subject of continuing attention in several disciplines. Researchers in social psychology have studied small groups from the point of view of their internal dynamics (Shaw, 1981), group versus individual performance (Hill, 1982) and coalition formation within groups (Murnighan, 1978). Political scientists have been interested in the formation of governments and the concomitant expectation of policies (Ordershock, 1986). Economists have deliberated on groups through the conceptual lens of social choice theory (Elster, 1983, p. 26-42).

1.1.2 Computer Supported Cooperative Work

The potential of computerized support for decision making has long been recognized by researchers in information systems (Keen and Scott-Morton, 1978). Organization theorists have noted that in an era of complexity and turbulence, decisions in organizations are increasingly being made in groups (Weick, 1979; Huber, 1984). With decision-related meetings becoming more frequent, the imperatives for increasing the efficiency and effectiveness of these sessions are strong (Gray, 1986; Huber and McDaniel, 1986). This has prompted organizational scientists to create new technologies for structuring
group processes (Deibecq, Van de Ven and Gustafson, 1975). More recently, information systems researchers have sought to apply computer support technologies to improve group decision making. The set of research efforts is collectively referred to as computer-supported cooperative work (CSCW).

Much of research in CSCW has concentrated on the development of technologically supported means of collecting, managing and displaying information that may be useful in decision situations (Kraemer and King, 1988). Notable results include the development of feasible group-support architectures (Applegate, Konsynski and Nunamaker, 1986), shared interfaces (Stefik, Foster, Bobrow, Kahn, Lanning and Suchman, 1987), meetings augmentation mechanisms (Cook, Ellis, Graf, Rein and Smith, 1987), data sharing devices (Greif and Sarin, 1987), and specialized tools for filtering communication (Malone, Grant, Turbak, Brobst and Cohen, 1987).

Kraemer and King (1988) note that researchers in CSCW have a bias toward developing technological aids that they presume will be needed by decision makers. The development of such tools is often not tied to specific aspects of individual and group decision processes. Consequently, it is not clear what impact these tools have on decision making1. The electronic boardroom is a case in point. Beyond purporting to facilitate group decisions by speeding up certain tasks

---
1. Kraemer and King suggest that some of these tools may have "too little" an impact for a good assessment of their effects to be made. This argument may be interpreted as suggesting that the class of tools represented by boardrooms or messaging (level 1 systems according to DeSanctis and Gallupe, 1987) improve communication but do not affect the decision processes strongly enough to enhance decision quality.
(e.g., voting) or providing more flexibility in meeting opportunities, it makes no specific claim about its impact on decision making. Its effectiveness as a decision aid, therefore, is not clear. This leads Kraemer and King to propose that more rigorous and detailed empirical research be conducted on the effectiveness of and experiences with existing systems. The argument also points to the need to develop tools and techniques specifically tied to individual and group decision processes.

There is some empirical research examining the impact of computerized support on groups, e.g., Applegate (1986), Eveland and Bikson (1988), and Zigurs (1988). The studies done so far differ considerably in the type of system used, choice of dependent variable, task and experimental methods, and are not strictly comparable except in the general sense of there being some kind of computerized support as a treatment. The most consistent set of results emerging from these studies is that the benefits of CSCW are predominantly affective (Kraemer and King, 1988). While the importance of affective benefits in group decision making is undeniable (Turoff and Hiltz, 1982), whether these benefits contribute to higher quality group decisions, is not clear. The potential for improving the quality of the group decision is evident when we recognize that groups do not perform in any optimal sense. While groups outperform their average member, their performance seldom reaches the level of the best member (Einhorn,

---

2. A notable exception is the study done by Gault, DeSanctis and Dickson (1988). This study measured decision quality as the proximity of a group's solution with that prescribed by the experts.
Hogarth and Klempner, 1977). Thus, though there is clearly scope for improving group performance with computer support, the existing set of CSCW techniques do not appear to effectively tap that potential. There is, therefore, a need to develop support tools that categorically aim at improving group decision quality.

1.1.3 The Research Proposed Here

The research described here addresses the two points made above - (a) the need to develop devices that have a clear nexus with individual and group decision processes, devices (b) that specifically aim to affect the quality of the decision. In this study, we consider the utility of decision feedback as a mechanism for aiding decision making. In examining the plausibility of incorporating such feedback in CSCW, we also make specific claims about the role of such feedback in improving group decisions. Furthermore, these effects are hypothesized to improve the quality of the group decision. There is, thus, a hypothesized linkage between the use of the tool and the quality of the group decision.

1.1.4 Relevance to Information Systems Research

The relevance of this research to information systems can perhaps be best stated in terms of the framework for group support research outlined by DeSanctis and Gallepe (1987). The authors propose a staged approach to developing computerized support. Thus, systems at level 1 should provide technical features simply aimed at facilitating communication, such as large screens, voting
compilation, facilities for anonymous ideas and electronic messaging. These features are found in computer-supported conference rooms and electronic boardrooms. Level 1 features seek to improve the decisions by facilitating information exchange among members. Most group support tools fall in this category. At level 2, group support is sought to be enhanced by providing techniques that directly affect the actual decision processes. Examples of such tools are problem structuring, decision modeling, simulation and preference formulation aids. Level 3 systems seek to actively affect the communication patterns (as opposed to merely facilitating them) through such devices as enforcing a set of rules for discussion.

The use of decision feedback as an aid straddles issues at levels 2 and 3. Insofar as it provides support for decisions, feedback as a tool clearly falls within the ambit of level 2. However, in providing feedback, the system also imposes a temporal structure on communications and decision making within the group. The timing of feedback assumes subjects to make decisions in a certain sequence, individually and then in the group mode. Doing this also alters the respective decision processes of individuals. Thus, recognition that providing feedback constitutes an intervention into group structure, implicitly affects issues at level 3 as well.

3. This individual-group sequence of decision making implicit in our discussion is widely prevalent in group decision making research, e.g., the risky-shift paradigm (Myers and Lamm, 1976) or the social decision scheme model (Davis, 1980).
1.2 Problem Domain

In this section, we outline the elements of the problem domain, thereby constituting the basis for stating the research question. Two principal elements impact the research question - the notion of feedback as a decision aid, and the recognition that conflict in groups has a cognitive aspect. Details of the theoretical underpinnings are addressed in chapter 2.

1.2.1 Cognitive feedback\textsuperscript{4}

The role of feedback mechanisms in affecting individual behavior is widely acknowledged (Wiener, 1948, 1950; Annett, 1969). Feedback is categorized as being of two types - outcome and cognitive. Outcome feedback is information about the accuracy of a response, while cognitive (or process) feedback provides information about the how and why of the accuracy of the response.

This research explores feedback as cognitive feedback - specifically whether providing cognitive feedback affects the quality of the decision in group situations. In doing so, it captures the interplay between decisions at two levels of analysis - the individual and the group.

\textsuperscript{4} The discussion here introduces the notion of feedback in a very general sense. Sections 2.4, 2.5 and 2.6 (chapter 2) elaborate on the concept of feedback, and provide explanations of the meaning and implications of outcome and cognitive feedback.
1.2.2 Cognitive conflict

The notion of cognitive conflict was first suggested by Hammond (1965). In broad terms, cognitive conflict can be understood as a counterpoint to the conventional conceptualization of conflict as caused by mixed motives.

Conflict in groups has been studied from several perspectives. For example, researchers in industrial relations have explored tasks involving multiple dimensions of dissent (as typically present in labor-management negotiations - McGrath, 1966). Economists, on the other hand, have studied tasks in which two (or more) parties' joint choices determine pay-offs to each (as in the prisoners' dilemma game - Pruitt and Kimmel, 1977). A common thread underlying such research is the imputation of differences of interest (or mixed-motives) as the cause of conflict among individuals. Hammond (1965) argues that this view of conflict is predicated on the notion that individuals are largely cognizant of their respective viewpoints and interests. This assumption, however, is tenuous in the light of evidence from research in human information processing. Human decision making is covert (internalized to the extent that not even the decision maker has full sense of it), not easily described, and is marked by inconsistencies (Balke, Hammond and Meyer, 1973). Moreover, not all conflicts can be attributed to differences in interest among individuals (Brehmer, 1976). Individuals in a group often disagree even in the absence of interest differentials. Such conflicts, known as cognitive conflicts, arise because group

---

5. Cognitive conflict is defined and surveyed in section 2.3 (Chapter 2).
members differ in their underlying cognitive views of the problem (Hammond, 1965), and can be difficult to resolve because their causes are (a) not understood properly, and (b) are often misdiagnosed as arising from mixed motives (Balke et al., 1973).

From the perspective of our research question, cognitive conflict is understood as conflict that is not marked by clear interest differentials among the parties. Since cognitive conflicts are commonly observed in group decision situations (Brehmer, 1976), the research question examines the effectiveness of decision feedback in the presence of such conflicts.

1.2.3 Research question

The research question, broadly stated, is: Does cognitive feedback improve the quality of group decision making?

1.3 Contribution

The contributions of this research are discussed from two perspectives: research in information systems and research on decision feedback.

1.3.1 Research in Information Systems

The utility of decision feedback in aiding complex decision tasks is well-documented (Hammond, Mumpower and Smith, 1977; Kleinmuntz, 1985; Lewis, Shields and Young, 1983; Sterman, 1989). Feedback has also been found to affect group performance (Zajonc, 1962; Zander and Wolfe, 1964). However, the issue of decision
feedback has not been explicitly addressed in CSCW research, even though its potential for enhancing group performance has been acknowledged (Nadler, 1979). The design of feedback has also been found to be crucial in enhancing decision performance (Ilgen, Fisher, and Taylor, 1979; Jacoby, Mazursky, Troutman, and Kuss, 1984). It is, therefore, important that research be conducted to determine how to design decision feedback.

The contribution of this research, is therefore, in establishing the importance of feedback as a decision tool. We add to the body of knowledge in CSCW by providing the possibility of incorporating feedback mechanisms in computerized support systems.

1.3.2 Research in Decision Feedback

Feedback can be administered as several referent levels\(^6\), and the problems associated with each level are quite distinct (Nadler, 1979). Cognitive feedback has been studied quite extensively in the context of the individual (section 2.5, chapter 2). However, beyond suggestions by Hammond and Brehmer (1973), the use of cognitive feedback as a decision aid has not been systematically extended to other referent levels.

The contribution of this research to the area of decision feedback is to (a) delineate issues that are appropriate to different referent levels and (b) present experimental evidence on the efficacy of cognitive feedback at these levels.

---

6. The question of the referent level is discussed in section 2.4.2, chapter 2.
1.4 Organization of the thesis

The discussion proceeds as follows. Chapter 2 reviews relevant research in human information processing, cognitive conflict and feedback. Chapter 3 details the experimental setting and chapter 4 discusses the data analysis and results. Chapter 5 proposes potential applications of the concept, limitations of the research and possible extensions for future research.
CHAPTER 2
THEORETICAL PREMISE

2.1 The Research Question - A Conceptual Framework

The research question encompasses two distinct but related, theoretical areas: cognitive conflict and decision feedback. This chapter presents an overview of each area and delineates the issues in order to provide a theoretical underpinning for the research question.

Cognitive Conflict

The notion of cognitive conflict is derived from research in social judgment theory (SJT). Section 2.2 outlines core principles of SJT that are germane to a discussion of cognitive conflict. A crucial element of SJT is the representation of human judgment. This is accomplished through the use of the lens model (Hursch, Hammond and Hursch, 1964; Tucker, 1964). Section 2.2.2 discusses the representation of human judgment in SJT, and more generally, issues surrounding the use of linear models in judgment research.

Another salient aspect of SJT relevant to subsequent discussion is the enumeration of a systematic taxonomy of tasks. This framework is a convenient reference point for further discussion of extant research in cognitive conflict (and decision feedback as well). Section 2.2.3 outlines a taxonomy of tasks widely used by SJT researchers.

The use of SJT as a way to model cognitive conflict was first suggested by Hammond (1965). Section 2.3 discusses cognitive conflict in three principal dimensions - a definition and formulation of
the concept in terms of the lens model (section 2.3.1), a description of
the cognitive conflict paradigm of experimentation (section 2.3.2), and
synthesis of research within the paradigm (section 2.3.3). Section
2.3.4 summarizes the central findings.

Decision Feedback

Feedback is a powerful mechanism for affecting behavior (Ilgen
et al., 1979) and is central to the acquisition of expertise (Hogarth,
1981). Section 2.4 discusses the elements of decision feedback that
pertain to the research question. Section 2.4.1 introduces the notion of
feedback *per se*. Section 2.4.2 makes a distinction between individual
and group feedback and clarifies the role of the underlying referent
setting.

Decision feedback is commonly classified as being of two types
- *cognitive* or *process* feedback (the two words are used
interchangeably) and *outcome* feedback (Hammond, Stewart, Brehmer
and Steinmann, 1975; Hogarth, 1981). This distinction is central to the
research question and is clarified in section 2.4.3. Section 2.5 surveys
cognitive feedback research. Section 2.6 spells out the factors that
arise in examining the nexus between cognitive feedback and the
underlying referent setting.

The Research Problem

The research problem can be framed in the following sequence:

(i) Individuals make decisions in groups.

(ii) Members in a group often have different mental models of the task
    / decision at hand.

(iii) As a result of (ii), there is disagreement, or, *cognitive conflict* in
groups.
(iv) Cognitive (or process) feedback has been found to be effective in improving decision making, by providing decision makers with greater insight into their judgment.

(v) So, (a) given the existence of cognitive conflict within groups, (b) does providing cognitive feedback help the group (c) improve decision quality?¹

Part (i) concerns basic issues in describing human judgment. These are covered in section 2.2. Parts (ii) and (iii), which deal with the nature and origin of cognitive conflict, are discussed in section 2.3. Part (iv) deals with the issue of feedback. This is covered in sections 2.4, 2.5 and 2.6. Section 2.7 defines part (v), i.e., the research problem, in the context of the issues emerging from the literature review. In this sense, it establishes the point of departure for the study. The remaining chapters elaborate on the implementation of the research question.

2.2 Social Judgment Theory

2.2.1 Probabilistic Functionalism

SJT is based on Brunswik’s theory of probabilistic functionalism (Brunswik, 1943, 1952), a theory that emphasizes the role of the environment in fashioning an organism’s coping mechanisms. Brunswik (1943) suggested that an organism’s environment is complex. However, the environment also contains considerable redundancy in information cues (Brunswik, 1943, 1952). An organism takes this redundancy into account in adapting to the environment,

¹ As is explained below, the quality of decision in this study is interpreted in terms of the lens model and its indices.
through the process of vicarious functioning. Vicarious functioning refers to the use of cues interchangeably so that different patterns of cues can lead to equivalent results (Postman and Tolman, 1959, p. 553). This ability to substitute for and make trade-offs between cues is seen as an essential survival skill in a complex environment (Brunswik, 1943).

According to Brunswik, the representation of human judgment involves two important considerations: (a) making a distinction between proximal variables, called cues, and a distal variable, called the criterion, and (b) modeling the task (or environment) and the inference mechanism of the individual within the same framework (Brehmer, 1979).

Brunswik also argued that the environment is uncertain, i.e., the relation between the cues and the criterion is probabilistic (Brehmer, 1979). In other words, cues do not permit a person to make inferences about the criterion with complete certainty. Brunswik maintains that all important inference tasks are probabilistic (Brunswik, 1943). Surveys of clinical tasks (Brehmer, 1976b) and social and political decision problems (Hammond, 1974) lend considerable credence to this assertion. Therefore, a conceptual framework for studying inference tasks must be able to handle the problem of uncertainty (Brehmer, 1979). Specifically, the concepts must enable us to express, not only the regularities of the task system, but also its irregularities. This means that statistical concepts, such as regression or the analysis of variance, are required to describe the task.
2.2.2 Description of Human Judgment

SJT uses statistical models (such as multiple regression and analysis of variance) to describe human judgment. The use of linear models was first suggested by Brunswik (1943) for studies of perception, and later by Hammond (1955) and Hoffman (1960) for the analysis of judgment. Since then, such models have been primarily used in studying the feedback/learning process or the integration of information from several cues (Slovic and Lichtenstein, 1971). Within this paradigm, such models capture the process of vicarious functioning in four important ways (Einhorn, Kleinmuntz and Kleinmuntz, 1979):

First, the additive combination function implies a fully compensatory system. Second, the degree to which the cues trade off (determined by the betas) is a function of the particular environment in which the judgment is made; that is, beta weights are determined by considering all the cues and their particular levels in the situation. Third, cue redundancy is incorporated in the model since the beta weights are determined by the correlational structure of the cues (and the correlation of the cues with the judgment). Moreover, the indeterminacy in estimating weights when cues are correlated parallels the organism's difficulty in this matter. Also, redundancy is defined with respect to the particular environment being studied. Fourth, the inconsistency and random error in judgment, resulting from the lack of cognitive control in executing one's strategy is explicitly defined and measured within regression procedures (p 467).

The Lens Model

The lens model (Hursch et al., 1964; Tucker 1964) describes tasks and a decision maker's inference through an organizing principle
for the cues and the criterion. Figure 2.1 presents a regression formulation of the lens model (Dudycha and Naylor, 1966). Lens model measures are summarized in table 2.1. The model has three components: (a) the task environment, defined by the cue set \( X = (X_1, X_2, \ldots, X_n) \), (b) the criterion event \( Y_0 \), and (c) the judge’s decision \( Y_s \). On the environmental (i.e., left) side of the model, the correlation \( r_{ie} \) between each cue \( X_i \) and the criterion event \( Y_0 \) is called the ecological validity of the cue. This measure can be interpreted as depicting the relevance of the \( i \)th cue in predicting the criterion event, independent of other cues. The relationship between the (entire set of) cues and the criterion is reflected in the correlation of the criterion event \( Y_0 \) and the prediction of the criterion event from the underlying model \( Y_0^* \). This measure is called environmental predictability \( R_0 \) and can be interpreted as reflecting the relevance of the cue set in predicting the criterion event.

A series of analogous measures describes the decision maker’s judgment process. The decision maker’s reliance on each cue, denoted by \( r_{is} \), is called the utilization coefficient. The correlation between the actual response and the prediction of the model is called consistency \( R_s \), and is interpreted as the decision maker’s cognitive control (Hammond and Summers, 1972). The similarity of the decision maker’s use of cues with the environmental relationships is called the matching index \( G \).

---

2. The model depicted here is univariate in that it has one dependent variable. The univariate model has long been the dominant mode of research in the lens model framework. Cooksey and Freebody (1985), and Stewart (1976) have recently proposed multivariate extensions of the model.
The judge's performance is summarized in the achievement index \( r_a \), which indicates the correspondence between the subject's response and the environmental event. \( r_a \) is denoted by the lens model equation (Tucker, 1964):

\[
ra = G \, Re \, Rs
\]

Equation (1) suggests that achievement is a function of three components - how closely the decision maker's mental model matches that of the environment (\( G \)), the predictability of the environment (\( Re \)), and the decision maker's consistency (\( Rs \)).
Figure 2.1

Lens Model - Two System View
(Dudycha and Naylor, 1966)

Achievement
\[ r_a = G_e R_s \]

Cue Set

Criterion Event
\[ = Y_e \]

Environmental Predictability
\[ R_e = Y_e Y_e \]

Predicted Criterion Event
\[ = Y_e \]

Matching Index
\[ G = r \hat{Y}_e \hat{Y}_s \]

Consistency
\[ R_s = Y_s Y_s \]

Predicted Judge's Decision
\[ = Y_s \]

Judge's Decision
\[ = Y_s \]
### TABLE 2.1: Lens Model Measures

<table>
<thead>
<tr>
<th>Name</th>
<th>Symbol</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ecological validity</td>
<td>$r_{le}$</td>
</tr>
<tr>
<td>Utilization coefficient</td>
<td>$r_{ls}$</td>
</tr>
<tr>
<td>Environmental predictability</td>
<td>$R_{e}$</td>
</tr>
<tr>
<td>Consistency</td>
<td>$R_{s}$</td>
</tr>
<tr>
<td>Achievement</td>
<td>$r_{a}$</td>
</tr>
<tr>
<td>Matching index</td>
<td>$G$</td>
</tr>
</tbody>
</table>

The Use of Linear Models in Studying Human Judgment

Linear models (of which the regression formulation of the lens model is an example) have been extensively used in judgment research, especially for studies of policy capturing, clinical judgment and multiple cue probability learning. Such models have been generally successful in predicting performance in a variety of tasks (Slovic and Lichtenstein, 1971), strongly suggesting that they capture a fundamental characteristic of judgment (Goldberg, 1970).

From another perspective, the goodness of fit of the linear model is posited to be more a function of the characteristics of the task and the insensitivity of the linear model than a characteristic of human behavior (Dawes and Corrigan, 1974; Slovic, Fischhoff and Lichtenstein, 1977). The predictive power of the linear model may occur despite serious flaws. For example, two models may be algebraically equivalent and yet suggest very different cognitive processes (Hoffman, 1960).

---

3. A recent study (Johnson, Meyer and Ghose, 1989) concludes that this robustness to misspecification of functional form may fail when there are negative correlations among attributes of a choice set.
The fundamental question in decision research is: what cognitive processes underlie judgment and choice? Judgment researchers have emphasized the desirability of studying decision making through multiple methods (Einhorn and Hogarth, 1981; Svenson, 1979). Studies that have directly compared different methods (e.g., linear models and process tracing) have found diverging results. For example, Einhorn et al. (1979) found linear models to provide a better fit of the judgment process than models elicited through process tracing methods. The authors attribute these differences more with respect to emphasis and descriptive level of detail than the underlying process uncovered. Billings and Marcus (1983), who also obtained diverging results, suggest that such lack of convergence is because the two methods represent different stages of the decision process. Process tracing examines the information acquisition process while linear model approaches study the integration or combination of information into decisions.

The essence of the discussion is captured in the recognition that all methods of representing judgment are at best paramorphic (Hoffman, 1960; Einhorn, 1971). Einhorn et al. (1979) point out that criticisms of linear models are based on the erroneous belief that the mathematical form of the model is isomorphic with the process it is supposed to represent:

Experts shown a linear model of their judgments may claim that they do not multiply cues by weights and then add up the products to form their judgment. This illustrates the misunderstanding referred to. An analogy might be that a quarterback does not compute a set of differential equations in throwing a touchdown pass,
although one might model that activity in such a way (p 469, note 5).

Both process tracing and linear models provide valuable insights into the judgment process, but of a different nature (Einhorn et al., 1979). The detailed level of analysis required in process tracing often precludes the use of large numbers of subjects or even decisions which would allow for the estimation of error in judgments about strategy use. For example, regression analyses can provide estimates of systematic error and error variance in judgments. The relative importance of a particular cue in affecting judgments is difficult to determine in a process tracing model, but regression weights and the intercorrelation matrix can provide some of this information. Alternatively, the results of a process tracing study may serve to better inform one regarding the types of nonlinear effects to include in a regression analysis. On the whole, therefore, linear models constitute a valid and useful mode of representing judgment4.

2.2.3 A Taxonomy of Tasks

Brehmer (1976, 1979) offers a classification of tasks. By this classification, a task may be described with respect to its contents as well as with respect to its formal characteristics. Formal characteristics comprise a set of general dimensions that can be used to describe all inference tasks. There are seven such dimensions, divided into two classes - (a) surface characteristics and (b) system characteristics.

4. This assumes conditional monotonicity, i.e., cues are monotonically related to the criterion (Dawes, 1975).
(a) Surface characteristics: They refer to the nature of the cues. There are three surface characteristics. The number of cues indicates the number of proximal variables to be taken into account in inferring the criterion. The metric characteristics of the cues specify whether the variables are nominal or quantitative. Cue intercorrelations denote the extent to which the cues tend to go together. Intercorrelations indicate the extent to which cues are interchangeable.

(b) System characteristics: These attributes clarify the relation between the cues and the criterion. Relative weights identify the importance of each cue in determining the criterion. The functional relation between a cue and the criterion indicates whether the relation is linear or non-linear. The integration rule specifies the manner in which cues are combined, such as, a simple additive or averaging rule or a configurual rule. The predictability of the system indicates the optimal combination rule.

2.3 Cognitive conflict

2.3.1 Defining cognitive conflict - a triple-system expansion of the lens model

Figure 2.2 presents a triple-system expansion of the lens model representing situations of interpersonal interaction (Brehmer, 1979). The lens model equation (i.e., equation 1) then becomes

\[ r_A = G R_{S1} R_{S2} \]  

(2)

5. Much of this section is based on Brehmer (1976, 1979), Hammond et al. (1966) and Rohrbaugh (1988).
where

\( r_A \) is the correlation between the judgments made by person \( S1 \) and those made by person \( S2 \). It indicates the level of agreement between \( S1 \) and \( S2 \).

\( G \) is the correlation between the linearly predictable variance in \( S1 \)'s judgment and \( S2 \)'s judgment. It denotes the similarity of their respective decision structures (i.e., policies), i.e., the extent to which the systematic features of their policies are similar.

\( R_{S1} \) and \( R_{S2} \) are the multiple correlations between the cues and judgments made by \( S1 \) and \( S2 \) respectively. They demonstrate the consistency of each subject's policies.

Cognitive conflict is defined as the disagreement between the subjects, i.e., when \( r_A \) takes on a value of less than unity. In equation 2, two possible sources of disagreement are identified. The first of these sources is a systematic difference between the policies of \( S1 \) and \( S2 \) such as, differences in weights, function forms or organizing principles. Such differences are revealed by the matching index, \( G \).

The second source of disagreement is the lack of consistency. Even if \( G \) equals 1, \( r_A \) will fall below unity when one or both of the persons involved is inconsistent (i.e., \( R_{S1} \) and/or \( R_{S2} \) is below unity). That is, two persons may give the same relative weights to the cues, employ the same functions, and use the same organizing principle but still fail to agree if they do not apply these weights, functions and principles in a consistent way. Thus, it is clear that the mere existence of disagreement does not necessarily imply fundamental differences in policy: inconsistency may well be another source.
Figure 2.2

Lens Model – Three System Expansion

Criterion Event
\[ = Y_e \]

Subject 1's Response
\[ S_1 \]

Predicted Response
\[ S'_1 \]

Predicted Response
\[ S'_2 \]

Subject 2's Response
\[ S_2 \]

\[ S_1 \text{ Consistency} \]
\[ R_{s1} \]

Matching Index
\[ G \]

\[ S_2 \text{ Consistency} \]
\[ R_{s2} \]

Agreement Between Ss
\[ r_A = GR_{s1} R_{s2} \]
In equation 2, cognitive conflict is conceptualized as a function of two particular attributes of human information processing, supported by evidence from several studies. First, humans lack insight into their judgment policies. Several studies have found a low correspondence between self-reported subjective policy weights and statistical weights (e.g., Hoffman, 1960; Slovic, 1969)\(^6\), suggesting that an individual is poor at understanding and implementing his/her judgment policy (Balke et al., 1973). Thus, two subjects, engaged in formulating a joint judgment, are not likely to have a precise notion of the extent to which their respective judgment policies (indicated by \(G\)) are similar or different (thereby implying a \(G\) of less than 1).

The second element in the formulation is the role of inconsistency. Another salient finding in studies in human information processing (particularly the "man versus model of man" literature) is that humans are inconsistent in applying their own judgment policies (Bowman, 1963; Hammond and Summers, 1972; Dawes and Corrigan, 1974; Hogarth, 1986, p. 193-195)\(^7\). Thus, one or more individuals in the joint judgment may be inconsistent (thereby causing \(R_{Si}\) to be less than unity).

In summary, the existence of interest differentials (or strategic motives) among the players is not a necessary condition for conflict to exist (Hammond, 1965; Brehmer, 1976; Rohrbaugh, 1988). Conflict

\(^6\) See, however, Reilly and Doherty (1989) for some evidence to the contrary.

\(^7\) Libby (1976) provides some conflicting evidence on the matter. See Goldberg (1976) for a very different interpretation of Libby's results.
can be created by characteristics of human information processing alone, and such conflict is referred to as cognitive conflict.

2.3.2 The cognitive conflict paradigm

The SJT paradigm for the study of cognitive conflict was developed by Hammond (1965). Experiments in this paradigm simulate a situation wherein two or more persons work together to solve a series of problems requiring them to make inferences from a set of cues. However, each person has a different mental model of the problem, and therefore, a different solution. The differences in mental models among the subjects stem from the fact that they use (the same set of) cues differently. Experiments in this paradigm have two distinct parts. The first part ensures that cognitive conflict is created within the group. The second part is the actual experimental stage, wherein groups in cognitive conflict are observed.

Creating Cognitive Conflict

In laboratory studies, differences in policies are usually created by training subjects to think differently about a given task. This is done through a multiple-cue probability learning (MCPL) exercise wherein a subject is administered a series of trials (Brehmer, 1980). In each trial, the subject is required to estimate the criterion value from a set of cue values, and is then prompted with the "actual" criterion value. A subject is induced a policy over a series of trials. Disagreement is created by inducing each subject to learn a different policy. Thus, cognitive conflict results from the fact that each subject (within a group) is trained differently.
Experimental Stage

In the second (i.e., experimental) stage, subjects are brought together and told that they have now learned their policy and are instructed to cooperate in solving a series of problems. When disagreement occurs, subjects discuss the case until they can reach a joint judgment agreeable to both of them. They are not informed of the fact that they have been trained differently. On each trial in the conflict stage, the subjects (a) study the cue values, (b) make individual judgments of the criterion variable, which (c) they communicate to each other, and if these judgments do not agree, they (d) discuss the case until they reach a joint judgment acceptable to both of them.

The paradigm is thus designed to evoke a very benign form of conflict. The subjects have no differences in interest, and no subject can make gains at the expense of the other. Their only concern is to find the correct answer.

2.3.3 Experimental studies of conflict

Brehmer (1976) provides a comprehensive survey of cognitive conflict studies done in the SJT paradigm. Such studies have concentrated on three major aspects of cognitive conflict: (a) the structure of cognitive conflicts (i.e., the relative importance of consistency and systematic differences in policy), (b) the effects of task characteristics, and (c) the effects of subject characteristics. Each aspect is reviewed in turn.
The Structure of Cognitive Conflicts

Early studies of policy conflict (e.g., Hammond, Todd, Wilkins and Mitchell, 1966) employed two-cue inference tasks. One cue was related to the criterion in a linear fashion, whereas the other was nonlinear. In the initial training, one subject in each pair was trained to rely on the first cue, and the other subject was trained to use the second cue. Since the cues were uncorrelated, the fact that different subjects relied on different cues led them to make different judgments. A principal result of these studies was that there was little or no reduction in overt judgmental differences (\(rA\) in equation 2 above) over a 20-trial sequence.

Brehmer (cf. 1976), who used the same task and procedure as the earlier studies, analyzed the subjects' estimates by means of the lens model equation (equation 2) to yield separate measures of policy similarity (\(G\)) and policy consistency (\(RSl\)). He found a rapid reduction in systematic differences (i.e., \(G\)) between the subjects' policies. At the same time, however, the consistency of their policies (i.e., \(RSl\)) decreased, thereby accounting for a larger proportion of the subjects' disagreement than did the systematic differences in policy. Thus, while there was little change in the amount of disagreement, the structure changed over time from systematic differences to inconsistency.

These results have been replicated over a wide variety of conditions, including subject characteristics such as sex (Hammond and Brehmer, 1973), and task conditions, both in task content (Hammond and Brehmer, 1973) and the formal characteristics of tasks (Brehmer, 1976).
The change in the structure of the disagreement thus seems to be a stable feature of policy conflict as studied in this paradigm. Brehmer (1976) attributes this finding to the manner in which subjects change their policies:

To maintain the consistency of their policies, subjects have to increase their dependency on new cues that they learn to use at the conflict stage (i.e., cues used by their opponents) at the same rate as they decrease their dependency on the cues they used initially. The subjects fail to do this, however. Instead, they decrease their dependency on the cues used initially at a faster rate than they learn to use new cues. This necessarily leads to a decrease in consistency, as can be seen from the fact that the multiple correlation between cues and judgments, which defines the subjects' consistency, is directly dependent on the sum of the correlations between the individual cues and the judgments, which define the dependency on the cues (p. 991).

**Effect of Task Characteristics**

The characteristics of the cognitive task have important effects on the structure of the inductive policies developed to cope with the task. Experimental results are summarized below.

* **Function form:** The function form factor affects the consistency of subjects' policies, but not the level of policy similarity (Brehmer, 1973). For nonlinear tasks in which the subjects develop nonlinear policies, the consistency of the policies is lower than for linear tasks, in which subjects develop linear policies (Hammond and Summers, 1972).

Studies examining differences in the formal characteristics of task structure report that subjects with non-linear policies give up more of their initial policies than subjects with linear policies (Brehmer, 1971, 1972).
* Relative weights: When relations between cues and criterion are linear, the distribution of cue validity does not affect the level of policy similarity (Brehmer, 1974a). However, for tasks containing linear and nonlinear cues, policy similarity is higher when the linear cue is more valid than the nonlinear cue. The cue validity distribution factor does, however, have a clear effect on the consistency of the subjects' policies. Consistency is higher when subjects use only one of the cues in the task than when they have to use multiple cues (Brehmer, 1974a).

* Cue Intercorrelations: Studies examining cue intercorrelations have found two interesting results. First, if the cues are positively intercorrelated, the judgments of two policies are similar, even if they rely on different cues. Consequently, cue intercorrelations introduce agreement in fact, despite disagreement in principle and thus hide the true nature of their differences from the subjects (Brehmer, 1975). Second, cue intercorrelations permit the subjects to have a high level of achievement even though they rely on cues that are themselves, invalid (Brehmer, 1975).

* Task predictability: Studies of learning show that the consistency of subjects' policies is a direct function of the predictability of the task (Brehmer, 1970; Dudycha and Naylor, 1966). This conclusion appears to hold in cognitive conflict studies as well, in that (a) subjects are able to reach the same level of policy similarity regardless of the level of the task, and that (b) variations in subject consistency are explained by the predictability of the task (Brehmer, 1973, 1975).
Thus, the characteristics of the task are important in two ways. First, they determine the structure of the subjects' policies. Tasks that require multiple cues, are non-linear and highly uncertain, lead to lower consistency (and consequently, higher conflict). Second, cue intercorrelations appear to directly impact policy change. Therefore, tasks with high cue intercorrelations lead to less policy change than tasks with orthogonal cue structures.

**Effect of Subject Characteristics**

Two kinds of subject characteristics are of interest: (a) the conventional individual difference variables such as personality and sex, and (b) the formal characteristics of the subjects' policies. There are no studies involving personality differences. Studies focusing on sex (Hammond and Brehmer, 1973) and nationality (Brehmer, 1974b) indicate that these factors do not affect the amount or structure of conflict. This is expected, given that conflict, as defined by the paradigm, mainly depends on the subjects' ability to adapt to the task. However, it should be noted that in prisoners' dilemma experiments, the sex of subjects has been shown to affect conflict (Pruitt and Kimmel, 1977).

2.3.4 Cognitive Conflict Research: Findings

**Conclusions**

Three definitive conclusions can be drawn from cognitive conflict studies. First, the structure of conflict changes as subjects interact with each other and with the task. The initial conflict results from systematic
differences in policies, whereas the persistence of conflict is due to inconsistencies on the part of subjects.

Second, these studies suggest that it may be useful to look for cognitive explanations of conflict (such as, the effect of task structure), especially in situations where there is no apparent reason for interest differentials to play a role. Even in the presence of mixed motives, cognitive reasons may still account for part of the conflict (Balke et al., 1973).

A third important conclusion points to the role of inconsistency in constructing a general theory of behavior. Studies of individual decision making, particularly the "man versus model of man" and cognitive control studies, have documented the role of inconsistency in individual decision making. Cognitive conflict studies add another dimension in clarifying the role of inconsistency in interpersonal interaction.

Conceptual Limitations of the Paradigm

Conceptually, the first question that must be asked is: how realistic is the paradigm in studying real life conflict situations? In a study of a labor-management dispute, Balke et al. (1973) found that representatives of the two sides were negotiating on what they thought were the most critical issues: duration of the employment contract, annualized wage increases, number and use of machine operators and the number of strikers to be recalled as part of the settlement. Analyses of the participants' policies (from a cognitive conflict viewpoint) revealed that only two of the four items were actually critical. Only the wage issue was important to management and only the
question of recall was important to the union. Policy differences among the management representatives themselves were wide, a fact they were not aware of. There was also a low degree of self-consistency among participants in deciding negotiating positions. The study of the Denver ammunition case (Hammond and Adelman, 1976) draws similar conclusions.

These (and other field studies) substantiate the general premise that cognitive conflict is a component of conflicts in general, even in situations where the parties involved have clear and strong interest differentials. However, neither the experimental research nor the field research provides any indication as to how and to what extent cognitive factors (versus interest differentials) shape a given conflict situation. In other words, available evidence does not allow us to specify the relation between the two types of conflict with any degree of precision. Results of cognitive conflict studies should, therefore, be used with caution and in situations where evidence clearly indicates that the conflict is driven largely by cognitive factors. Arguably, organizational analogs of such situations are abundant. However, the presence of interest differentials would require a somewhat different analysis.

A second important conceptual question relates to the manner in which cognitive conflict is defined. In the context of the interpersonal conflict paradigm, cognitive conflict is essentially viewed as a problem of information integration, given a particular set of cues. The problem can also be conceived, in a complementary sense, as one of information acquisition: cognitive conflict arising from the use of
different cues by different subjects. This is a realistic scenario: different individuals performing the same task are known to form mental models consisting of different cues (Gentner and Stevens, 1983). It is also implied by the notion of vicarious functioning - the intersubstitutability of cues by different individuals, all performing the same task. Redefining cognitive conflict by including the information acquisition aspect would undoubtedly enrich the scope of explanation offered by the paradigm. This particular formulation of the problem, however, has not been addressed in cognitive conflict research.

Finally, how does the nature of conflict evolve in the absence of outcome information (i.e., information about judgmental accuracy)? The non-outcome information case has not been sufficiently well investigated for definitive conclusions to be formed (Brehmer, 1976). The absence of outcome information is, however, a common characteristic of individual and group decisions in a variety of non-laboratory situations (Hogarth, 1986, p. 222; Ashton, 1985). Studies of information processing indicate that providing outcome information decreases consistency in individuals (Schmitt, Coyle and King, 1976; Arkes, Dawes and Christensen, 1986). Since consistency ($R_{ij}$ in equation 2 above) appears to be the cornerstone of cognitive conflict explanations, the following question would therefore be interesting: is the nature of cognitive conflict different in the absence of outcome information? This issue is important and should be studied.

Methodological Limitations

Methodologically, the main issue is group size. With the exception of Cvetkovich (1973) who used groups of five, experiments
within the cognitive conflict paradigm have been conducted with dyads. Group studies typically involve groups of three or more (Hill, 1982). Evidence also indicates that the size of the group affects the nature and quality of the group decision (viz, Einhorn et al., 1977; Bray, Kerr and Atkin, 1978). It is an open question as to whether the cognitive conflict findings hold for dyads or for small groups in general.

2.4 Decision Feedback

2.4.1 Feedback as a general concept

Wiener (1948, 1950) is credited with popularizing the concept of feedback (cf. Nadler, 1979). At its most basic level, feedback is information received by an individual about his or her past behavior (Annett, 1969). It provides some information about the correctness or adequacy of a response (Bourne, 1966). Thus, feedback is the process by which an environment returns to individuals a portion of the information in their response output necessary to compare their present strategy with a representation of an ideal strategy (Doherty and Balzer, 1988). Feedback is crucial to behavior (Norman, 1981) in that it enables individuals to understand and improve their judgments, improve their expertise in a judgment task, and reduce their commitment to incorrect judgment strategies (Hogarth, 1981).

Research on individual feedback is vast and diverse, and is surveyed in Annett (1969). In general, individual feedback can be understood in the larger context of information processing (Hogarth, 1986, pp. 206-207), wherein the decision maker attends to the
feedback, perceives / interprets the feedback, integrates elements of
the feedback, and reacts to the feedback.

The use of feedback in groups is based on the view that
receiving feedback is a potentially significant event in the life of a group
performing work (Bowers and Franklin, 1976). Through feedback, a
group may obtain information about the quantity and quality of its
output as well as knowledge about the effectiveness of the methods
used to achieve desired levels of performance (Nadler, 1979).

2.4.2 Individual versus Group feedback: the Referent Level

Nadler (1979) points out that the referent level to which
feedback is provided, has important implications for the effectiveness
of the feedback. Thus, feedback provided to groups is intrinsically
different from feedback to individuals:

Feedback received by individuals in a group may be very
different than feedback received individually because that
information may be confounded by the actions of other
group members. Second, even if the data can be
interpreted and understood, the individual is limited by
the inherent nature of the group in his or her ability to act
on that information (p. 312).

Nadler (1979) suggests that the referent level be incorporated in
the design and analysis of feedback by envisaging the delivery of
feedback at three referent levels:
(a) Individual feedback, i.e., information provided to an individual
regarding his performance with no intervening social factors.
Therefore, both the unit of decision making and the feedback pertain to
the decision processes of the same individual. This is the typical format of analysis in research on decision feedback.

(b) Individual feedback in groups (or interpersonal feedback), where the individual's behavior occurs in a group setting. Here, the unit of decision making is still the individual, the crucial difference being that feedback pertains to the decision processes of the other individuals in the group. The introduction of interpersonal feedback has important implications for individual information processing.

(c) Group feedback, where the information reflects the functioning of the group as a unit. Here, both the unit of decision making and the feedback pertains to the group as an entity. It should be clarified that the group being referred to is an actual group and not any form of aggregation of individual decisions. The use of the group as the referent level raises two issues. First, which of the findings from individual feedback research can legitimately be transported to the domain of the group, and second, what are the additional issues that arise in shifting the level of analysis? These are further discussed in section 2.6

The Effect of Feedback at Different Referent Levels

Studies that have examined the linkage between feedback and the referent level (reviewed by Nadler, 1979) differ in the type of task, size of group, nature of feedback and dependent variable. This makes the results somewhat ambiguous and difficult to interpret. Nevertheless, one conclusion appears to predominate - performances

---

8. This point is elaborated in section 2.6.3.
seem to be best when individuals have feedback about their own performance, that of other group members and that of the group as a whole (Zajonc, 1962; Zander and Wolfe, 1964; Matsui, Kakuyama and Uy Onglatco, 1987).

2.4.3 Two Conceptions of Feedback - Cognitive and Outcome

Decision feedback is often characterized as being of two types - cognitive or process feedback (the two words are used interchangeably) and outcome feedback (Hammond et al., 1975; Hogarth, 1981). Outcome feedback is information about the accuracy or correctness of the response. Cognitive feedback represents information regarding the how or the why that underlies this accuracy. This distinction can be further clarified by using the concepts of predictive and explanatory value of information (Jacoby et al. 1984):

Predictive value refers to the correlation between accurate information regarding one's past performance and the ability of this information to predict future levels of performance. In contrast, explanatory value represents information regarding why certain relationships occurred. This may or may not be accompanied by the ability to predict if (or when) these relationships will occur in the future, that is, predictive and explanatory value are conceptually independent (pp. 531-2)

Studies in multiple-cue probability learning have demonstrated that while subjects clearly learn from outcome feedback in simple (i.e., two-cue, linear) tasks (O'Connor, Doherty and Tweney, 1989), learning from outcome feedback in complex, uncertain tasks does not happen (Brehmer, 1980; Hoffman, Earle and Slovic, 1981), especially for tasks with low predictive value (Hammond and Summers, 1972). Furthermore, providing outcome feedback in such tasks reduces
response consistency (Schmitt et al., 1976; Arkes et. al., 1986), causing accuracy to decrease (Hammond, Summers and Deane, 1973; Schmitt et al., 1976). The detrimental effects of outcome feedback have persisted even when subjects are told that the task is probabilistic and they should not expect to be right every time (Brehmer and Kuylenstierna, 1978). Experienced decision makers seem to understand this. In a study of how security analysts use information, Jacoby et al. (1984) found that in contrast to poorer performers, better performing decision makers were more likely to ignore outcome-only feedback.

2.5 Cognitive feedback

2.5.1 Defining Cognitive Feedback within the Lens Model Framework

Cognitive feedback has been studied extensively, mainly through operationalizations of the lens model. According to Doherty and Balzer (1988), cognitive feedback provides a subject with information describing the following relationships (notations refer to table 2.1 and figure 2.1):

(a) Between cues and the criterion \((Y_e)\), i.e., information about the task. This is known as task information. Three kinds of relational indices characterize task information - overall task predictability \((R_e)\), cue intercorrelations \((r_{ij})\), and correlations between cues and the

\[9. \text{ Much of this section (particularly sections 2.5.1 and 2.5.3, and tables 2.3 -2.5) is based on material derived from Balzer, Doherty and O'Connor (1989) and an earlier draft of the paper. Dr. Michael E. Doherty provided valuable guidance in enhancing the researcher's understanding of cognitive feedback.} \]
criterion \((r_{\theta})\). Task information can also be information about another person, as in the cognitive conflict paradigm (i.e., equation 2 - section 2.3.1).

(b) *Between cues and the person's inferences*, i.e., information about the person's cognitive state. This is known as cognitive information. This component, as implied by the lens model, largely mirrors task information (the only difference being that there is no equivalent for cue intercorrelations, or \(r_{ij}\)).

(c) *Between cognitions and the distal objects*. This third category comprises indices of *functional validity information*, or, information about the relation of the cognitive system to the task system (Balzer et al., 1989). These indices include the achievement correlation \((r_{\theta})\), the matching index \((G)\), and the correlation between the residuals from the predictions of those models \((C)\).

Table 2.2 summarizes descriptions of various forms of cognitive feedback.
TABLE 2.2: Components of Cognitive Feedback
(Doherty and Balzer, 1988)

<table>
<thead>
<tr>
<th>Component</th>
<th>Representation</th>
<th>Measure</th>
</tr>
</thead>
<tbody>
<tr>
<td>TI</td>
<td>Task Predictability</td>
<td>$R_{a}$</td>
</tr>
<tr>
<td></td>
<td>Ecological Validity</td>
<td>$r_{ie}$</td>
</tr>
<tr>
<td></td>
<td>Cue Intercorrelation</td>
<td>$r_{ij}$</td>
</tr>
<tr>
<td>CI</td>
<td>Subject Consistency</td>
<td>$R_{S}$</td>
</tr>
<tr>
<td></td>
<td>Cue Utilization</td>
<td>$r_{is}$</td>
</tr>
<tr>
<td>FVI</td>
<td>Achievement</td>
<td>$r_{2}$</td>
</tr>
<tr>
<td></td>
<td>Matching Index</td>
<td>$G$</td>
</tr>
</tbody>
</table>

Note:
CI: Cognitive information
TI: Task information
FVI: Functional validity information

The Relation between Task Information and Cognitive Information

Insofar as they purport to depict mirror images of a decision situation, cognitive information and task information bear some resemblance. However, the respective interpretations (and therefore effects) of cognitive and task information are not necessarily the same. Human decision making is often studied in terms of a two-step model wherein the individual selects, and then implements, a decision strategy (Dudycha and Naylor, 1966; Hammond and Summers, 1972). Within this framework, cognitive information putatively aids the decision maker by providing feedback on the decision maker's policy structure (policy feedback), as well as, feedback on the consistency with which the decision strategy is being implemented (consistency feedback) (Hammond and Brehmer, 1973). The components of task information, on the other hand, are task structure and task predictability. While policy structure and task structure bear a close analogy, the notions of
task predictability and the decision maker's consistency are entirely
different. Consistency is the manifestation of a person's cognitive
control (Hammond and Summers, 1972), which is a characteristic of
human information processing. Task predictability, on the other hand,
is caused by environmental uncertainty.

2.5.2 Operationalizations of Cognitive Feedback

Each index in table 2.2 can be operationalized as a number of
statistical measures and can be presented in a variety of forms. The
discussion concentrates on presentation formats.

Cognitive feedback is provided in three formats - graphically,
verbally and numerically (Balzer et al., 1989). Graphic feedback is
rendered in different combinations of the following:
(a) visual relationships between task outcomes and a subject's
judgments (Hammond and Boyle, 1971),
(b) bar graphs representing the subject's ecological validities and
utilization coefficients (Balke et al., 1973),
(c) a best fit curve representing the subject's function form and the
scatter around the resulting line (Hammond, 1971), and
(d) visual comparisons between the subject's policy weights and the
task (Hammond, 1971).

Purely verbal descriptions of weights and function forms have
been used by, e.g., Deane, Hammond and Summers (1972). In that
study, subjects in one experimental condition were given verbal
descriptions of environmental weights and function forms, while other
subjects received information pictorially.
Some researchers have presented cognitive feedback numerically, by providing statistical measures. For example, Newton (1965) presented all feedback in the form of correlation coefficients.

2.5.3 The Effectiveness of Cognitive Feedback

Mode of Comparison

Tables 2.3, 2.4 and 2.5 summarize overall results of effectiveness studies. This summary is meant to be illustrative, and as such, is not exhaustive. Overall results are split by dependent variable used (matching or knowledge, achievement and control, respectively), type of feedback (task information and cognitive information or task information only) and comparison made (cognitive feedback versus outcome feedback or cognitive feedback versus no feedback).

Overall Results

Tables 2.3 - 2.5 indicate a consistent overall pattern of results. We draw three conclusions. First, for all three dependent variables, cognitive feedback appears to be more effective than the other conditions (i.e., outcome feedback or no feedback). Second, there is no incremental performance differential in providing task information and cognitive information, versus providing task information only (e.g., Nystedt and Magnusson, 1973; Steinmann, 1974. Hammond and Boyle, 1971, is an exception). Finally, due to lack of evidence, no

---

10. For a comprehensive review of studies examining the effectiveness of cognitive feedback, see Balzer et al. (1989).
definitive conclusions can be drawn regarding the effectiveness of cognitive information alone.

**TABLE 2.3 (a): Effectiveness of cognitive feedback: Knowledge**  
*CFB studies using CI and TI*

<table>
<thead>
<tr>
<th>Comparison</th>
<th>Result</th>
<th>Studies</th>
</tr>
</thead>
<tbody>
<tr>
<td>CFB vs OFB</td>
<td>CFB&gt;OFB</td>
<td>Hoffman et al. (1981); Lindell (1976);</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Nystedt and Magnusson (1973); Steinmann (1974)</td>
</tr>
<tr>
<td>CFB vs NFB</td>
<td>CFB&gt;NFB</td>
<td>Fero (1975)</td>
</tr>
<tr>
<td></td>
<td>n.s.</td>
<td>Schmitt et al. (1977)</td>
</tr>
</tbody>
</table>

**TABLE 2.3 (b): Effectiveness of cognitive feedback: Knowledge**  
*CFB studies using TI only*

<table>
<thead>
<tr>
<th>Comparison</th>
<th>Result</th>
<th>Study</th>
</tr>
</thead>
<tbody>
<tr>
<td>CFB vs NFB</td>
<td>(TI&gt;no TI)</td>
<td>Schmitt et al. (1976)</td>
</tr>
</tbody>
</table>

**TABLE 2.4 (a): Effectiveness of cognitive feedback: Achievement**  
*CFB studies using TI and CI*

<table>
<thead>
<tr>
<th>Comparison</th>
<th>Result</th>
<th>Study</th>
</tr>
</thead>
<tbody>
<tr>
<td>CFB vs OFB</td>
<td>CFB&gt;OFB</td>
<td>Hammond and Boyle (1971); Hoffman (1981);</td>
</tr>
<tr>
<td></td>
<td>n.s</td>
<td>Schmitt et al. (1977)</td>
</tr>
<tr>
<td>CFB vs NFB</td>
<td>(TI&gt;no TI)</td>
<td>Balke et al. (1973); Fero (1975)</td>
</tr>
</tbody>
</table>
**TABLE 2.4 (b): Effectiveness of cognitive feedback: Achievement**  
*CFB studies using TI only*

<table>
<thead>
<tr>
<th>Comparison</th>
<th>Result</th>
<th>Study</th>
</tr>
</thead>
<tbody>
<tr>
<td>CFB vs OFB</td>
<td>No studies</td>
<td></td>
</tr>
<tr>
<td>CFB vs NFB</td>
<td>n.s.</td>
<td>Schmitt et al. (1976)</td>
</tr>
</tbody>
</table>

**TABLE 2.5 (a): Effectiveness of cognitive feedback: Control**  
*CFB studies using TI and CI*

<table>
<thead>
<tr>
<th>Comparison</th>
<th>Result</th>
<th>Study</th>
</tr>
</thead>
<tbody>
<tr>
<td>CFB vs OFB</td>
<td>CFB&gt;OFB</td>
<td>Adelman (1981), Hoffman et al. (1981), Schmitt et al. (1977)</td>
</tr>
<tr>
<td>CFB vs NFB</td>
<td>(TI&gt;no TI) n.s</td>
<td>Balke et al. (1973); Fero (1975) Schmitt et al. (1976).</td>
</tr>
</tbody>
</table>

**TABLE 2.5 (b): Effectiveness of cognitive feedback: Control**  
*CFB studies using TI only*

<table>
<thead>
<tr>
<th>Comparison</th>
<th>Result</th>
<th>Study</th>
</tr>
</thead>
<tbody>
<tr>
<td>CFB vs OFB</td>
<td>CFB&gt;OFB</td>
<td>Hammond and Boyle (1971), Nystedt and Magnusson (1973)</td>
</tr>
<tr>
<td>n.s</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Sample Sizes

Sample sizes of the studies reported vary widely. While many studies have used large samples (e.g., Hoffman et al., 1981 - N = 182; Lindell, 1976 - N = 90), several studies have notably small sample sizes (e.g., Nystedt and Magnusson, 1973 - N = 12; Balke et al., 1973 - N = 6). The fact that results seem to hold across different sample sizes perhaps affirms their robustness.

2.5.4 Cognitive feedback research: Issues

Problems in Comparison

Measuring the effectiveness of cognitive feedback poses a major problem. Most studies have used some combination of cognitive information and task information. A few have used task information only. Except for Flack and Summers (1971 - N = 2), there does not appear to be any study using cognitive information alone. Therefore, while existing evidence allows us to make comparisons of cognitive feedback versus outcome feedback or cognitive feedback versus no feedback, the relative efficacies of different forms of cognitive feedback (or combinations thereof) cannot be gauged.

Operationalization of Feedback

As indicated above, a variety of operationalizations (i.e., formats and/or statistical measures) have been used in these studies. Studies using format of presentation (e.g., Deane et al., 1972) as a variable have not found any effect resulting from presentation format. However, Balzer et al. (1989) point out that since cognitive feedback is used to provide a "mental model" of an individual's judgment policy, different
operationalizations of such feedback may produce different mental models. This makes comparison difficult.

Task Information versus Cognitive Information

According to Doherty and Balzer (1988), since providing cognitive information along with task information does not cause performance improvements, providing task information alone should suffice. This conclusion is open to question. It has been argued above that the feedback contained in task information and cognitive information can be interpreted differently. Second, a fundamental conclusion from studies of human information processing - that subjects lack self insight into their own judgment policy - is a strong rationale for providing cognitive information. Given a strong premise and the absence of studies providing cognitive information only (with the exception of Flack and Summers, 1971), it is not appropriate to conclude that cognitive information is not effective. Finally, providing task information requires knowledge of task outcomes. However, in numerous decision situations, outcomes are not available with any degree of accuracy or are irrelevant (Hogarth, 1986, p. 222; Ashton, 1985). In such situations, cognitive information is really the only form of feedback available. Therefore, lack of evidence about the effectiveness of cognitive information is a significant gap in the literature.

Task Content

Cognitive feedback studies implicitly assume that the formal characteristics of the task are far more important than task content, thereby implying that cognitive feedback methods generalize across
contents (Balzer et al., 1989). A study that did manipulate task content (Adelman, 1981) found significant effects arising from task content. This causes some concern about the assumption, especially since cognitive feedback studies have used such a wide variety of task contents, e.g., predicting GPAs (Schmitt, Coyle and Saari, 1977) hypothetical medical diagnosis (Brehmer and Hagafors, 1986), and predicting the order of finish in horse races (Rohrbaugh, 1981).

**Subjects' Understanding of Statistical Concepts**

Research by Fong, Krantz and Nisbett (1986) demonstrates that training in statistics has an effect on how people reason about problems. In particular, (a) purely abstract forms of statistical training affect reasoning about concrete events, and (b) even when statistical rules are learnt in particular content domains, they may be abstracted from those domains to a very great degree, sufficient to allow their application to quite different domains (Holland, Holyoak, Nisbett and Thagard, 1986). Since cognitive feedback is primarily statistical information, training in statistics is a factor that should affect a subject's ability to comprehend and use the feedback. It is not clear from the literature whether this factor has been taken into account.

Published cognitive feedback studies usually do not indicate if subjects were trained on how to use statistical measures or whether any self-reports were obtained on their use of the feedback provided. If training was not given, it must be asked if subjects actually (a) understood the statistical concepts underlying the feedback and/or (b) whether they used them appropriately.
Timing of Feedback

Studies of feedback have shown that the timing of feedback has performance implications in terms of effectiveness versus cognitive resources (Wickens, 1984). Immediate feedback is "costly", requiring more cognitive resources. Delayed feedback, on the other hand, is less effective. However, studies in cognitive feedback have not considered this factor. Experiments in cognitive feedback have administered feedback over periods ranging from immediate feedback (Hoffman et al., 1981) to a week following subjects' judgments (Schmitt et al., 1976). This is an important gap in the literature, one with explicit design implications.

Referent level of Feedback

Studies in cognitive feedback have concentrated on the individual, and have not explicitly addressed the issue of the referent level. This limits the usefulness of existing research in providing

---

11. Some studies (e.g., Rohrbaugh, 1979, 1981) have used cognitive feedback in comparing the relative efficacies of social judgment analysis (SJA) versus the Delphi and the Nominal Group Technique. In these studies, individuals in SJA groups were provided with graphic descriptions of their respective policy structures at the start of the experiment. However, because of the following reasons, these studies do not provide specific spin-offs for our purposes:
(i) These studies did not consider the lens model variables (consistency, knowledge, agreement), and therefore are not particularly helpful in extending findings on cognitive feedback research to the group level.
(ii) Comparison groups were put through structured processes (i.e., Delphi or Nominal Group Technique) and thus were not really control groups in the conventional sense of the term.
(iii) Feedback was provided only at the beginning of the session. It is reasonable to expect the participants would have altered their respective policies during the course of the session (section 2.3.3). However, they did not have the benefit of feedback on their revised policies during the course of the experiments.
design guidelines for cognitive feedback at various referent levels. This gap in the literature needs to be addressed in future research.

2.6 Cognitive feedback and the underlying referent level

This section sketches out a basis for including the referent level in the study of cognitive feedback (or more specifically, issues involved in providing feedback at the interpersonal and group levels). It is, in a sense, an integration of issues discussed in sections 2.3, 2.4 and 2.5. The essence of the discussion here is captured in two questions raised in section 2.4.2. The first is to determine the specific conclusions from individual feedback research that can be meaningfully transported to the domain of interpersonal and group research. The second issue is to enunciate additional factors that arise in extending the unit of analysis to the interpersonal and group levels.

The arguments are made in three parts. Section 2.6.1 demonstrates the plausibility of the premise that findings from research on cognitive conflict can be extended to groups as such. Section 2.6.2 examines the implications of doing so in terms of making specific predictions. Section 2.6.3 examines the final piece of the puzzle: providing feedback at the interpersonal level.

2.6.1 The plausibility of extending findings from Individual Literature

The validity of extrapolating findings from the individual literature to the group level rests on the assumption that the underlying
mechanisms involved in both are at least comparable. This argument may be questioned on the grounds that while individuals are constrained in their performance by cognitive limitations, groups do not have such shortcomings. Group performance, however, is bounded by "process losses" (Steiner, 1972). Process losses may occur at various stages of the decision making process (cf. Chalos and Pickard, 1985) and in ways remarkably similar to the individual decision making process, e.g., in the aggregation of information (Tuckman and Lorge, 1962), in the integration of information (Howell, Gettys, Martin, Nawrocki and Johnston, 1970), or at the decision making / choice stage (Laughlin and Branch, 1972). Thus, individual and group decision processes can indeed be considered analogous, especially in studying (a) the respective decision sequences, and (b) consequences arising from deficiencies in information processing capabilities.

2.6.2 Cognitive Feedback at the Group Level

The validity of the point made above is echoed in a demonstrated lack of optimality in (decision) performance that characterizes not only individuals, but groups as well. It is this lack of optimality that provides the rationale for using decision aids (such as

12. While decision making at the individual and group levels has received much attention, the issues studied at each level are quite different. At the individual level, the emphasis has largely been on probabilistic inference, heuristics and biases, etc (Einhorn and Hogarth, 1981). Research at the group level, on the other hand, has tended to focus on such topics as group polarization (Myers and Lamm, 1976) and social decision schemes (Davis, 1980). Little research has been done on determining how factors that affect individual decisions work to influence decision making at the group level (Chalos and Pickard, 1985, is an exception). Hence the need for this exercise in a priori extrapolation from the individual to the group level.
cognitive feedback) (cf. Lewis et al., 1983)\textsuperscript{13}. This argument - that
cognitive feedback improves decision quality not only in individuals, but
also in groups\textsuperscript{14} - becomes clearer when we explore its implications in
terms of the lens model indices of decision quality - achievement,
consistency and knowledge.

Achievement

Group research that has focused on achievement (i.e., the level
of accuracy that groups can attain) concludes that while groups
perform better than individuals\textsuperscript{15}, they do not perform in any optimal
sense whatsoever. \textit{Ex post} analyses show that groups usually
outperform the average member, but seldom attain the level of the best
individual judgment (Yetton and Bottger, 1982). \textit{Ex ante} analyses also
demonstrate that best members consistently outperform the group in
decision accuracy (Einhorn et al., 1977; Lorge, Fox, Davitz and
Brenner, 1958)\textsuperscript{16}. This lack of optimality (representing, in a lens model
sense, an achievement index close to unity) provides \textit{a priori}

\textsuperscript{13}. The authors make the point that the use of decision aids is appropriate only in
situations where (a) there is an opportunity cost in making low quality decisions,
and (b) the divergence (from optimality) is related to cue selection or cue
combination (and not due to, say, goal conflict between an individual and an
organization). It is the second factor that we refer to here.

\textsuperscript{14}. Note here that we seek to simply establish the validity of general principle that
providing cognitive feedback across different referent levels is useful. The actual
design and content of feedback would vary across referent levels.

\textsuperscript{15}. According to the interactional theory, groups have greater resources. This
enables them to perform better than individuals (Barnlund, 1959). Others argue
more on the lines of the Central Limit Theorem in asserting that groups reduce
random errors associated with individual judgment and therefore, perform better
(Hogarth, 1978).

\textsuperscript{16}. This lack of performance is attributed to process losses.
justification in extrapolating individual results to the group level: that there is room for improving group performance. In other words, groups receiving cognitive feedback can be expected to perform better (in a lens model sense) than groups not receiving cognitive feedback.

**Consistency**

The notion of random errors described above (Hogarth, 1978) can be interpreted as the lack of consistency in implementing a policy. Groups have a built-in mechanism for reducing inconsistency, and can therefore be expected to be more consistent than individuals. This, however, does not imply that groups already perform at such an increased level of accuracy that they are as consistent as their decision rules (Hogarth, 1978). In other words, the enhanced performance does not reach such a level as to neutralize the effects of cognitive feedback. Therefore, extrapolation from the individual level here is warranted: groups receiving feedback should be more consistent than groups not receiving such feedback.

**Knowledge**

Knowledge (in the lens model sense) is a function of extent to which an individual learns the relation among a set of variables, over multiple instances (Brehmer, 1980). Studies of outcome feedback demonstrate that such learning is problematic in complex, uncertain tasks, and often leads to misspecification of relations. Groups are also prone to a similar phenomenon: superstitious learning (Levitt and March, 1988). Superstitious learning occurs in situations where the

---

17. In a review of MCPL research, Brehmer (1980) concludes that experience with a task does not necessarily lead to learning the relations in the task.
links between actions and outcomes are misspecified (as in probabilistic tasks), often leading to incorrect hypotheses being formed. The mechanisms involved in superstitious learning bear a close resemblance with those associated with MCPL in complex, uncertain tasks and inhibit the formation of knowledge. Thus, we can expect the effect of cognitive feedback on groups be similar to that on individuals: the provision of cognitive feedback should enhance knowledge.

2.6.3 Cognitive Feedback at the Interpersonal Level

Interpersonal Feedback: Consistency

Providing interpersonal information has important implications on an individual's consistency. The argument here draws on the explanation of cognitive conflict offered by Brehmer (1976) and discussed in section 2.3.3. With the introduction of interpersonal information, an individual repeatedly recalibrates his/her judgment policy in the light of the policies of others (which are also being continuously reformulated). Doing so entails changing the person's dependence on cues, and in doing so, the subject loses consistency (Brehmer, 1972).

The role of cognitive feedback in such situations is to enable the individual to retain cognitive control while (a) shaping his/her policy, and at the same time, (b) incorporating information on the policies of others in formulating the policy. Therefore, individuals with cognitive feedback will display significantly greater consistency than individuals without the benefit of such feedback.
Interpersonal Feedback: Matching of Individuals' Policies

An insightful finding of the cognitive conflict literature has been that the structure of the conflict changes over time (sections 2.3.3 and 2.3.4). Cognitive conflict initiates as disagreement over policy (i.e., a low matching index). Over time, however, these policy differences tend to reduce considerably.

With the introduction of cognitive feedback, individuals will recognize their own and others' policies better. In consequence, they can be expected to match their respective policies quicker. Evidence from cognitive conflict research indicates, however, that eventually these policies converge anyway. Thus, while the provision of cognitive feedback would clearly be useful in the beginning, its value over time is not clear. Thus, whether cognitive feedback will create a lasting performance differential in terms of matching, is a matter of conjecture.

Interpersonal Feedback: Agreement among Individuals' Decisions

Equation 2 posits that agreement (among individual decisions) is a function of the two factors discussed above - the respective consistencies of the individuals, and the extent to which their policies are similar. A priori, we expect cognitive feedback to enhance individual consistency, and increase (or at the very least, sustain) matching among the individuals' policies. The net effect of cognitive feedback, therefore, will be to increase agreement among individuals.

18. The basis for this argument is that individuals demonstrate a lack of self-insight into their judgment policies.
2.7 Framing the Research Question

The framework sketched above can be seen as establishing the *beginning*, or point of departure, for defining the research problem. In summary, we explore the impact of cognitive feedback on group decision making. In doing so, we examine the following issues:

* **Cognitive feedback at multiple referent levels.** Feedback at the individual level has been studied extensively. This study examines feedback in the other two dimensions - interpersonal and group.

* **Cognitive feedback where task outcomes are absent.** Section 2.5.4 concludes that the effect of providing cognitive information *only*, has not been determined in any precise sense. This is important, because, as section 2.3.4 points out, there are numerous tasks where outcome feedback is not easily available or may be misleading or irrelevant. In such instances, cognitive information may be the only feedback available. This study addresses this issue at the interpersonal and group levels.

These, then, are conceptually the principal features of the study. Methodologically, the study addresses/clarifies the following issues that have been lacking in prior research on cognitive conflict and cognitive feedback.

* **The use of small groups.** Section 2.3.4 points out that cognitive conflict research has traditionally been carried out through dyads. We seek to extend the findings to small groups as they are traditionally defined. In this study, we use groups of three members.
* Training subjects on the use of feedback. It is not clear from
cognitive conflict studies whether subjects were trained on how to use
feedback. As section 2.5.4 points out, this acts as a potential confound
on measuring the effectiveness of feedback. This study
operationalizes this issue explicitly by (a) providing training on
feedback at the start of the experimental phase, and (b) verifying that
subjects actually understood the implications of each component of
feedback.\footnote{19}

* The effect of feedback (over time). The impact of feedback is
measured, not only across conditions (feedback versus no feedback),
but over time as well.

The research question is further developed as follows. Chapter
3 translates the question into tangible hypotheses, explains the
experimental design adopted, the design of feedback and the training
of subjects. Chapter 4 reports the results of the experiment. Chapter 5
revisits the important issues addressed in this chapter, and discusses
the implications of the results.

\footnote{19. The appendix (A3.3 and A3.4) provides details on how this was done.}
CHAPTER 3
METHOD

3.1 Scope

This chapter describes an experimental procedure for testing the research question. Section 3.2 presents the hypotheses. Section 3.3 describes the experimental design and its rationale. The choice of subjects for the experiment and the randomization procedure are detailed in section 3.4, and the experimental task is explained in section 3.5. Section 3.6 discusses the design of the feedback mechanism.

As figure 3.1 indicates, the experiment is divided into three phases - the training phase, the experimental phase and a post-experiment debriefing phase. Section 3.7 presents a description, rationale and results of the training phase. The experimental phase is covered in section 3.8. Section 3.9 focuses on the post-experiment debriefing phase, and section 3.10 summarizes the discussion in this section.

3.2 Hypotheses

3.2.1 General Framework for setting hypotheses

Hypotheses are discussed in terms of lens models indices - consistency, matching and agreement\(^1\). For each variable, hypotheses

\(^1\) Details on the calculation of these measures at the individual and group level are provided in sections 4.2.2 and 4.2.3 (chapter 4), respectively.
are formulated at two referent levels - the individual and the group. Hypotheses are usually stated in the null form. However, since this study has several hypotheses, the hypotheses here are formulated in alternate fashion for reasons of clarity. Also, with the exception of \( H_4 \), the reasons underlying the hypotheses have already been discussed in chapter 2. Those arguments are not repeated here, beyond pointing to the relevant sections.

3.2.2 \( H_1 \): Consistency

Section 2.6 postulates that individuals and groups will display enhanced levels of consistency with the benefit of cognitive feedback. This proposition is reflected in \( H_{11} \) and \( H_{13} \). With respect to the time dimension, the argument (that cognitive feedback enables the retention of cognitive control) translates to the following. The provision of feedback will enable subjects to be uniformly consistent over time. On the other hand, subjects not receiving cognitive feedback will not be able sustain a uniform level of consistency (which, as stated above, will be lower than that of those receiving feedback). In statistical terms, we expect this lack similarity between the two experimental conditions to be manifested as a significant interaction between condition (i.e., type) and the time factor (i.e., blocks)\(^1\) - \( H_{12} \) and \( H_{14} \).

\(^1\) An interaction occurs when an independent variable has different effects on a dependent variable at different levels of another independent variable (Kerlinger, 1986, p. 230). Here, we hypothesize the effect of independent variable time on dependent variable consistency to be contingent on another independent variable, type. Hence the reason to expect the interaction.
Individual Level

H_{11}: Individuals receiving cognitive feedback will be more consistent than individuals receiving no feedback.

H_{12}: Individuals receiving feedback will display a uniform level of consistency over time, whereas subjects not receiving feedback will not display a uniform level of consistency.

Group Level

H_{13}: Groups receiving cognitive feedback will be more consistent than groups not receiving cognitive feedback.

H_{14}: Groups receiving feedback will display a uniform level of consistency over time, whereas those not receiving feedback will not display a uniform level of consistency.

3.2.3 H_2: Matching Index^2

Individual Level

The discussion in section 2.6.2 suggests two conclusions about the matching index among individuals. First, there will be a significant difference in the individual matching index (H_{21}) between the two conditions. Second, over time, the individual matching index will increase for both conditions. Thus, in statistical terms, we hypothesize a main (block) effect - H_{22}.

H_{21}: The individual matching index in the cognitive feedback condition will be higher than that in the no-feedback condition.

H_{22}: The individual matching index will increase over time.

---

2. The matching index is formulated at two levels - individual and group. The individual matching index is a measure of how close the respective individual decision models are to each other. The group matching index measures the extent to which the group decision model is close to each individual decision model. To maintain clarity, the respective measures will be described as individual matching index and group matching index.
**Group Level**

We postulate above that with feedback, group members will attain greater convergence in their respective (individual) policies. Subsequent to the individual decisions, the same individuals will formulate a group decision on the same set of cases. It is, therefore, logical to expect that the group policy in feedback groups will be more similar to the respective individual policies than in no-feedback groups (H$_{23}$). Also, drawing an analogy with the individual situation, we predict that over time, the group matching index will increase, thereby implying a block effect (H$_{24}$).

H$_{23}$: Groups getting feedback will have a higher group matching index (i.e., match with the members' individual policies) than groups not receiving feedback.

H$_{24}$: The group matching index will increase over time.

### 3.2.4 H$_3$: Agreement

**Individual Level**

*Per* section 2.6.2, we postulate that the agreement index in the feedback condition will be greater than the agreement index in the no feedback condition (H$_{31}$). Further, since the individual matching index is expected to increase over time (H$_{22}$), the agreement index will also increase over time (equation 2).

H$_{31}$: The agreement index in the feedback condition will be greater than that in the no-feedback condition.

H$_{32}$: The agreement index will increase over time.
Group Level

At the group level, the agreement index is a measure of the extent to which the group decision is similar to the respective individual decisions on the same cases. The hypotheses on agreement are formulated on the lines of those advanced for hypotheses H_{23} and H_{24} respectively.

H_{33}: Groups getting feedback will have a higher agreement index (i.e., agreement with the members' individual decisions) than groups not receiving feedback.

H_{34}: The agreement index will increase over time.

3.2.5 H_{4}: The Number of Iterations

Iterations as a surrogate for cognitive effort

In analyzing the tradeoff between effort and accuracy in judgment, Beach and Mitchell (1978) note the importance of cognitive effort. The use of feedback requires cognitive effort (Wickens, 1984). Therefore, measuring cognitive effort employed by the subjects provides some indication of the use of feedback. In studies of human information processing, decision time is often used as a surrogate measure for cognitive effort (e.g., Jarvenpaa, 1989). This study captured the decision behavior of subjects by recording the number of times they iterated in formulating a decision. This figure denotes the number of times a decision maker seeks to reformulate his/her policy. Being closely tied to the actual decision process, it is a more concrete measure of cognitive effort than decision time.
The premise of this study is that cognitive feedback provides subjects with insights into their policies and consistency. It is, therefore, logical to expect that decision makers with access to feedback will use it to refine their judgment policies to a greater extent than those with no access to feedback. This implies that individuals in the former category will iterate more often than those in the latter category.

$H_{41}$: Individuals getting feedback will iterate more than those getting no feedback.

3.3. Experimental Design

3.3.1 Description of design

The research design is illustrated in table 3.1. The experiment used a factorial design with two components in order to capture the feedback condition and learning. These components are - between-subjects and within-subjects.

Between Subjects

The fundamental objective driving the experiment is to determine the effect of cognitive feedback, i.e., ascertain whether there are any significant differences between feedback and no-feedback conditions. That is the rationale for the between-subject component.

Within Subjects

In addition to ascertaining if systematic differences exist among experimental conditions, feedback studies also seek to study the effect
within each condition over time\textsuperscript{3}. This is referred to as a within-subject design (Barlow and Hersen, 1984, p. 66) and involves multiple measurements over different points of time.

In this experiment, the within subjects aspect was operationalized by using blocks as time intervals. The conceptual basis for using blocks is clarified by Hogarth (1981):

Learning involves the use of feedback to generate, modify, maintain, or abandon hypotheses, a process that occurs across time. In discrete tasks, however, the methodological constraints of technique (e.g., regression analysis) typically force investigators to assume that subjects are applying a single judgmental rule. Thus, ... the learning of judgmental rules is investigated by comparing rules inferred from performance achieved in different blocks of trials as opposed to monitoring behavior over time on a trial-by-trial basis (p. 202).

Learning within the lens model paradigm typically involves providing the subject with the other side of the model, i.e., task information.

When the lens model is extended to study interpersonal situations, providing the subject with information about others' cognitive systems is analogous to providing task information (Doherty and Balzer, 1988). Therefore, there is strong reason to assume some process of learning in such experiments. It is thus important that the experimental design (and subsequent analysis) be comparable to that of other studies of learning. That, briefly, is the rationale for the blocks.

Subjects were required to perform the task over four blocks, each block containing 10 trials\textsuperscript{4}.

\textsuperscript{3} For example, Adelman (1981), Hoffman et al. (1981), Schmitt et al. (1976), and Schmitt et al. (1977).

\textsuperscript{4} The literature on cognitive feedback does not provide specific guidance on the optimum number of blocks (or, for that matter, on the number of trials to be used within each block). In past studies, experimenters have used varying numbers of
### Table 3.1

<table>
<thead>
<tr>
<th>Between Subjects</th>
<th>Cognitive feedback</th>
<th>Block1</th>
<th>Block2</th>
<th>Block3</th>
<th>Block4</th>
</tr>
</thead>
<tbody>
<tr>
<td>No feedback</td>
<td>Block1</td>
<td>Block2</td>
<td>Block3</td>
<td>Block4</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Within Subjects</th>
</tr>
</thead>
</table>

#### 3.3.2 Phases of the Experiment

Figure 3.1 summarizes the overall sequence of the experiment. The experiment was carried out in three phases: training, experimental task and debriefing. These phases are detailed in sections 3.7, 3.8 and 3.9, respectively.

---

blocks. For example, Adelman (1981) used 3 blocks, each having 30 trials. Hoffman used 8 blocks (of 25 trials each), while Schmitt et al. (1976) used 3 blocks (of 15 trials each). The pilot sessions were conducted with five blocks (of 10 trials each). However, the sessions became unwieldy and subjects became restive in the last block. The actual experimental sessions were restricted to four blocks (of ten trials each).
Figure 3.1

Experimental Sequence

(a) Overall Sequence

Individual Training → Experimental Task → Individual Debriefing

(b) Experimental Task

Block 1 | Block 2 | Block 3 | Block 4

(c) Sequence within Each Block

Individual Task | Group Task | Individual Task
3.3.3 Experimental Setting

The sessions were conducted in a closed room in the presence of the experimenter. Figure 3.2 shows the layout of the experiment room and the seating of the group members. Except for the instructions and questionnaires (which were on printed material), the experiment was run entirely on computer terminals with software written by the experimenter.

![Diagram of experimental room layout](image-url)
3.4 Subjects

3.4.1 Choice of Subjects

The experiment was conducted with 90 volunteer students in a business school, 27 of whom were undergraduates and 63 were graduate students. Participants were divided into 30 groups (of three in each group). Each group was then assigned to the cognitive feedback or no feedback condition.

Randomization

Randomization was sought to be effected at three levels, viz:

(i) assigning groups to conditions,
(ii) assigning individuals to groups, and,
(iii) assigning individuals within groups, i.e., determining who was to be member 1, member 2 and member 35.

Randomization at levels (i) and (iii) were executed successfully. With respect to level (ii), the initial composition of each group was determined by randomly assigning individuals to groups. However, in order to accommodate individual time preferences, the composition of several groups (12 out of 30) had to be changed, some of them repeatedly. In this sense, therefore, randomization at level (ii) could not be implemented completely.

5. As sections 3.7.2 and 3.8.1 explain, the assignment of individuals within each group is relevant to training and the experimental task. In training, each member in a group is trained on a different policy, depending on whether the individual is member 1, 2 or 3. In the group part of the experimental task, member 2 acts as the group scribe.
How does the lack of complete randomization impinge on the results? Randomization among the conditions (i.e., level (i)) is clearly important, since it has a direct bearing on the results of the experiment (Kerlinger, 1986). Randomization at level (iii) is also important, because each member in a group is trained differently, and the attainment of cognitive conflict is contingent on the effectiveness of the training. Randomization at level (ii) ensures that groups are not distinctly different from each other in terms of skill differential. The task does not require particular skills (such as, knowledge of, or previous experience in, the task), and conversely, there is no reason to believe that the presence of such skills would enhance performance (in terms of the lens model). Thus, randomization was indeed effected where it mattered, and the absence of complete randomization at level (ii) does not pose any threat to the internal validity of the results.

Participant Profiles: Implications for Results

Two types of subject characteristics could potentially have affected results in this experiment: demographic factors and task-specific factors. Demographic factors were operationalized as age, sex, full time work experience, and level of education (i.e., whether the subject was an undergraduate or graduate student). The task-specific factor was operationalized by asking subjects whether they had any prior experience in the task.

Table 3.2 (a) profiles the subjects in terms of demographic factors. There were no significant differences among the two conditions. An inspection of education levels shows that most of the subjects (63 out of 90) were graduate students. The distribution of
subjects across conditions was not uniform in this respect. The proportion of graduate students to undergraduates was 3:2 in the feedback condition, and 4:1 in the no-feedback condition. One argument might possibly be that graduate students are better versed in statistical concepts than undergraduates and would therefore be in a position to make better use of the feedback than undergraduates (Fong, Krantz and Nisbett, 1986)6. If that is indeed the case, the results obtained here, then, are conservative in the sense that, given a more even distribution, we might possibly expect feedback groups to perform better (than in this experiment), and no-feedback groups to perform worse. That, in no way, detracts from the direction of performance differences posited in section 3.2.

Table 3.2 (a) - Subject Profiles: Demographics

Age:

<table>
<thead>
<tr>
<th>Age</th>
<th>Cognitive Feedback</th>
<th>No Feedback</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>20-25</td>
<td>24</td>
<td>24</td>
<td>48</td>
</tr>
<tr>
<td>26-30</td>
<td>12</td>
<td>15</td>
<td>27</td>
</tr>
<tr>
<td>31-35</td>
<td>5</td>
<td>1</td>
<td>6</td>
</tr>
<tr>
<td>36-40</td>
<td>4</td>
<td>3</td>
<td>7</td>
</tr>
<tr>
<td>41-</td>
<td>0</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Total</td>
<td>45</td>
<td>45</td>
<td>90</td>
</tr>
</tbody>
</table>

6. However, given that most undergraduates were seniors and most graduate students were in the first year of the MBA program, there is no reason to expect wide differences in statistical skills/knowledge.
### Sex:

<table>
<thead>
<tr>
<th>Age</th>
<th>Cognitive Feedback</th>
<th>No Feedback</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>M</td>
<td>34</td>
<td>32</td>
<td>66</td>
</tr>
<tr>
<td>F</td>
<td>11</td>
<td>13</td>
<td>24</td>
</tr>
<tr>
<td>Total</td>
<td>45</td>
<td>45</td>
<td>90</td>
</tr>
</tbody>
</table>

### Work Experience:

<table>
<thead>
<tr>
<th>Work Experience</th>
<th>Cognitive Feedback</th>
<th>No Feedback</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>9</td>
<td>12</td>
<td>21</td>
</tr>
<tr>
<td>1-3</td>
<td>21</td>
<td>15</td>
<td>36</td>
</tr>
<tr>
<td>4-6</td>
<td>8</td>
<td>6</td>
<td>14</td>
</tr>
<tr>
<td>7-9</td>
<td>4</td>
<td>6</td>
<td>10</td>
</tr>
<tr>
<td>10-</td>
<td>3</td>
<td>6</td>
<td>9</td>
</tr>
<tr>
<td>Total</td>
<td>45</td>
<td>45</td>
<td>90</td>
</tr>
</tbody>
</table>

### Level of Education:

<table>
<thead>
<tr>
<th>Program</th>
<th>Cognitive Feedback</th>
<th>No Feedback</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Undergraduate</td>
<td>18</td>
<td>9</td>
<td>27</td>
</tr>
<tr>
<td>Graduate</td>
<td>27</td>
<td>36</td>
<td>63</td>
</tr>
<tr>
<td>Total</td>
<td>45</td>
<td>45</td>
<td>90</td>
</tr>
</tbody>
</table>

Table 3.2 (b) indicates the distribution of subjects with respect to prior experience in the task. There were no significant differences between the two conditions.
Table 3.2 (b) - Subject Profiles: Task-Specific Factor

Prior Involvement in Personnel Selection:

<table>
<thead>
<tr>
<th>Prior involvement in personnel Selection?</th>
<th>Cognitive Feedback</th>
<th>No Feedback</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>No</td>
<td>23</td>
<td>25</td>
<td>48</td>
</tr>
<tr>
<td>Yes</td>
<td>22</td>
<td>20</td>
<td>42</td>
</tr>
<tr>
<td>Total</td>
<td>45</td>
<td>45</td>
<td>90</td>
</tr>
</tbody>
</table>

With respect to the factors specified above, therefore, the distribution of subjects across the two conditions was fairly uniform. This suggests that the characteristics of subjects over the two conditions were very homogeneous with respect to factors that may conceivably affect subject behavior in the context of the task. Thus, any experimental differences that may exist between the conditions are not likely to be related to differences in subject characteristics.

3.4.2 The use of Students as Subjects

The use of students as experimental subjects raises questions of external validity, viz, the issue of generalizability of results across persons (Cook and Campbell, 1979). In research on managerial decision making, the practice of using students has been criticized on the reasoning that business managers and students may not come from a homogeneous population (Abdel-Khalik, 1974). However, empirical evidence suggests that decisions made by students and business managers may in fact be quite similar if the task does not
require specific prior experience and skills that are unique to business managers (Ashton and Kramer, 1980).

This experiment is a study of how feedback affects individuals and groups. Nothing in the experiment, conceptually or methodologically, requires any experience with the task on the part of the subject or gives persons with experience (in the task) any added advantage. Therefore, the generalizability of the results within the context of the task is not a serious issue here.

**Experts versus Novices**

At a more general level, however, the external validity of the entire class of experiments dealing with policy capturing in one form or another, has been questioned, specifically in the light of differences between experts and novices. According to Brehmer and Brehmer (1988), the use of novices versus experts in policy capturing has two consequences.

*First*, (the novices) will not have any developed judgment policy, but will have to develop such a policy in the course of the policy capturing session. Studies of this kind are therefore best seen as experiments on how people transform what abstract declarative knowledge they may have about some domain into concrete judgments about actual cases. In other words, they are experiments on what Brehmer (1987) has called *policy construction*. *Second*, because these subjects have had little experience with the actual judgment task, they will have had no chance to learn the characteristics of the domain of real cases, such as the intercorrelations between cues and the distribution of cue values (p. 79).

The issue of the differences between cognitive processes of experts and novices is outside the scope of discussion here7. Suffice it

---

7. The issue of expert-novice differences has been studied extensively in knowledge acquisition for expert systems (LaFrance, 1989, provides a useful review of that literature). With respect to possible expert-novice differences in policy capturing
to say that reasons such as these make the task of applying feedback to actual organizational settings, considerably more complex.

3.5 Task

Human information processing is highly contingent on demands placed by the task (Payne, 1982). It is, therefore, important to delineate the role of the task in the context of the experiment. This section explains the logic underlying the selection of the task for this experiment.

3.5.1 Description

Details about instructions about the presentation of the task to subjects are contained in the appendix (A2.1 and A2.2). Briefly, the task entails screening candidates for a particular position from a large pool of applications. Subjects are given profiles of a set of candidates. These profiles are in the form of ratings on a 1-9 scale for each of three dimensions - work experience, test scores and education. Based on ratings on individual dimensions, subjects are then required to arrive at overall ratings of these candidates on a scale of 1-9.

---

8. Such exercises are frequently conducted by large corporations as the front end of the process of personnel selection.
Appropriateness of Task to the Experiment

The task is appropriate to the experiment for several reasons. First, tasks constructed in the manner described above are especially suitable for policy capturing, and therefore, have been widely used in lens model studies (e.g., Dougherty, Ebert and Callender, 1986; Lane, Murphy and Marques, 1982). Second, it is also a task where outcome information is typically absent or irrelevant and therefore appropriate to the context of the experiment. Finally, such tasks are routinely performed in groups or committees.

3.5.2 Task Characteristics

The elements comprising Brehmer's framework of tasks (section 2.2.3) primarily relate to the cue-criterion relationship. Since this experiment does not provide information about outcomes as either cognitive or outcome feedback, the framework as a heuristic is of limited value. Two elements apply specifically: task content and complexity.

Task content is not known to affect the structure of cognitive conflict (Hammond and Brehmer, 1973). In cognitive feedback literature as well, a general assumption is that results apply across a wide spectrum of tasks (section 2.5.4). The assumption made here - that task content does not affect results in such experiments - is warranted by available evidence.

Task complexity can be defined in many ways, and these definitions are largely contingent on the underlying theoretical framework (Wood, 1986; Campbell, 1988). Studies of human
information processing consider the number of cues to be a measure of task complexity (Payne, 1982; Wood, 1986). For cognitive conflict tasks, the number of cues appears to have no impact on the structure of conflict (Hammond and Brehmer, 1973). The role of task complexity in the effectiveness of cognitive feedback remains unclear (section 2.5.3). In general, however, performance in many tasks decreases if too many cues are processed (Ogilvie and Schmitt, 1979). Studies in individual and group information processing (e.g., Chalos and Pickard, 1985) have typically used between two and four cues. Hence the choice of three cues for this experiment.

3.6 Cognitive Feedback

3.6.1 Policy feedback

Policy feedback, whose purpose is to aid the subject in selecting a policy, can be provided through information on:

* Cue validity distribution, i.e., relative weights of cues in the decision rule (Hammond and Boyle, 1971), and,

* Functional relationships between cues and criterion, e.g., whether the cue-criterion relationship is linear or non-linear (e.g., Hoffman et al., 1981) or the slope of the line.

Policy feedback provided in the experiment is shown in figure 3.3. Feedback is provided on relative cue weights and not on function forms, for the following reasons:

9. It may be clarified here that cognitive feedback, in this instance, involves the delivery of cognitive information only (section 2.5.1, chapter 2).
(a) Information on cue weights is intuitively easy to understand and use (in refining a decision strategy). On the other hand, information stating that a certain function form is nonlinear is not only not understood easily, it is also difficult to apply in modifying a decision strategy.

(b) An important factor in designing feedback is the cognitive load imposed on the decision maker. From the perspective of the cost-benefit model (Beach and Mitchell, 1978), the marginal benefit accrued from function form information is not likely to exceed the extra cognitive effort required in processing the information.

Cue weights were calculated as follows:

(a) Beta weights were calculated from a multiple regression of cue values and the subject's estimates.

(b) The weights were then transformed to represent percentages of the sum of squared weights (as in Hoffman et al., 1981). Thus, beta values of 0.4, 0.8 and 0.2 were shown as 0.2, 0.75 and 0.05, respectively.10

(c) Transformed weights on the three cues were then displayed as a horizontal stacked bar, on a 1-100 scale (i.e., the transformed weights added up to 100).

---

10. In the pilot sessions, both forms of weights were experimented with. Subjects receiving weights as percentages comprehended the meaning of the information much more easily than subjects who were given original beta values.
Enter a number between 1-9, then press END. Press HOME if estimates are final.

FIGURE 3.3 - POLICY FEEDBACK
3.6.2 Consistency feedback

Consistency feedback provides the subject with guidance in implementing his/her decision strategy consistently. Thus, consistency feedback should enable the recipient to retain cognitive control (cf. Hammond and Summers, 1972). This implies two objectives in providing consistency feedback:

(a) Indicating the extent to which the subject is consistent with his/her underlying mental model, and,

(b) Aiding the subject in improving his/her level of consistency.

Objective (a) is achieved by furnishing an overall index of consistency, i.e., providing $R_{S\ell}$. Objective (b) is attained by providing the subject consistency scores on individual cases, i.e., criterion values the subject would have given if the person had been entirely consistent with his/her underlying model.

Figure 3.4 illustrates consistency feedback provided in the experiment. The overall index (or $R_{S\ell}$) was shown as a bar chart, on a 0-1 scale. It was calculated as the multiple correlation between the cues and estimates given by the subject. Consistency scores were calculated as follows:

(a) Beta weights were derived from a multiple regression of cue values and the subject's estimates.

(b) Criterion values were computed by multiplying cue values with respective beta weights.
(c) If any of the resultant criterion values were greater than 9 or less than 1, all criterion values computed were rescaled so as to lie between 1 and 9.\textsuperscript{11}

\textsuperscript{11}. It was observed from the pilot sessions that subjects were confused when the calculated values were outside the 1-9 range that they were using in making the original estimates. In order to resolve this problem, the actual experiments used the scaling procedure.
### (A) INDEX OF OVERALL CONSISTENCY

**CONSENSUS**

<table>
<thead>
<tr>
<th>MEMBER 1</th>
<th>MEMBER 2</th>
<th>MEMBER 3</th>
<th>GROUP</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**WORK** | **TEST** | **EDUCATION** | **MEMBER 1'S ESTIMATE** | **MEMBER 2'S ESTIMATE** | **MEMBER 3'S ESTIMATE** | **GROUP ESTIMATE**
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>1</td>
<td>4</td>
<td>3</td>
<td>5</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>1</td>
<td>9</td>
<td>1</td>
<td>4</td>
<td>6</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>8</td>
<td>2</td>
<td>7</td>
<td>5</td>
<td>4</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>6</td>
<td>6</td>
<td>1</td>
<td>5</td>
<td>3</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>9</td>
<td>3</td>
<td>7</td>
<td>4</td>
<td>4</td>
<td>2</td>
<td>6</td>
</tr>
<tr>
<td>1</td>
<td>4</td>
<td>9</td>
<td>6</td>
<td>3</td>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td>6</td>
<td>5</td>
<td>8</td>
<td>5</td>
<td>4</td>
<td>2</td>
<td>6</td>
</tr>
<tr>
<td>2</td>
<td>8</td>
<td>9</td>
<td>4</td>
<td>6</td>
<td>1</td>
<td>8</td>
</tr>
</tbody>
</table>

Enter a number between 1-9, then press END. Press HOME if estimates are final.

### (B) CONSISTENCY SCORES

**FIGURE 3.4 - CONSISTENCY FEEDBACK**
3.6.3 *Interpersonal Feedback.*

Figure 3.5 displays interpersonal feedback. Such feedback involves providing the decision maker with (a) policies of other group members, and (b) an indication of how closely their respective mental models are.

The display of policies is provided through policy feedback. The relative similarity of mental models is indicated through the *individual matching index* \((G)\). \(G\) was calculated as follows:

(a) A correlation was drawn between the model values of the decision policies (i.e., consistency scores) of any two subjects. This constituted the individual matching index between the two subjects. Since each group had three members, individual matching indices were calculated between members 1-2, members 1-3 and member 2-3.
MATCHING INDEX

AGREEMENT BTW.

<table>
<thead>
<tr>
<th>0</th>
<th>162</th>
<th>163</th>
<th>263</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>WORK EXP.</th>
<th>TEST SCORE</th>
<th>EDUCATION</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>1</td>
<td>9</td>
<td>1</td>
</tr>
<tr>
<td>8</td>
<td>2</td>
<td>7</td>
</tr>
<tr>
<td>6</td>
<td>6</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>6</td>
<td>1</td>
</tr>
<tr>
<td>9</td>
<td>3</td>
<td>7</td>
</tr>
<tr>
<td>1</td>
<td>4</td>
<td>9</td>
</tr>
<tr>
<td>1</td>
<td>6</td>
<td>5</td>
</tr>
<tr>
<td>3</td>
<td>2</td>
<td>8</td>
</tr>
<tr>
<td>2</td>
<td>9</td>
<td>9</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>MEMBER 1'S ESTIMATE</th>
<th>MEMBER 2'S ESTIMATE</th>
<th>MEMBER 3'S ESTIMATE</th>
<th>GROUP ESTIMATE</th>
</tr>
</thead>
<tbody>
<tr>
<td>3 3</td>
<td>5 6</td>
<td>3 3</td>
<td>3 3</td>
</tr>
<tr>
<td>3 6</td>
<td>4 6</td>
<td>4 4</td>
<td>4 4</td>
</tr>
<tr>
<td>4 5</td>
<td>5 6</td>
<td>5 5</td>
<td>5 5</td>
</tr>
<tr>
<td>5 4</td>
<td>2 4</td>
<td>4 3</td>
<td>2 3</td>
</tr>
<tr>
<td>4 6</td>
<td>3 3</td>
<td>3 3</td>
<td>3 3</td>
</tr>
<tr>
<td>3 5</td>
<td>2 3</td>
<td>1 1</td>
<td>2 2</td>
</tr>
<tr>
<td>6 9</td>
<td>8 9</td>
<td>6 8</td>
<td>6 9</td>
</tr>
</tbody>
</table>

Enter a number between 1-9, then press END. Press HOME if estimates are final.

FIGURE 3.5 - INTERPERSONAL FEEDBACK
3.6.4 Screen Design

Figure 3.6 displays the entire screen with all the feedback. The screen was designed on the basis of established guidelines in computer interface design (cf. Schneiderman, 1987). Since information presented through graphs has better properties than numerical information (Jarvenpaa and Dickson, 1988), feedback on decision rules, overall consistency and matching was provided with graphs. Given that the task required subjects to process considerable information, a systematic approach to design adopted here ensured that the information load was manageable\textsuperscript{12}.

\textsuperscript{12} Concerns were expressed during the pilot sessions as to whether subjects (a) understood what each element of feedback meant, and (b) whether they managed to retain the information through the course of the experiment. Concern (a) was addressed by training subjects in the feedback condition on the meaning and use of each piece of information appearing on the screen. This was done at the start of the experimental phase (details in the appendix, section A3.3). Concern (b) was addressed by circulating a brief questionnaire after the individual part (but before the group part) in block 1 (table 3.5). The questionnaire (appendix, A3.4) displayed each element of feedback and asked the participant to identify it on a multiple choice format. Only two (out of 45) participants did not identify one piece of feedback correctly.
Enter a number between 1-9, then press END. Press HOME if estimates are final.

**FIGURE 3.6 - ENTIRE SCREEN FOR EXPERIMENT**
3.7 The Training Phase

3.7.1 Basis for training

The starting point of a cognitive conflict experiment is to ensure that members of a group have dissimilar policies (section 2.3.2). This is typically done through training, wherein subjects are induced a particular decision structure through a multiple-cue probability learning exercise (see Brehmer, 1980, for a conceptual overview of MCPL).

Rohrbaugh (1979, 1981) used a different method of inducing conflict. Before the experiment, he inferred the decision structure of each subject from responses over a set of trials. This information was then used to cluster subjects (through INDSCAL) into homogeneous blocks. Each group was then formed with members from different clusters, i.e., subscribing to dissimilar decision structures.

The latter approach is problematic in that it gives less control to the experimenter in inducing a particular set of mental models. It is also cumbersome in administration and the assignment of subjects to groups. For these reasons, the former approach (which is the typical one) was chosen.

3.7.2 Training Procedure

The appendix (A2) details the procedure used for training. Figure 3.7 displays the training screen. The subject was given a set of cases. Each case represented a candidate’s profile on three dimensions - work experience, test scores, and education (each given on a 1-9 scale). The subject was required to give the candidate an
overall rating (again, on a 1-9 scale). This was followed by outcome feedback in the form of the "actual" committee estimate for that candidate. This process was repeated for 60 trials. No time limit was stipulated. The training took between 10 and 20 minutes.

The crux of the training was to induce a different mental models for each member of a group (i.e., create 3 different mental models within the group). This was done by varying the outcome feedback as follows:

(a) For member 1 of a group, the committee estimate on each trial was calculated as:

Criterion value = 0.8 * Cue1 + 0.1 * Cue2 + 0.1 * Cue3

(b) For member 2, the committee estimate was:

Criterion value = 0.33 * Cue1 + 0.33 * Cue2 + 0.33 * Cue3

(c) For member 3, the committee estimate was:

Criterion value = 0.1 * Cue1 + 0.1 * Cue2 + 0.8 * Cue3

The criterion value was rounded up to the nearest whole number. The predictability of the task was high (with an R² of 0.9).

------------------

Figure 3.7 goes here

------------------

3.7.3 Verifying the effectiveness of Training

The effectiveness of a multiple-cue probability learning exercise is measured through the lens model indices of matching and achievement¹³ (Brehmer and Hagafors, 1986). Since the explicit

¹³. G and r_a respectively in equation 1, section 2.2.2.
purpose of training is to induce a mental model, the matching index is a better indication of the effectiveness of learning in this context.

Matching

Table 3.3 provides means and standard deviations of subjects’ matching indices (G). The matching scores are uniformly high. This is consistent with past results for tasks of high predictability (Hammond and Summers, 1972; Brehmer and Hagafors, 1986). There were no significant differences in performance (a) among members 1, 2 and 3 (i.e., among members with different mental models), and (b) between subjects in the cognitive feedback and no feedback conditions. The appropriate conclusion is that the induction of mental models was indeed effective.

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall Index</td>
<td>0.95</td>
<td>0.08</td>
</tr>
<tr>
<td>Member 1</td>
<td>0.93</td>
<td>0.08</td>
</tr>
<tr>
<td>Member 2</td>
<td>0.97</td>
<td>0.08</td>
</tr>
<tr>
<td>Member 3</td>
<td>0.94</td>
<td>0.07</td>
</tr>
<tr>
<td>CFB</td>
<td>0.94</td>
<td>0.08</td>
</tr>
<tr>
<td>NFB</td>
<td>0.95</td>
<td>0.07</td>
</tr>
</tbody>
</table>

Notes:
1. CFB: Cognitive feedback
   NFB: No feedback
2. Differences among members (1, 2 and 3) were not significant.
3. Differences among subjects in the two conditions were not significant.
Achievement

Table 3.4 provides means and standard deviations of subjects' achievement indices ($r_a$). The scores are high (though less than the corresponding matching indices). This is consistent with past results for tasks of high predictability (Hammond and Summers, 1972; Brehmer and Hagafors, 1986). There were no significant differences in performance (a) among members 1, 2 and 3, and (b) between subjects in the cognitive feedback and no feedback conditions. These results reinforce the point that training was effective.

Table 3.4: Achievement ($r_a$) for training

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall Index</td>
<td>0.84</td>
<td>0.09</td>
</tr>
<tr>
<td>Member 1</td>
<td>0.84</td>
<td>0.09</td>
</tr>
<tr>
<td>Member 2</td>
<td>0.85</td>
<td>0.09</td>
</tr>
<tr>
<td>Member 3</td>
<td>0.85</td>
<td>0.08</td>
</tr>
<tr>
<td>CFB</td>
<td>0.84</td>
<td>0.09</td>
</tr>
<tr>
<td>NFB</td>
<td>0.84</td>
<td>0.09</td>
</tr>
</tbody>
</table>

Notes:
1. CFB: Cognitive feedback
   NFB: No feedback
2. Differences among members 1, 2 and 3 were not significant.
3. Differences among subjects in the two conditions were not significant.

Correspondence with Subjective Weights

As a further check, subjects were asked (at the end of the training session) to assign weights to each dimension so that they added up to 100. The objective was to assess if (and to what extent)
subjects had insight into their own judgment policies (cf. Hoffman, 1960; Reilly and Doherty, 1989).

Figures 3.8 (a - c) display subjective weights plotted against regression-derived weights on each cue. The pattern of randomness and the absence of a clear trend indicates that subjective weights (i.e., responses elicited from subjects) do not bear a clear relationship with the "actual" weights. This confirms past findings (summarized in Hammond et al. 1975) that subjects do not exhibit much self-insight.
Figure 3.8 (a) - Training Data
FIGURE 3.8 (b) - TRAINING DATA
FIGURE 3.8 (C) — TRAINING DATA

Subjective Weight

Actual Weight

Cue 3: Education
3.8 Experimental Phase

The experiment was performed over four blocks, each block containing ten trials. The appendix (A3.1, A3.2 and A3.3) contains details of instructions given to subjects at the start of the experiment.

Each block required subjects to perform two tasks - the individual task and the group task. Table 3.5 provides a step-by-step description of the sequence of tasks subjects were required to perform within each block.

3.8.1 Individual Task

The individual task can best be described in terms of the following scenario:

Subject S sits at the terminal and sees a screen containing instructions. After pressing any key, he/she sees ten cases (in the bottom half of the screen), each representing a candidate's profile on three dimensions - work experience, test scores, and education. S is asked to make estimates on each case, i.e., arrive at a value between 1 and 9 (the bottom row always contains instructions on what is to be done next). In doing so, he/she can scroll up or down and change values of estimates. After filling in all ten estimates, S presses the END key. If S happens to be in the feedback condition, the screen is refreshed with feedback (figure 3.6). In any case, the bottom half is refreshed with the latest estimates of the other members of the group. S is free to iterate, i.e., repeat the process described above, any number of times. When S has finalized the estimates, he/she presses the HOME key instead of the usual END key, thus signifying the end of S's part of the individual task. No time limit is stipulated and S is not allowed to interact with other members of the group.
The individual part of the task was completed when all three members pressed the HOME key. In the cognitive feedback condition, each subject was administered a short questionnaire in block 1 (appendix, A3.4), after completion of the individual task\textsuperscript{14}.

3.8.2 Group task

After completion of the individual task, members worked as a group, making a set of group estimates on the same cases. The screens were the same and the procedure for iterating was the same as above. Each group had a scribe, a member who entered estimates once the group as a whole agreed on a value\textsuperscript{15}. The scribe for each group was member 2. No time limit was imposed.

3.8.3 Manipulation Check

After the completion of the group task, each group member was asked to make independent estimates on the same cases, without the knowledge of the other members. No cognitive feedback was provided at this stage.

At the end of the experimental treatment, each subject was given a short questionnaire (appendix, A3.5).

\textsuperscript{14} See note 11.

\textsuperscript{15} In arriving at estimates during the group task, members were allowed to converse freely.
Table 3.5 - Sequence of tasks within each block of the experimental phase

**Overall sequence**

* Four blocks, conducted immediately after one another.

**Sequence within each block**

* Individual Task
* Group Task
* Manipulation Check

**Individual Task:** At the start of each block, each individual is given ten cases. The individual does the following:

(i) Each subject makes estimates on all ten cases.
(ii) Subjects in cognitive feedback condition receive feedback.
(iii) Subject studies feedback and revises estimates.
(iv) Subject is free to iterate between steps (ii) and (iii) indefinitely till he/she is satisfied with the estimates.

**Group Task:** After the completion of the individual task (i.e., after all members of a group have finalized their estimates), the group is given the same ten cases. The group does the following:

(i) Group makes estimates on all ten cases.
(ii) Groups in cognitive feedback condition receive feedback.
(iii) Group studies feedback and revises estimates.
(iv) Group iterates between steps (ii) and (iii) till the group members are satisfied with the estimates.

**Manipulation Check:** After the completion of the group task, each individual goes through a manipulation check. This involves the following:

(i) Each subject is given the same ten cases along with (a) his/her individual estimates on the cases and (b) the group estimates.
(ii) The subject is required to make a fresh set of estimates (which may be, but need not be, identical to the subject's earlier estimates).
(iii) No feedback is provided in either of the conditions.
3.9 Debriefing

The objective in conducting the debriefing session was to have subjects record their own profiles as well as for the experimenter to get a sense of how the experiment was viewed by the subjects (details of the questionnaire are in the appendix, section A4).

In response to the question of whether subjects took the task seriously, only five individuals responded with a score of less than four (on a scale of 1-5). This corresponds with the experimenter's observation that subjects were very serious in performing the task. Subjects were also asked (on a five point scale) whether they found the task to be interesting or realistic. Table 3.6 indicates the answers to those questions. The means indicate that subjects (including ones with experience in the task) did find the task interesting and realistic (differences among the two groups - experienced and not experienced - were not significant). This points to two conclusions. First, since the task was perceived to be interesting, we expect that the scale of cognitive resources that subjects committed to the task was similar to what decision makers commit in real-life tasks. Second, subjects who had experience in the task did not view the task as an artificial one.
Table 3.6 - Subject Attitudes toward Task

<table>
<thead>
<tr>
<th></th>
<th>Was task Interesting?</th>
<th>Was task Realistic?</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>Std dev</td>
</tr>
<tr>
<td>Overall</td>
<td>3.80</td>
<td>1.13</td>
</tr>
<tr>
<td>Subjects with no experience in pers. sel.</td>
<td>3.94</td>
<td>1.14</td>
</tr>
<tr>
<td>Subjects with prior experience in pers. sel.</td>
<td>3.65</td>
<td>1.11</td>
</tr>
</tbody>
</table>

Note: Questions were asked on a five point scale.

3.10 Summary

In this chapter, we discussed the experimental setting. Four aspects of the experimental set-up are especially important to the study and bear repeating:

(i) The experiment was conducted with thirty groups (of three members each). Fifteen groups were assigned to the feedback condition and the others to the no-feedback condition. The experiment was conducted over four blocks of ten trials each. Table 3.1 illustrates the research design and table 3.5 provides a step-by-step description of the experimental phase of the task.

(ii) The first part of the experiment consisted of training the subject, in order to ensure that each member of a group had a different decision model of the task. The training was done through an MCPL task consisting of 60 trials. The results of the training, measured by the
matching index and achievement (summarized in tables 3.3 and 3.4), indicate that the training was effective. (iii) Cognitive feedback was provided in three components: policy feedback, consistency feedback and interpersonal feedback. Policy feedback was provided through stacked bar charts of regression-derived cue weights. Consistency feedback was provided through an overall consistency index and individual consistency scores on each case. Interpersonal feedback was provided through bar-graph representations of the matching index.

(iv) Subjects performed the task of personnel screening. They were given scores (on a scale of 1-9) on three dimensions - work experience, test scores and education. They were required to evaluate each case (i.e., each candidate) on a scale of 1-9. Outcome feedback (i.e., feedback on the "accuracy" of their responses) was not provided. Subjects were undergraduate and graduate business students. Differences (if any) among the populations do not pose a threat to the internal validity of the results.
CHAPTER 4
RESULTS

4.1 Scope

This chapter reports the results of the study and discusses their implications. Section 4.2 specifies the method of analysis, the operationalization of dependent variables and sample sizes. Sections 4.3 and 4.4 report the results in the context of individual and group data, respectively. Section 4.5 integrates the results.

4.2 Framework for the Analysis of Data

4.2.1 Model for Analysis

The data was analyzed using a Type(2)$^1 \times \text{Blocks}(4)$ design, with repeated measures on the latter. The analyses were conducted on SAS$^R$ statistical software, using the General Linear Models procedure.

The Use of Repeated Measures

When several measurements are taken on the same experimental unit (in this case, each individual or group), the measurements tend to be correlated with each other. When the measurements represent qualitatively different things (such as height, weight, etc), this correlation can be included in the analysis by using multivariate methods, such as MANOVA. When the measurements

$^1$ That is, the type of feedback (cognitive feedback or no feedback).
are responses to different levels of some experimental factor of interest (in this case, *time*, operationalized as blocks), the correlation can be taken into account through a repeated measures analysis of variance (SAS/STAT Guide, 1987, p. 602).

Repeated measures analysis can be performed by univariate or multivariate methods. LaTour and Miniard (1983) point out that all the univariate approaches present difficulties. The conventional method often results in Type I errors, whereas the Greenhouse-Geisser procedure (or a modified version of it) understates the true level of significance. Thus, except in designs where within-subject factors have two levels (in which case the univariate and multivariate procedures produce identical results), the multivariate procedure should be used. This study reports results output by the multivariate method¹.

4.2.2 Dependent variables - Individual Data

Table 4.1 displays the dependent variables used in the analysis and the respective sample sizes.

---

¹. In this study, the univariate and multivariate methods produced identical results in every analysis.
Table 4.1 - Individual data: Sample Sizes for Dependent Variables

<table>
<thead>
<tr>
<th>Variable</th>
<th>Lens Model Index</th>
<th>N per Block</th>
</tr>
</thead>
<tbody>
<tr>
<td>Consistency</td>
<td>$R_{si}$</td>
<td>90a</td>
</tr>
<tr>
<td>Matching Index</td>
<td>$G$</td>
<td>30b</td>
</tr>
<tr>
<td>Agreement</td>
<td>$r_A$</td>
<td>30b</td>
</tr>
<tr>
<td>No of Iterations</td>
<td>N.A.</td>
<td>90a</td>
</tr>
</tbody>
</table>

Notes: a - One datapoint for each member. 
b - Composite of 3 datapoints.

Consistency

Consistency is defined as the multiple correlation between the cues and the criterion (Brehmer 1979). Consistency for a subject in each block is calculated from estimates made by the subject in his/her final iteration in that block. Thus, for each group, there are three data points in each block, indicating the respective consistencies of three members.

Matching Index

The matching index ($G$) is defined as the correlation between model estimates (of criterion values) of two members (Brehmer 1979). The matching index in each block is calculated as follows:

(i) The matching index for any two members is calculated. Thus, three data points are obtained - $G_{12}$, $G_{13}$, $G_{23}$. 
(ii) The matching indices are then averaged to form one composite index in each block for each group\(^2\).

**Agreement**

The index of agreement \((r_\text{A})\) is defined as the correlation between criterion values of two members (Brehmer 1979). Agreement in each block is calculated as follows:

(i) Agreement for any two members is calculated. Thus, three data points are obtained - \(r_{12}, r_{13}, r_{23}\).

(ii) The three measures are then averaged to form one composite measure in each block for each group.

**Number of Iterations**

The number of iterations made by each subject was recorded (by the software) at the time of the experiment.

### 4.2.3 Dependent variables - Group Data

Table 4.2 displays the dependent variables used in the analysis and the respective sample sizes.

---

\(^2\) The averages taken for matching and agreement indices, both at the individual and group levels, were arithmetic means.
<table>
<thead>
<tr>
<th>Variable</th>
<th>Lens Model Index</th>
<th>N per Block</th>
</tr>
</thead>
<tbody>
<tr>
<td>Consistency</td>
<td>$R_{si}$</td>
<td>30a</td>
</tr>
<tr>
<td>Matching Index</td>
<td>$G$</td>
<td>30b</td>
</tr>
<tr>
<td>Agreement</td>
<td>$r_A$</td>
<td>30b</td>
</tr>
<tr>
<td>No of Iterations</td>
<td>N.A.</td>
<td>30a</td>
</tr>
</tbody>
</table>

Notes: 
- a - One datapoint in each block.
- b - Composite of 3 datapoints.

**Consistency**

Consistency in each block is calculated from estimates made by the *group* in their final iteration in *that* block. Thus, for each group, there is one data point in each block.

**Matching Index**

The matching index in each block is calculated as follows:

(i) The matching index between the group model and the model of *each* member (from his/her last iteration in *that* block) is computed.

Thus, three data points are obtained - $G_{1g}$, $G_{2g}$, $G_{3g}$.

(ii) The matching indices are then averaged to form one composite index in each block for each group.

**Agreement**

The agreement index in each block is calculated as follows:

(i) The agreement between the group estimates and that of each members (from his/her last iteration in *that* block) is calculated.

Thus, three data points are obtained - $r_{A1g}$, $r_{A2g}$, $r_{A3g}$. 
(ii) The three measures are then averaged to form one composite measure in each block for each group.

4.3 Experimental Results: Individual Data

4.3.1 Consistency:

As figure 4.1 demonstrates, means for the feedback condition are higher in all the blocks. This conclusion is also supported in table 4.3 (a), which indicates a significant difference in consistency attributable to type. Furthermore, the possibility that individual consistency is a function of membership in a particular group (i.e., the type(group) factor) is ruled out. Thus, hypothesis H\textsubscript{11} is confirmed.

Table 4.3 (a) - Individual Data: Between Subjects Analysis
Dependent Variable: Consistency (n = 90)

<table>
<thead>
<tr>
<th>Effect</th>
<th>F Value</th>
<th>degrees of freedom</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type</td>
<td>83.80</td>
<td>1, 59</td>
<td>0.0001</td>
</tr>
<tr>
<td>Type (Group\textsuperscript{3})</td>
<td>1.23</td>
<td>1, 28</td>
<td>0.2485</td>
</tr>
</tbody>
</table>

3. The type(group) effect denotes any effect due to an individual belonging to a particular group.
Table 4.3 (b) indicates a type and block interaction effect, which is hypothesized in $H_{12}$. Thus, $H_{12}$ is confirmed. However, there is a significant main (block) effect as well\textsuperscript{4}, thereby indicating an increase in individual consistency over time. This suggests some form of learning effect\textsuperscript{5}. An inspection of figure 4.1 suggests different behavior patterns among the two conditions. Individuals in the feedback condition reached a high degree of consistency at the end of the first block (mean = 0.92, block 1) and stabilized over the remaining blocks (mean = 0.94, block 4). This can be interpreted as suggesting that cognitive feedback enables individuals to retain cognitive control over time. For individuals not receiving feedback, the picture is less clear. By the second block, these individuals appear to have lost cognitive control\textsuperscript{6}. However, by the third block, these individuals seem to have

\textsuperscript{4} When there is a main effect and an interaction effect, interpretation becomes a complex issue. Some writers (e.g., Petersen, 1985) assert that in the presence of interaction, the main effects have no real meaning, whether significant or not. Adopting that position, we can assert here that the main effect is not important. However, other writers (e.g., Kerlinger, 1986) suggest that both the main effect and interaction should be explained if the main effect is strong and the interaction is ordinal. In this case, the problem is not serious, since the interaction effect was hypothesized and the interaction is not ordinal. Nevertheless, given that we have a strong main effect, further interpretation is desirable.

\textsuperscript{5} The phenomenon of learning over time has been studied extensively in industrial psychology (Yelle, 1979, provides a useful survey). However, for the following reasons, findings from that literature are not useful for the purposes of this study. (i) Learning in such studies is typically measured as time taken (or marginal time taken) to complete a particular task (e.g., Bailey, 1989). While time is an important variable with implications for efficiency, the emphasis here is on measuring learning by gauging decision quality. Quality does not appear to be a popular variable in learning curve studies. (ii) Learning curve studies tasks are not oriented toward studying decision behavior, i.e., tasks where the principal output is a decision.

\textsuperscript{6} A possible explanation for this lies in the multiple levels of decisions required of the individuals, and the resultant demand on their cognitive capacities placed by the decision sequence. As table 3.6 explains, subjects were required to perform
recovered some of their consistency. This appears to be the result of learning over time and is reflected in the block effect. In summary, there is some learning in both conditions, but much of it is explained by the interaction between condition and time.

Table 4.3 (b) - Individual Data: Within Subjects Analysis

<table>
<thead>
<tr>
<th>Effect</th>
<th>Wilks' Lambda</th>
<th>F value</th>
<th>degrees of freedom</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Block</td>
<td>0.605</td>
<td>12.61</td>
<td>3</td>
<td>58</td>
</tr>
<tr>
<td>Block x Type (Group)</td>
<td>0.799</td>
<td>4.84</td>
<td>3</td>
<td>58</td>
</tr>
<tr>
<td>Block x Type</td>
<td>0.278</td>
<td>1.11</td>
<td>84</td>
<td>174</td>
</tr>
</tbody>
</table>

A trend analysis of the respective means (table 4.3 c) provides further insight. Quite clearly, there is a quadratic trend, and this trend is accounted for by the type (of condition). This conclusion is supported by figure 4.1, which indicates that while the feedback means display more or less a linear pattern, the no-feedback means are aligned in a (negative) quadratic fashion. The existence of non-linearity in the no-feedback condition clearly indicates a lack of uniformity in the dependent variable for that condition only. This reinforces the assertion in H₁₂.

---
7. For the multivariate analyses, identical F values were obtained for Pillai’s Trace, Hotelling-Lawley Trace and Roy’s Greatest Root.
Table 4.3 (c) - Individual Data: Trend Analysis for Blocks
Dependent Variable: Consistency (n = 90)

<table>
<thead>
<tr>
<th></th>
<th>Linear F-value</th>
<th>p</th>
<th>Quadratic F-value</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>5.39</td>
<td>0.0237</td>
<td>29.90</td>
<td>0.0001</td>
</tr>
<tr>
<td>Type</td>
<td>1.44</td>
<td>0.2351</td>
<td>13.51</td>
<td>0.0005</td>
</tr>
<tr>
<td>Type(Group)</td>
<td>1.21</td>
<td>0.2614</td>
<td>0.90</td>
<td>0.6153</td>
</tr>
</tbody>
</table>

4.3.2 Matching among Members' Models

Figure 4.2, which displays the means for the cognitive feedback and no feedback conditions over blocks, indicates a higher matching index for the feedback condition. However, evidence in table 4.4 (a), does not warrant the conclusion that this difference is significant. Thus, H21 is not accepted.

Table 4.4 (a) - Individual Data: Between Subjects Analysis
Dependent Variable: Matching Index (n = 30)

<table>
<thead>
<tr>
<th>Effect</th>
<th>F Value</th>
<th>degrees of freedom</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type</td>
<td>1.64</td>
<td>1, 28</td>
<td>0.2113</td>
</tr>
</tbody>
</table>


From figure 4.2 and table 4.4 (b), we can conclude that the matching index increases over time. Hypothesis H22 is, therefore, accepted.

Table 4.4 (b) Individual Data: Within Subjects Analysis
Dependent Variable: Matching Index (n = 30)

| Effect      | Wilks' Lambda | F value | degrees of freedom | p     |
|-------------|................|---------|--------------------|-------|
| Block       | 0.387         | 13.68   | 3                  | 26    | 0.0001 |
| Block x Type| 0.990         | 0.08    | 3                  | 26    | 0.9682 |

As the figure and table 4.4 (c) indicate, the trend is linear. Thus, we conclude that matching is a function of time, and not of feedback.

Table 4.4 (c) - Individual Data: Trend Analysis for Blocks
Dependent Variable: Matching Index (n = 30)

<table>
<thead>
<tr>
<th></th>
<th>Linear</th>
<th>Quadratic</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>F-value</td>
<td>p</td>
</tr>
<tr>
<td>Mean</td>
<td>23.59</td>
<td>0.0001</td>
</tr>
<tr>
<td>Type</td>
<td>0.05</td>
<td>0.8239</td>
</tr>
</tbody>
</table>

4.3.3 Agreement among Members' Estimates

Figure 4.3 displays means for the agreement index over blocks. The figure, along with evidence in table 4.5 (a), confirms hypothesis H31 - that agreement in the feedback condition is greater than that in the no-feedback condition.
Table 4.5 (a) - Individual Data: Between Subjects Analysis

Dependent Variable: Agreement (n = 30)

<table>
<thead>
<tr>
<th>Effect</th>
<th>F Value</th>
<th>degrees of freedom</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type</td>
<td>8.41</td>
<td>1, 28</td>
<td>0.0072</td>
</tr>
</tbody>
</table>

From table 4.5 (b), hypothesis $H_{32}$ - that agreement increases over time - is accepted. Also, as table 4.5 (c) indicates, this increase follows a linear trend.

Table 4.5 (b) Individual Data: Within Subjects Analysis

Dependent Variable: Agreement (n = 30)

<table>
<thead>
<tr>
<th>Effect</th>
<th>Wilks' Lambda</th>
<th>F value</th>
<th>degrees of freedom</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Block</td>
<td>0.270</td>
<td>23.37</td>
<td>3, 26</td>
<td>0.0001</td>
</tr>
<tr>
<td>Block x Type</td>
<td>0.971</td>
<td>0.26</td>
<td>3, 26</td>
<td>0.8568</td>
</tr>
</tbody>
</table>

Table 4.5 (c) - Individual Data: Trend Analysis for Blocks

Dependent Variable: Agreement (n = 30)

<table>
<thead>
<tr>
<th></th>
<th>Linear F-value</th>
<th>p</th>
<th>Quadratic F-value</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>36.01</td>
<td>0.0001</td>
<td>0.49</td>
<td>0.4886</td>
</tr>
<tr>
<td>Type</td>
<td>0.05</td>
<td>0.8332</td>
<td>0.25</td>
<td>0.6211</td>
</tr>
</tbody>
</table>
Figure 4.3 - Individual Data: Agreement
Dependent Variable: Feedback

Block

+ No Feedback

Cognitive Feedback
4.3.4 Number of Iterations

Figure 4.4 displays means of the number of iterations over blocks. From the figure and table 4.6 (a), the hypothesis H41 - that individuals in the feedback condition iterate more often than those in the no-feedback condition - is accepted. However, there is also a significant type(group) effect. Members in some groups appear to have iterated more often than those in other groups. A plausible explanation for this can be found in evoking the notion of norm development.8

Table 4.6 (a) - Individual Data: Between Subjects Analysis

<table>
<thead>
<tr>
<th>Effect</th>
<th>F Value</th>
<th>degrees of freedom</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type</td>
<td>56.61</td>
<td>1, 59</td>
<td>0.0001</td>
</tr>
<tr>
<td>Type( Group)</td>
<td>2.46</td>
<td>1, 28</td>
<td>0.0018</td>
</tr>
</tbody>
</table>

8. The suggestion here is that groups developed some pattern of norms when subjects were performing their individual task. Thus, if one or more members iterated repeatedly, others followed suit. (Notes made by the experimenter indicate that this pattern was noticeably pronounced in eight of the thirty groups. These groups did not have any distinguishable peculiarities with respect to subject characteristics discussed in chapter 3.) Groups are known to develop performance and social norms in order to maintain behavioral consistency (Shaw, 1981, p. 279). However, the development of norms is thought to be largely dependent on the processes of group interaction. This result suggests that the provision of interpersonal feedback engenders a mechanism similar to that of group interaction, even though the feedback is directed at the individual. This finding also reinforces the suggestion made in section 2.4.2 that interpersonal feedback is a complicated notion: it encompasses phenomena at both individual and group levels, and yet is distinct from either.
Table 4.6 (b) points to a block effect as well (in addition to the type(group) effect discussed earlier). Table 4.6 (c) (as well as figure 4.4) indicates a quadratic trend in \textit{both} conditions.

<table>
<thead>
<tr>
<th>Effect</th>
<th>Wilks' Lambda</th>
<th>F value</th>
<th>degrees of freedom</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Block</td>
<td>0.841</td>
<td>3.66</td>
<td>3</td>
<td>58</td>
</tr>
<tr>
<td>Block x Type</td>
<td>0.950</td>
<td>1.01</td>
<td>3</td>
<td>58</td>
</tr>
<tr>
<td>Block x Type (Group)</td>
<td>0.164</td>
<td>1.72</td>
<td>84</td>
<td>174</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Effect</th>
<th>Linear F-value</th>
<th>p</th>
<th>Quadratic F-value</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>9.48</td>
<td>0.0031</td>
<td>7.15</td>
<td>0.0096</td>
</tr>
<tr>
<td>Type</td>
<td>2.50</td>
<td>0.1193</td>
<td>0.29</td>
<td>0.5948</td>
</tr>
<tr>
<td>Type(Group)</td>
<td>1.66</td>
<td>0.0507</td>
<td>1.22</td>
<td>0.2562</td>
</tr>
</tbody>
</table>
In looking for a plausible explanation of results in tables 4.6 (b and c), correlations were computed between the number of iterations and the other dependent variables. Of these measures, the association between consistency and the number of iterations offers some insight. First, the overall correlation is significant (prob. < 0.001). The correlation is significant in the feedback condition (prob. < 0.05), thereby indicating that subjects iterated on the basis of feedback, improving their consistency with more iterations. However, the correlation between consistency and the number of iterations is not significant in the no-feedback condition. Thus, it appears that subjects failed to enhance their consistency with additional iterations.

Correlations were then computed within each condition, over four blocks. None of the correlations in the no-feedback condition were significant. In the feedback condition, the correlations in blocks 1 and 2 were significant (prob < 0.10), whereas those in blocks 3 and 4 were not. Figure 4.4 indicates that for the feedback condition, the number of iterations in the last two blocks remained quite high. This suggests that while feedback was still being used, its rate of use had declined in comparison with the level of consistency attained. This may have been caused by a learning effect which, over time, enabled subjects to attain a certain level of consistency with less input by way of feedback. This

9. If that is indeed the case, it may be possible that the need for feedback declines over time. A conclusion such as this can have important implications for the design of feedback. However, data from this experiment does not warrant such an inference (figure 4.4 shows the number of iterations to be increasing in the last block). If iterations denote the use of feedback, then feedback is evidently still being used in the last block. The only way to answer the question definitively is to conduct process tracing, a point that is elaborated in chapter 5.
may also account for the quadratic trend in the means for the feedback condition, whereas the trend for consistency is linear (figure 4.1).

4.4 Experimental Results: Group Data

4.4.1 Consistency:

Figure 4.5 displays means for consistency over the blocks. From table 4.7 (a), we conclude that feedback groups did indeed have greater consistency than no-feedback groups. Thus, hypothesis H13 is accepted.

Table 4.7 (a) - Group Data: Between Subjects Analysis

*Dependent Variable: Consistency* (n = 30)

<table>
<thead>
<tr>
<th>Effect</th>
<th>F Value</th>
<th>degrees of freedom</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type</td>
<td>16.45</td>
<td>1, 28</td>
<td>0.0004</td>
</tr>
</tbody>
</table>

Table 4.7 (b) points to a block x type interaction effect. We can, therefore, accept hypothesis H14.

Table 4.7 (b) - Group Data: Within Subjects Analysis

*Dependent Variable: Consistency* (n = 30)

<table>
<thead>
<tr>
<th>Effect</th>
<th>Wilks’ Lambda</th>
<th>F value</th>
<th>degrees of freedom (num denom)</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Block</td>
<td>0.507</td>
<td>8.39</td>
<td>3 26</td>
<td>0.0005</td>
</tr>
<tr>
<td>Block x Type</td>
<td>0.738</td>
<td>3.07</td>
<td>3 26</td>
<td>0.0454</td>
</tr>
</tbody>
</table>
However, there is also a strong block effect. The figure and table 4.7 (c) also point to a quadratic trend. This trend is more pronounced in the no-feedback condition (which is reflected in the block x type effect). A comparison with figure 4.1 is appropriate here. For the feedback condition, the means for individual data are more or less in a straight line, whereas the means for group data are in a curve. For the no-feedback condition, the group means are higher than the means for individual data. This is consistent with previous findings that groups are more consistent than individuals (section 2.6.1). However, the shape of the curve is different. Individuals appear to hit a low in block 2, but recover in block 3. As explained above, this may be caused by the cognitive load placed on the individual, which is eventually alleviated somewhat by a learning effect. In the case of the group, the means in blocks 2 and 3 are about the same. In one sense, unlike the individual case, the group mean in block 2 fell, but not as much as in the individual case. In another sense, however, the group mean in block 3 did not recover. An inspection of the corresponding values of the individual and group means suggests that the former explanation has greater credence. Thus, while the group could not cope adequately in block 2, its performance did not decline as much as in individuals. This is in accordance with the view mentioned above that groups have a mechanism for reducing random errors, and that this mechanism holds up over time\textsuperscript{10}.

\textsuperscript{10} An analysis of the group process provides another perspective on the results. See tables 4.10 (a) and (b) and the accompanying explanation.
Table 4.7 (c) - Group Data: Trend Analysis for Blocks
Dependent Variable: Consistency (n = 30)

<table>
<thead>
<tr>
<th></th>
<th>Linear F-value</th>
<th>p</th>
<th>Quadratic F-value</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>0.11</td>
<td>0.7433</td>
<td>25.18</td>
<td>0.0001</td>
</tr>
<tr>
<td>Type</td>
<td>0.16</td>
<td>0.6878</td>
<td>8.25</td>
<td>0.0077</td>
</tr>
</tbody>
</table>

4.4.2 Matching with Members' Individual Models

Figure 4.6 - which displays means for the matching index - and table 4.8 (a), serve to confirm hypothesis H23 that groups getting feedback will have a higher matching index (i.e., match with the members' individual policies) than groups not receiving feedback.

Table 4.8 (a) - Group Data: Between Subjects Analysis
Dependent Variable: Matching Index (n = 30)

<table>
<thead>
<tr>
<th>Effect</th>
<th>F Value</th>
<th>degrees of freedom</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type</td>
<td>4.86</td>
<td>1, 28</td>
<td>0.0359</td>
</tr>
</tbody>
</table>

Table 4.8 (b) confirms H24, which predicts the matching index to increase over time. As table 4.8 (c) indicates, the trend is linear.

Table 4.8 (b) - Group Data: Within Subjects Analysis
Dependent Variable: Matching Index (n = 30)

<table>
<thead>
<tr>
<th>Effect</th>
<th>Wilks' Lambda</th>
<th>F value (num denom)</th>
<th>degrees of freedom</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Block</td>
<td>0.382</td>
<td>14.00</td>
<td>3</td>
<td>26</td>
</tr>
<tr>
<td>Block x Type</td>
<td>0.898</td>
<td>0.98</td>
<td>3</td>
<td>26</td>
</tr>
</tbody>
</table>
### Table 4.8 (c) - Group Data: Trend Analysis for Blocks

*Dependent Variable: Matching Index (n = 30)*

<table>
<thead>
<tr>
<th></th>
<th>Linear F-value</th>
<th>p</th>
<th>Quadratic F-value</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>23.56</td>
<td>0.0001</td>
<td>0.12</td>
<td>0.7337</td>
</tr>
<tr>
<td>Type</td>
<td>0.04</td>
<td>0.8528</td>
<td>1.25</td>
<td>0.2729</td>
</tr>
</tbody>
</table>
Figure 4.4 - Individual Data
Dependent Variable: Number of Iterations
Cognitive Feedback

Block

1.50
1.75
2.00
2.25
2.50
2.75
3.00
3.25
3.50
3.75
4.00
+ No Feedback
4.4.3 Agreement with Members' Individual Estimates

Figure 4.7 displays means for the cognitive feedback and no-feedback conditions. Hypothesis H33, which predicts that feedback groups will have a higher agreement index than no-feedback groups, is accepted from evidence in table 4.9 (a).

Table 4.9 (a) - Group Data: Between Subjects Analysis
Dependent Variable: Agreement (n = 30)

<table>
<thead>
<tr>
<th>Effect</th>
<th>F Value</th>
<th>degrees of freedom</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type</td>
<td>17.82</td>
<td>1, 28</td>
<td>0.0002</td>
</tr>
</tbody>
</table>

Hypothesis H34, i.e., notion that agreement will increase over time, is also accepted. However, as figure 4.7 and the trend analysis (table 4.9 c) indicate, the alignment of the means for the no-feedback condition is somewhat ambiguous. The explanation is to be found in the results obtained for consistency (table 4.7 c). Since agreement is a function of consistency, it is likely that values obtained for consistency have cascaded to the results in table 4.9 (c).

Table 4.9 (b) - Group Data: Within Subjects Analysis
Dependent Variable: Agreement (n = 30)

<table>
<thead>
<tr>
<th>Effect</th>
<th>Wilks' Lambda</th>
<th>F value</th>
<th>degrees of freedom</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Block</td>
<td>0.258</td>
<td>24.88</td>
<td>3</td>
<td>26</td>
</tr>
<tr>
<td>Block x Type</td>
<td>0.857</td>
<td>1.44</td>
<td>3</td>
<td>26</td>
</tr>
</tbody>
</table>
Figure 4.7 - Group Data
Dependent Variable: Agreement

- No Feedback
- Cognitive Feedback
Table 4.9 (c) - Group Data: Trend Analysis for Blocks
Dependent Variable: Agreement (n = 30)

<table>
<thead>
<tr>
<th></th>
<th>Linear F-value</th>
<th>p</th>
<th>Quadratic F-value</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>29.46</td>
<td>0.001</td>
<td>0.84</td>
<td>0.3675</td>
</tr>
<tr>
<td>Type</td>
<td>0.00</td>
<td>0.9986</td>
<td>1.60</td>
<td>0.2157</td>
</tr>
</tbody>
</table>

4.4.4 Groups' use of Feedback

Group decision making is prone to process losses, which often results in lower decision quality. One way to overcome process losses has been to structure group decisions (Delbecq et al., 1975). Since cognitive feedback is posited to improve decision quality, then it is useful to study if and in what manner group processes were structured over the two conditions. This can be done in the context of the two-stage model of human information processing that forms the basis for cognitive feedback - selection of a decision rule and its implementation (section 2.5).

Formulation of Decision Rule

From table 4.10 (a) indicates whether groups formulated a decision rule. We learn that groups in the feedback condition were more likely to explicitly formulate a decision rule than groups in the no-feedback condition. Thus, groups in the feedback condition had a better structure of the decision process than groups in the no-feedback condition, implying that the provision of feedback induces some structure in the group process.
Table 4.10 (a) - Group Data: Analysis of Notes
Formulation of Decision Rule

<table>
<thead>
<tr>
<th>Condition</th>
<th>Decision Strategy Formulated?</th>
<th>Yes</th>
<th>No</th>
</tr>
</thead>
<tbody>
<tr>
<td>Feedback</td>
<td></td>
<td>12</td>
<td>3</td>
</tr>
<tr>
<td>No Feedback</td>
<td></td>
<td>7</td>
<td>8</td>
</tr>
</tbody>
</table>

Note: A group was considered to have formulated a decision rule if the group explicitly went through the exercise of doing so.

The results of table 4.10 (b), which indicates the specific point of time a group formulated the decision rule, are also insightful. Groups with feedback formulated their decision rules fairly early on in the task (blocks 1 and 2). This may explain why the consistency of these groups was steady over all the blocks (figure 4.5). However, the no-feedback groups either did not formulate a policy or did so late in the task (blocks 3 and 4). This serves to explain why the mean consistency of no-feedback groups fell early on (no systematic decision procedure), and rose in block 4 (after some groups developed a systematic decision procedure). Notes indicate that as the task progressed, some no-feedback groups realized that a decision rule would be of help, and proceeded to formulate one.

Table 4.10 (b) - Group Data: Analysis of Notes
Time of formation of Decision Rule

<table>
<thead>
<tr>
<th>Condition</th>
<th>Block1</th>
<th>Block2</th>
<th>Block3</th>
<th>Block4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Feedback</td>
<td>9</td>
<td>2</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>No Feedback</td>
<td>2</td>
<td>1</td>
<td>4</td>
<td></td>
</tr>
</tbody>
</table>
The results of tables 4.10 (a) and (b), therefore, lead to an interesting conclusion regarding group processes. Groups with cognitive feedback realize the importance of structure (in their decision processes) very early on. This leads them to formulate a decision rule (which then constitutes the basis for subsequent action). On the other hand, groups without feedback do not appear to grasp the notion of structure initially. Eventually, some of them do, and when they do formulate a decision rule, their performance (in terms of consistency) improves. Feedback groups recognize that, in order to take advantage of the feedback, they need to adopt some structure in their decision processes (a realization that is not apparent in the no-feedback groups). A reasonable inference that one can draw, is that, the provision of cognitive feedback has the effect of inducing groups to adopt some structure in their processes.

**Use of Feedback**

The feedback provided can be isolated into four components: consistency index, consistency scores, policy weights and matching index (section 3.5). The notes provide some indication on (a) which components of feedback were actually used by the groups (table 4.10 - c), and (b) the manner in which they were used (tables 4.10 - d, e and f).

Table 4.10 (c) indicates that consistency scores and policy weights were used with some regularity. On the other hand, the indices of consistency and matching were used more sparingly. The reason for this difference lies in the utility of feedback. Policy weights
help a group in formulating a decision rule. Consistency scores provide a concrete aid for a group to improve their consistency. On the other hand, the index of consistency is abstract in that it acts as a signal of overall consistency, but does not help in enhancing it. The matching index, while denoting the extent to which members' individual policies converge, does not help in formulating a group policy.

### Table 4.10 (c) - Group Data: Analysis of Notes

*Use of feedback by groups*

<table>
<thead>
<tr>
<th>Type of Feedback</th>
<th>Block1</th>
<th>Block2</th>
<th>Block3</th>
<th>Block4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Consistency Index</td>
<td>3</td>
<td>4</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Consistency Scores</td>
<td>12</td>
<td>10</td>
<td>12</td>
<td>13</td>
</tr>
<tr>
<td>Policy Weights</td>
<td>14</td>
<td>12</td>
<td>12</td>
<td>13</td>
</tr>
<tr>
<td>Matching Index</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>

**Note:** Use of a particular type of feedback by a group was inferred from specific references made by group members about that piece of feedback. So, if a group member directed the attention of the group to consistency scores, the group was considered to have used consistency scores in that block.

Table 4.10 (d) illustrates the manner in which consistency feedback was used by groups. The most persistent pattern was the use of consistency scores in revising estimates. This happened in two ways: groups sought to improve their consistency, or the consistency score was used as a means of settling disagreements among group members. In a few instances, consistency scores were used in
tandem with consistency index. In those cases, a low consistency
index triggered off a warning to the group, which then used the
consistency scores to improve their consistency. In most of the cases,
however, the groups were able to operate on the basis of the scores
alone. It may be possible that the index of consistency is used only
when consistency falls below a certain threshold (of tolerance).

<table>
<thead>
<tr>
<th>Manner of use</th>
<th>Consistency Index</th>
<th>Consistency Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Feedback is recognized and activates revision of estimates</td>
<td>5</td>
<td>29</td>
</tr>
<tr>
<td>Feedback is recognized but causes no action</td>
<td>5</td>
<td>10</td>
</tr>
<tr>
<td>Other/Not clear</td>
<td>3</td>
<td>8</td>
</tr>
</tbody>
</table>

Note: The count represents the number of times a type of feedback was used. Thus, if a group used the consistency index over 3 blocks, the count for that group would be 3.

Policy weights were used primarily formulating the group
decision rule (table 4.10 e)\textsuperscript{11}. They were occasionally used for
revising group estimates, when the group felt that the current policy

\textsuperscript{11} In a few instances, differences in initial policy weights (among individuals) were so strong that some group members suspected that they had been trained on different policies (as indeed they had been). When asked, the experimenter stated that the question could not be answered then, but that everything would be clarified later (and was, after the debriefing session).
was not what was conceived by group members when they formulated their estimates.

Table 4.10 (e) - Group Data: Analysis of Notes

<table>
<thead>
<tr>
<th>Manner of use</th>
<th>Policy Weights</th>
</tr>
</thead>
<tbody>
<tr>
<td>Feedback is used for revising group decision rule</td>
<td>29</td>
</tr>
<tr>
<td>Feedback is used for formulating group decision rule</td>
<td>5</td>
</tr>
<tr>
<td>Feedback is recognized but causes no action</td>
<td>12</td>
</tr>
<tr>
<td>Other/Not clear</td>
<td>5</td>
</tr>
</tbody>
</table>

The matching index was not used by groups in meaningful sense (table 4.10 f). Some groups briefly noticed the index, but did not use it.
Table 4.10 (f) - Group Data: Analysis of Notes

<table>
<thead>
<tr>
<th>Manner of use</th>
<th>Matching Index</th>
</tr>
</thead>
<tbody>
<tr>
<td>Feedback is recognized and activates revision of estimates</td>
<td>0</td>
</tr>
<tr>
<td>Feedback is recognized but causes no action</td>
<td>3</td>
</tr>
<tr>
<td>Other/Not clear</td>
<td>1</td>
</tr>
</tbody>
</table>

Group Behavior: Other Dimensions

The notes also reveal an impressionistic picture of the general patterns of group behavior. The groups were task-oriented, in the sense that they concentrated on the task rather than on the maintenance aspect. That is indeed what one would expect, given that (a) the groups were ad hoc, (b) lasted only for the duration of the task (an hour or so), and (c) the nature of the task required serious effort toward the task. That was evident in the patterns of discussion among the groups, which were very task-oriented. Moreover, the fact that group members differed in their policies did not appear to cause any conflict among them. That is also what one would expect, given that the conflict was strictly cognitive, and a rather benign one. Disagreements among group members were usually resolved, first through discussion, then by majority rule. In feedback groups, disagreements were often settled through the use of consistency scores (though the reverse, i.e., majority rule overriding consistency
scores, was also in evidence). From the perspective of the group members, the scores were supposed to represent "better" decisions. The scores also had the property of being neutral (making them amenable to acceptance by all group members). On the other hand, when the majority rule produced a solution, consistency sometimes became a secondary consideration.

4.4.5 Comparison with Cognitive Conflict Studies

While this study was based on the interpersonal conflict paradigm, it had two important differences: (a) there was no outcome information, and (b) the groups were not dyads, but were three member groups. A juxtaposition of the results of this study against findings from the cognitive conflict literature (section 2.3) offers some interesting perspectives. First, the broad findings regarding the change in the structure of the conflict, are replicated here. Disagreement indeed starts out as policy conflict, but is sustained by inconsistency (figures 4.1 and 4.2). However, the extent of disagreement appears to reduce over time (figure 4.3). This is somewhat different from the common conception of disagreement as enduring over time. Evidently, the absence of outcome information causes groups to reduce disagreement over time. This stands to reason, because, given the lack of information about "correctness", group members have fewer reasons to disagree anyway.
4.5 Summary of Results

The results of the study are summarized in tables 4.11 (a) and (b). At the individual level, the provision of feedback (a) enables higher consistency and (b) greater agreement among group members. The matching of individual policies increases over time, but is not (significantly) enhanced by cognitive feedback. In using feedback, individuals also iterate through their decision processes more often than individuals who do not get feedback.

At the group level, feedback enables groups to attain higher consistency. Groups with feedback also make decisions whose underlying models are more similar to those of their individual members than groups without feedback. Groups with feedback also make estimates closer to the estimates of the individual members, than the no-feedback groups. An analysis of notes on group processes indicates that cognitive feedback groups are more likely to formulate decision rules early in the task, thus imparting some structure to the group process. The cognitive feedback groups appear to rely extensively on two components of feedback: policy and consistency scores. The consistency index is used sparingly. There was no evidence that groups used the matching index at all.
Table 4.11 (a): Summary of Results
Individual Data

Dependent Variable: Consistency

1. \( H_{11}: \) CFB individuals will be more consistent than NFB individuals. 
   Accepted \((p < 0.0001, n = 90): Table 4.3 (a)\)

2. \( H_{12}: \) CFB individuals will be uniformly consistent over time, whereas NFB individuals will not be uniformly consistent over time. 
   Accepted \((p < 0.0045, n = 90): Table 4.3 (b)\)

Dependent Variable: Matching

3. \( H_{21}: \) The CFB matching index will be higher than the NFB matching index. 
   Not Accepted \((p > 0.1000, n = 30): Table 4.4 (a)\)

4. \( H_{22}: \) The matching index (in both conditions) will increase over time. 
   Accepted \((p < 0.0001, n = 30): Table 4.4 (b)\)

Dependent Variable: Agreement

5. \( H_{31}: \) The CFB agreement index will be higher than the NFB agreement index. 
   Accepted \((p < 0.0072, n = 30): Table 4.5 (a)\)

6. \( H_{32}: \) The agreement index (in both conditions) will increase over time. 
   Accepted \((p < 0.0001, n = 30): Table 4.5 (b)\)

Dependent Variable: No. of Iterations

7. \( H_{41}: \) CFB individuals will iterate more than NFB individuals. 
   Accepted \((p < 0.0001, n = 90): Table 4.6 (a)\)

Notes:
1. CFB: Cognitive Feedback
   NFB: No Feedback
2. Hypotheses are discussed chapter 3.
3. Tables containing the results are in chapter 4.
Table 4.11 (b): Summary of Results

*Group Data*

<table>
<thead>
<tr>
<th>Dependent Variable: Consistency</th>
</tr>
</thead>
</table>
| 1. \( H_{11} \): CFB groups will be more consistent than NFB groups.  
*Accepted* \((p < 0.0001, \ n = 30)\): *Table 4.7 (a)* |
| 2. \( H_{12} \): CFB groups will be uniformly consistent over time, whereas NFB groups will not be uniformly consistent over time.  
*Accepted* \((p < 0.0005, \ n = 30)\): *Table 4.7 (b)* |

<table>
<thead>
<tr>
<th>Dependent Variable: Matching</th>
</tr>
</thead>
</table>
| 3. \( H_{21} \): The CFB matching index will be higher than the NFB matching index.  
*Accepted* \((p > 0.0359, \ n = 30)\): *Table 4.8 (a)* |
| 4. \( H_{22} \): The matching index (in both conditions) will increase over time.  
*Accepted* \((p < 0.0001, \ n = 30)\): *Table 4.8 (b)* |

<table>
<thead>
<tr>
<th>Dependent Variable: Agreement</th>
</tr>
</thead>
</table>
| 5. \( H_{31} \): The CFB agreement index will be higher than the NFB agreement index.  
*Accepted* \((p < 0.0002, \ n = 30)\): *Table 4.9 (a)* |
| 6. \( H_{32} \): The agreement index (in both conditions) will increase over time.  
*Accepted* \((p < 0.0001, \ n = 30)\): *Table 4.9 (b)* |

Notes:  
1. CFB: Cognitive Feedback  
2. NFB: No Feedback  
3. Hypotheses are discussed in chapter 3.  
4. Tables containing the results are in chapter 4.
CHAPTER 5
CONCLUSIONS AND LIMITATIONS

5.1 Scope

This chapter revisits salient findings of the study and places them in the larger context of research on decision feedback and information systems. In doing so, we sketch out some conceptual and methodological issues that limit the scope and applicability of this study and therefore need to be addressed in future research. Section 5.2 discusses the findings of the study in the context of decision feedback research. Section 5.3 outlines potential applications of cognitive feedback in the design of information systems. Section 5.4 discusses limitations of this study (and more generally, of the underlying theoretical frameworks). Section 5.5 summarizes the work.

5.2 Conclusions of the Study

We argued in chapter 2 (note 12) that although decision making has been studied extensively at the individual and group levels, not much research has been done on determining how factors (such as cognitive feedback) that affect individual decisions work to influence decision making at the group level. This study extends cognitive feedback research from the individual level, and in doing so, provides a better understanding of processes at work in interpersonal and group information processing. In this section, we briefly restate findings of
past research and revisit the results of the study in the context of these findings.

5.2.1 A Recapitulation of Past Research

Cognitive Feedback

Decision feedback is categorized into two types: outcome and cognitive. Outcome feedback, or feedback on the accuracy of responses, has been found to be dysfunctional, especially for complex tasks and/or low environmental predictability. Cognitive feedback, on the other hand, has attractive properties in terms of its utility as a decision aid1.

Referent Level of Feedback

Researchers in social psychology have found the referent level of feedback to be a factor in determining the usefulness of feedback. Three referent levels are of essence2:

(a) Individual feedback, i.e., feedback wherein the unit of decision making and feedback pertain to the decision processes of the individual.

(b) Interpersonal feedback, i.e., feedback, wherein the unit of decision making is the individual, but feedback pertains to the decision processes of other individuals in the group.

(c) Group feedback, i.e., feedback on group decisions. Here, the unit of decision making as well as the feedback itself, pertains to the group as an entity.

---

1. Sections 2.4 and 2.5, chapter 2, survey research on decision feedback.

2. The implications of providing feedback at multiple referent levels are discussed in sections 2.4.2 and 2.6, chapter 2.
Cognitive Conflict

When different individuals have different mental models of a situation, making decisions in a group often leads to cognitive conflict. While cognitive conflict originates as differences in individual decision models, these mental models tend to converge over a period of time. In the process of recalibrating their respective models, however, individuals lose cognitive control and become inconsistent in implementing their decision models. This leads to a persistence of cognitive conflict\(^3\).

Absence of Outcome Feedback

Studies of feedback have concentrated on situations where outcome information is available. The effectiveness of cognitive feedback in the absence of outcome information (a common characteristic of individual and group decisions) has not been investigated carefully in past research.

5.2.2 Research Question

The question posed by the study was: (a) with the existence of cognitive conflict within groups, (b) does providing cognitive feedback help improve decision quality (c) at multiple referent levels (d) in the absence of outcome feedback?

\(^3\) Research in cognitive conflict is examined in section 2.3, chapter 2.
5.2.3 Findings of the Study

Lens Model Indices

Decision quality was operationalized in terms of three lens model indices: consistency, matching and agreement.  
* Consistency: The study found cognitive feedback to enhance consistency at the interpersonal and group levels. Individuals/groups having the benefit of cognitive feedback retained an enhanced level of consistency in a uniform manner; others (i.e., no-feedback individuals/groups) fluctuated in their cognitive control and hence displayed erratic levels of consistency.  
* Matching: Subjects receiving cognitive feedback recorded a higher matching index than subjects not receiving feedback. This difference was significant at the group level, but not at the individual level. The matching index also increased over time for both conditions.  
* Agreement: Recipients of cognitive feedback (individuals as well as groups) registered greater values in the agreement index than the no-feedback individuals and groups.

Groups’ use of Feedback

An analysis of notes on the group process offer some interesting insights into group behavior:  
* Formulation of Decision Rule: The delivery of cognitive feedback to groups was associated with greater structure in these groups. This was evidenced by the fact that more groups in the feedback condition formulated a decision rule than the corresponding number in the no-feedback condition. Also interesting is the time of formulation of the
decision rule. Groups in the feedback condition that did formulate a rule typically did so early in the task (blocks 1 and 2), whereas the corresponding groups in the no-feedback condition formulated such rules late in the task (block 3). The provision of feedback appears to induce a sense of structure in groups, perhaps by triggering a recognition that creating structure in the group process (in this case, by formulating a group decision rule) enables feedback to be used productively.

*Use of Feedback:* Groups appeared to have used two components of feedback quite extensively: policy feedback and consistency scores. The consistency index, on the other hand, was used very rarely.

### 5.2.4 Consistency in Decision Making

Perhaps the salient finding of the study is the role of consistency in interpersonal and group information processing. Since consistency is so prominent a variable in this study, it is important to reinterpret consistency in the overall sense of the term and delineate its role in the context of findings reported here.

**What is Consistency?**

Hogarth (1982) points out that consistency is a relative concept that has several meanings. From one perspective, *logical consistency* refers to conformance with normative principles. This version of consistency has been studied extensively in economics in the context of interpersonal and intertemporal consistency (Elster, 1983). From another perspective, *process consistency* refers to the regularity with which people apply behavioral rules in responding to given stimuli, as
in making judgments. Process consistency itself can be viewed in two ways. One concerns the consistency with which a particular person responds to the same stimulus configuration on two independent occasions. An example of this is the measurement of test-retest correlations. The other form of process consistency concerns the application of the same behavioral rule across several stimuli. It is this form of consistency that has been studied extensively in the context of "man versus model-of-man", cognitive control and cognitive feedback research.

Is Consistency always Desirable?

Research in behavioral decision theory emphasizes the importance of consistency in decision making (Hogarth, 1986). However, the significance of inconsistency in behavioral processes should not be overlooked (Hogarth, 1982), especially in probabilistic environments (Brunswik, 1952). In such environments, suggests Brunswik, we mitigate the effects of uncertainty by using cues and behavioral rules in interchangeable fashion. While such intersubstitutability implies behavioral inconsistency by definition, it is more valid over time than decision strategies that place reliance on a single (imperfectly valid) cue. In competitive environments too, inconsistency in behavioral rules can be functional in the form of strategic behavior, by increasing the difficulty others have in predicting one's action (Hogarth and Makridakis, 1981).

Consistency in the Context of the Research Findings

There are two conclusions we draw with respect to consistency (i.e., application of a rule across multiple stimuli or cases). First, we
demonstrate that the problem of inconsistency exists not only for individuals (as past research in bootstrapping suggests), but at the interpersonal and group levels as well. At the interpersonal level, the presence of inconsistency signifies cognitive conflict, and our results broadly replicate findings in the cognitive conflict literature. At the group level, inconsistency signifies deficiencies in the capability to integrate information.4

The second conclusion we draw is with respect to the utility of cognitive feedback in enhancing consistency at multiple referent levels. This study provides evidence that cognitive feedback improves consistency not only at the level of the individual (as past research shows), but at other the interpersonal and group levels as well. This effect is manifest not only across conditions, but over time as well.

5.2.5 Feedback and Group Behavior

Our analysis of interpersonal and group decision behavior points to two interesting conclusions. First, the provision of feedback is associated with greater structure in group process. Structured group processes have long been envisaged as a method of minimizing process losses and thereby enhancing the quality of group

4. Groups are thought to be a powerful device for minimizing random errors associated with individual decision making, and as a consequence, considerably more consistent (chapter 2, note 15). Our study confirms this general pattern (of groups being more consistent than individuals), but the fact that groups benefited from feedback also reveals a lack of optimality in group performance. This is in consonance with the argument made in chapter 2 (section 2.6.1) that both individuals and groups are subject to deficiencies in their information processing mechanisms, thereby establishing an a priori justification for providing decision feedback at both levels.
performance (Delbecq et al., 1975). The results presented here confirm this general premise.

The second conclusion we draw is regarding interpersonal decision behavior. Apart from its connotations with respect to information processing (i.e., cognitive conflict research), this aspect of behavior has not been studied carefully in past research. The possibilities are intriguing, as suggested by the manner in which individuals iterated in their decision task (that is, reformulated their respective decision rules)\(^5\). In the individual component of the task, individual behavior was influenced by that of other group members, thereby suggesting that the provision of feedback appears to trigger a mechanism similar to that of group interaction. While more research is needed for a clearer understanding of this mechanism, this study offers a glimpse of the complexity of such behavior.

5.3 Applications of Cognitive Feedback

5.3.1 Cognitive Feedback in CSCW

Chapter 1 (section 1.1.2) pointed out a tendency in CSCW research of developing systems without tying them specifically to individual and group decision processes. These systems are feature oriented, i.e., built on the presumption that providing a set of apparently useful features will somehow affect the overall decision (process or outcome). Such a practice creates two problems. First, it is difficult to assess the impact/benefits these systems may have on decision

---

\(^5\) We refer to the results discussed in section 4.3.4.
making (Kraemer and King, 1988). Claims about the utility of these systems are based more on anecdotal evidence than tangible results (e.g. project NICK - Cook et al., 1987). Second, some features that these systems provide create "side-effects", i.e., trigger complex processes in decision making without actually planning or accounting for them (e.g., COLAB - Stefik et al., 1987). Such effects should be considered explicitly as design principles in building group support systems. These two points are elaborated below, the first by outlining the utility of feedback as a decision aid, and the second by specifying implications for the design of multiuser interfaces.

Cognitive Feedback as a Decision Aid

A preliminary survey of empirical research on the effectiveness of systems for supporting group work indicates the benefits of such systems to be primarily affective (Kraemer and King, 1988). While affective benefits are important in measuring the quality of group decisions (DeSanctis and Gallepe, 1987), they do not by themselves justify the construction and use of systems for group support. Systems built for collaborative work must also make claims of improving decision quality in an individual or collective sense. Cognitive feedback as a decision aid enhances overall decision quality and thereby fulfills these desiderata.

In suggesting that decision support systems be considered from a systemic perspective, Ariav and Ginzberg (1985) provide a set of

6. Since we concentrate on the no-feedback situation, this study operationalizes decision quality as consistency, matching and agreement. If outcome information is available (and reliable), decision accuracy would be another measure of decision quality.
heuristics for the design of such systems. Using these heuristics as a 
broad basis, the following sketches out a set of principles for the 
incorporation of cognitive feedback in group support.

* Environment: An important aspect of the environment is the type of 
task supported. The class of tasks that closely maps with the task 
used in the study (thereby justifying the transferability of results) are 
those requiring inductive inference, i.e., tasks where decision makers 
draw inferences (regarding judgment or choice) from an available set 
of cues (Hammond et al., 1980). Organizational analogs of such tasks 
being performed in collective settings are arguably abundant, e.g., 
evaluations by internal audit committees (Libby, 1982), decisions by 
loan committees (Chalos and Pickard, 1975), student selection by 
admission committees (Hogarth, 1986), screening of applications by 
human resource specialists, selection of personnel by interview 
committees (Dougherty et al., 1986).

* Role: In the context of group support, we interpret role as the type of 
decision support afforded by the system (or part of the system). 
Cognitive feedback provides support specifically as a decision-analytic 
aid (which is a level 2 feature in the framework proposed by DeSanctis 
and Gallupe, 1987), as opposed to features that facilitate the exchange 
of information (i.e., level 1 features).

* Components/Arrangement: We envisage an arrangement that 
involves a tight coupling of all three components of a decision support 
system - model management and dialog management. The model 
management function will provide facilities for the computation of 
feedback by containing algorithms (such as regression equations,
discriminant functions or tools for modeling categorical data), and selecting the appropriate set of algorithms. The role of dialog management will be in the delivery of feedback. This will entail (a) building capabilities for the display of different types of graphical formats (e.g. stacked bars, pie charts, plots, etc) and (b) selecting formats based on the type of task and user preferences. The data management function will be instrumental in providing the data for computation.

**Designing Multiuser Interfaces: Implications for WYSIWIS**

The second implication of our study regarding the design of systems for CSCW involves the "side-effects" argument made at the beginning of the section. A case in point is WYSIWIS (What you see is what I see), proposed by Stefik, Bobrow, Foster, Lanning and Tatar (1987b) as a conceptual basis for the design of multiuser interfaces. The first generation of multiuser interface implementations (e.g., Stefik et al., 1987) are based on variations of WYSIWIS.

The key concept in WYSIWIS is the exchange of interpersonal information, as operationalized through stampsheets (Stefik et al., 1987) or graphical rooms (Card, Pavel and Farrell, 1985). The arguments in WYSIWIS, e.g., the frequency with which screens should be refreshed, primarily relate to technical capabilities/constraints. This study reveals an aspect of WYSIWIS not considered previously, i.e., the issue of interpersonal information processing. In other words, what effects do the introduction of devices for the exchange of interpersonal information (such as stampsheets) have on information processing, and therefore, on decision quality? Issues such as information load,
the frequency of change of information and decision maker consistency are fundamental to decision quality and (a) should be investigated and (b) included as design principles for multiuser interfaces.

5.3.2 Knowledge Acquisition in Expert Systems

Given that knowledge in organizations is distributed among several experts (Huber, 1984), developing multiple-expert knowledge acquisition techniques in general, and providing explicit techniques for resolving conflicts in particular, are critical to developing knowledge based systems where knowledge is gleaned from several experts (Shaw and Woodward, 1988). Here, we outline the potential for cognitive feedback to serve as a knowledge acquisition device in the context of multiple experts.

We envisage a two-phase mechanism for knowledge acquisition. In the first phase, each expert develops a knowledge base of his own7. This is followed by the second phase, a two step mechanism closely modeled on the framework for the use of feedback (Hogarth, 1986):

(a) Incorporation: In this step, the knowledge bases created from the first phase are analyzed. Feedback is then provided on the degree of affinity between them and their key differences. Thus, participants are

---

7. The task of doing this is similar to the individual decision making/feedback situation in the sense that an individual is asked to formulate a set of decision rules and is provided feedback. This is followed by the "group" decision aspect. Note that this sequence is similar to the individual-group sequence of decision making predicated earlier.
provided with interpersonal feedback in the form of cognitive information.

(b) *Interpretation and adjustment:* In this step, the experts interpret the feedback and adjust their respective knowledge bases accordingly. The steps (a) and (b) are repeated till one of two events occurs - the knowledge bases are identical, or no further areas of consensus can be found.

The crux of this strategy is to develop a method for comparing the knowledge bases. The comparison can be made from two perspectives - integration and acquisition of information. In the first, we identify the set of cues used by both experts and point out differences in their respective cue-combination patterns (this is similar to the feedback provided in this experiment). In the second, we point out cues used by one but not the other, thereby attempting to uncover differences in their "conceptual schemas".

A number of issues must be resolved for this mechanism to be feasible. First is the question of *validation*. Shaw and Gaines (1986) argue that it is not sufficient to show that reasonable expert systems can be developed; knowledge engineering methods must also be *validated*. The issue here is: what would be a suitable method of validating the technique? The resultant expert system can be evaluated in terms of its outputs if appropriate outcome information exists. The answer is less clear in the absence of outcome information. A possible approach would be to compare the combined

---

8. In which case, it is probably more efficient to construct models of the environment instead (von Winterfeldt, 1988).
knowledge base with the individual knowledge bases in terms of such measures as consistency of output.

The second issue centers around what mechanism should be triggered if conflicts cannot be resolved. One approach is to compare the unresolved rules individually and retain the ones deemed superior (by an a priori method of comparison).

Finally questions about specific operationalization of the feedback or specific domains in which the mechanism would be effective, need to be resolved.

5.4 Limitations and Further Research

In defining the scope of an experiment, a researcher seeks to remove potential confounds and establish greater control. While control strengthens the credibility of the results, it also constrains them in that it narrows the boundaries of the study. A consequence is the exclusion of issues that are, otherwise, interesting as well as relevant.

Here, we make the point that the theoretical basis for studying feedback should be extended to capture issues that are not addressed in the lens model/cognitive conflict paradigm. First, we argue that the concept of cognitive conflict needs redefinition. Then we refer to the cognitive feedback framework - the representativeness of tasks, components of feedback, and the use of feedback.

5.4.1 Cognitive Conflict: the Need for a Richer Perspective

In cognitive conflict experiments, subjects are presented a set of cues predetermined by the experimenter (as in this study), so that the
researcher can then study how people integrate information. In the process, however, a very important element of human judgment, the acquisition of information, is neglected. What makes a study of information acquisition interesting and important is that in most judgment tasks, cues are intersubsitutable (Hogarth, 1981). Besides, there is evidence that different individuals performing the same task do indeed form mental models consisting of different cues (Gentner and Stevens, 1983). The researcher could, for example, present subjects with a large number of cues and observe (a) how many and which cues are selected and (b) what implications the selection of cues have for performance. Studies in information acquisition in the lens model literature are extremely rare (Chalos and Pickard (1985) is an exception) and do not appear to have been conducted at all in the cognitive conflict paradigm either9.

5.4.2 Issues in Cognitive Feedback Research

The Representativeness of Tasks

An oft-repeated argument made by Brunswik (1943, 1952) was that laboratory situations should retain the representativeness of tasks - i.e., the structure of the task used in an experiment should approximate its structure in the environment. Thus, if the cue inter-

9. To study the acquisition of information in this experiment would have opened up another level of analysis - cognitive conflicts caused not only by cue integration, but by the choice of cues as well. However, to get some sense of the possibilities, subjects were asked at the end of the experiment what other cues they might have considered in making such a decision. Most subjects mentioned one cue, usually one of two types of variables - work-related (e.g. past performance) or demographic (age or personality of applicant). This may suggest that the most important variables (perceived by subjects) were already factored into the task.
correlations of a task are usually high, the laboratory version of the
task should replicate that structure. Tasks used in experimental
situations, however, are seldom representative (Brehmer and Brehmer,
1988), usually for sound experimental reasons. For example, tasks in
policy capturing studies routinely use orthogonal cues so that the
utilization of each cue can be observed with greater precision\(^{10}\).
However, as Phelps and Shanteau (1978) demonstrated, this practice
impinges on the validity of the conclusions.

In their study, Phelps and Shanteau (1978) compared the
decision behavior of livestock judges using orthogonal cues as well as
correlated cues. When judges made decisions from orthogonal
(dichotomous) cues, analyses of variance of their judgment policies
yielded few interactions, leading the authors to conclude that the
judgments followed a linear model. There were 9-11 main effects,
suggesting that the judges used 9-11 cues. In the second condition,
the same judges were given highly correlated cues. Multiple
regression analyses yielded few significant beta weights, indicating that
the judges used very few cues. Whether the differences between the
two conditions indicate real differences in judgment policies or whether
they are merely statistical artifacts, is not clear. In other words,
commonly held conclusions about human judgment inferred from policy
capturing studies may be more appropriate to laboratory settings than

\(^{10}\) The same is true of tasks employed in cognitive conflict studies (and, for that
matter, this study). The argument here is that strong intercorrelation among cues
often masks policy differences, creating misleading results (section 2.3.3, chapter
2).
actual environmental setting. This is clearly a limitation of such studies (including this research), and should be rectified.

Another way in which the representativeness of tasks impinges on the results is the format of information presented (Brehmer and Brehmer, 1988). The effect of format on decision behavior is well-known (Payne, 1982). Studies in the lens model framework adopt a rigid mode of presentation, which creates the problem of "paper people". Instead of being given real people (or cases), subjects are presented with a set of "coded" cues (such as high/low or 1-9 as in this experiment). Arguably, instances of tasks where information is presented in such fashion, are numerous (e.g., student screening, loan evaluation, etc.). However, using coded information may lead us to underestimate of one form of expertise, that of extracting (or evaluating) cue information. Thus, two physicians, given the same X-ray plate, may come up with different diagnoses (see Libby, 1982, for some disquieting possibilities), not because they integrate the cues differently, but because they interpret the cue values differently in the first place. Considering that physicians go through elaborate training before they can correctly interpret X-ray plates, this is a real possibility. In presenting the information already coded, we may not only be getting an incomplete picture of clinical expertise, but may actually be creating some artifacts, as in presenting a physician with an orthogonal cue set. Such cases may simply not exist in their experience, leaving them with the problem of producing a judgment for a case they may feel they know nothing about. This may result in inconsistency.
Components of Cognitive Feedback

In this experiment, cognitive feedback was treated as a discrete entity in the sense that subjects either received feedback or they did not. The feedback presented was composed of three distinct elements - policy feedback, consistency feedback and interpersonal feedback (section 3.6, chapter 3). Since no attempt was made to study the impact of individual elements of feedback, we cannot make a definitive statement about the relative effectiveness of one form against another. The need to evaluate the differential impacts is an important one, however - for two reasons. First is the issue of cognitive load. The three elements of feedback added up to a considerable amount of information, which must have placed a load on the subjects' cognitive capacities. If the role of feedback is indeed to supplement a decision maker's cognitive limitations, it may be counter productive to actually add to his/her cognitive load by providing the feedback. On the other hand, if it can be determined that a more parsimonious repertoire (say, two elements) is just as effective, a more acceptable design principle can be postulated.

In a more fundamental sense, the lack of disaggregation of the elements leaves some questions unanswered. The elements of feedback are based on a two-step model of decision making - the formulation of a strategy and its implementation. Thus policy feedback helps the strategy formulation aspect, whereas consistency feedback aids in implementation. The question is: do we know that to be the case? An experimental design, such as the one outlined in table 5.1,
would have provided some answers in that direction (another facet of this problem is discussed in section 5.2.3).

**Table 5.1: A Revised Experimental Design**

<table>
<thead>
<tr>
<th>Policy and Consistency feedback</th>
<th>Block1</th>
<th>Block2</th>
<th>Block3</th>
<th>Block4</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Between Subjects</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Policy feedback</td>
<td>Block1</td>
<td>Block2</td>
<td>Block3</td>
<td>Block4</td>
</tr>
<tr>
<td>Consistency feedback</td>
<td>Block1</td>
<td>Block2</td>
<td>Block3</td>
<td>Block4</td>
</tr>
<tr>
<td>No feedback</td>
<td>Block1</td>
<td>Block2</td>
<td>Block3</td>
<td>Block4</td>
</tr>
</tbody>
</table>

**Within Subjects**

**How do Decision Makers Actually Use Feedback?**

Cognitive feedback studies adopt a "black-box" perspective: they assume that subjects somehow incorporate feedback into their decision processes and that this is reflected in the dependent variables, e.g., enhanced consistency. Beyond concepts gleaned from general information processing frameworks (such as Hogarth, 1986), we have little or no understanding of how the decision maker actually uses feedback in formulating and executing decision strategies. For example, do decision makers use feedback only when it serves to confirm their understanding? At what stages of the decision process does the decision maker observe the feedback? Is feedback used only initially, or continually? One might add that such studies require the use of process-tracing methods such as information boards or protocol
analysis. The use of such techniques among lens model researchers is virtually nonexistent (Brehmer (1979) even suggests that such methods should not be used)\textsuperscript{11}.

5.5 Summary

In a fundamental sense, this exercise sought to explore the effects of decision processes usually attributed to individuals, across multiple levels of aggregation. In addition to establishing the efficacy of cognitive feedback over multiple referent levels, this study marks an initial step in bridging the gap in our understanding of how factors that affect individual decisions work to influence decisions at other levels of analysis. The results contribute to our knowledge of decision aids create a basis for formulating systematic design principles in the realm of CSCW research. More work remains to be done in sharpening our understanding of the components of cognitive feedback, and how decision makers actually use feedback in judgment/choice. The results will lead to improved design principles in building group decision aids and multiuser interfaces.

\textsuperscript{11} In this experiments, subjects were asked at the end of the individual task in block 1 (A3.4, appendix) which feedback elements they had used in formulating their decisions. Most subjects indicated that they used all the elements (the performance of those who did not use all elements did not differ from the rest). Subjects were also asked to indicate the manner in which they used the information. Most responded with stock answers, such as: \textit{improve my decision rule}. One would hesitate to make much of these responses. They were elicited after the training. The validity of such post-experimental reports is open to question (Ericsson and Simon, 1980).
APPENDIX

EXPERIMENTAL MATERIALS AND INSTRUCTIONS PROVIDED TO SUBJECTS

A1 Description of the Experiment

A1.1 Written description given to subjects

* This experiment seeks to gain an understanding of how individuals and groups use information for making decisions. The task is to screen candidates in personnel selection for a company. The experiment has three parts.

  PART ONE is a training session, and requires decisions to be made individually.

  PART TWO involves making decisions on your own, as well as, in consultation with other members of your group.

  PART THREE is the debriefing stage of the experiment.

* You will be given a set of instructions at the beginning of each part.

* At the end of each part, you fill out a brief questionnaire.

* The experiment is run on a computer terminal, and at each stage, the bottom row prompts you on what to do next.

* Please follow the guidelines strictly. The system prompts, along with the instructions, will guide you at every stage.

* When in doubt, please ask the experimenter.

PLEASE TURN OVER FOR FURTHER INSTRUCTIONS
A1.2 Further clarification of the experiment by the experimenter

The task and purpose of the experiment were repeated orally in order to confirm that the written instructions had been fully understood. Questions by subjects were also answered.

A2 Training

A2.1 Written Instructions provided to subjects about Training

* Company XYZ needs to recruit personnel at the managerial level and has received applications of varying quality from several interested candidates.

* The number of applications is very large, and the recruitment committee has done an initial screening of the applications. The committee has used three criteria for evaluating candidates:

  Work Experience
  Test Scores
  Education

* Each candidate has been evaluated on each of these criteria, on a scale of 1-9 (1 implies very poor, 9 implies very good). Based on the ratings on these criteria, the candidate gets an overall rating of suitability - again, on a scale of 1-9. For example, a set of candidates may be rated in the following manner:

<table>
<thead>
<tr>
<th>Work Experience</th>
<th>Test Scores</th>
<th>Overall Education</th>
<th>Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>9</td>
<td>6</td>
<td>8</td>
<td>7</td>
</tr>
<tr>
<td>4</td>
<td>8</td>
<td>9</td>
<td>6</td>
</tr>
<tr>
<td>8</td>
<td>3</td>
<td>7</td>
<td>6</td>
</tr>
</tbody>
</table>

Thus, if a candidate got a 9 on work experience, 6 on test scores and 8 on education, the committee rated him/her as 7 (on an overall basis).

* The committee wants to train you in performing this task, i.e., screening candidates. The objective of the training is to ensure that your judgment matches their decisions.

1. Notes indicate that questions at this stage were quite cursory, e.g., "Is there any time limit?", or, "do we do this individually?"
* In order to do this, you are given profiles of several candidates, one at a time, on the criteria mentioned above.

* Your task is to give the candidate an overall rating of between 1 and 9 (1 implying very poor, 9 implying very good) - just as the committee did. After you rate a candidate, you are given the committee's overall evaluation of that candidate. Your task is to come as close as you can to the committee's evaluation in subsequent cases.

* This task is strictly individual, and you are not to talk to or consult with other members of your group.

* After completing the task, please answer the questions in the enclosure.

* Please make sure you understand the instructions. When in doubt, ask the experimenter.

YOU ARE NOW READY TO PROCEED WITH THE TASK

A2.2 Further clarification of training by the experimenter

The preceding instructions were repeated orally. The objective of the experiment - that the subjects had to match the committee judgments and thereby infer the committee's decision rule - was emphasized. Questions, (mainly from subjects seeking clarifications on exactly what they were required to do), were answered.

A2.3 Questions answered by the subject after completing training

1. Describe (in words, equation, etc) what decision rule you followed in making your estimates.
2. Distribute 100 points among the three criteria you used for reaching your overall estimate, in accordance with the importance you assigned them (the total should add up to 100).

   Work Experience    ______
   Test Scores        ______
   Education          ______
   Total:             ______ (out of 100)

3. How clear were the instructions regarding the task in this part of the experiment?

   Not at all Clear
   Clear          ______  ______
   Very Clear     ______  ______
   1              2          3  4  5

4. Your general comments:

   _______________________________________________________
   _______________________________________________________
   _______________________________________________________

A3 Experimental Stage

A3.1 Written instructions

* You are about to start part two of the experiment. The task is the same - evaluating candidates with respect to their overall suitability.

* The important difference is this: you no longer have the committee ratings of candidates. So, unlike PART I, you have no way of finding out if you came close to the committee ratings.

* The general sequence is this:

   1. You are given a set of TEN cases (i.e., profiles of ten candidates);

   2. YOU (i.e., individually) make estimates on these candidates.
3. You can then see how the other members of the group have rated them. You can revise your own estimates accordingly. Remember, this is an individual task. While you can observe the scores of the other members appearing on your screen (and take them into account in revising your estimates), you should not talk to or consult with other members of your group.

4. After individual members are done, the GROUP as a whole comes together and makes estimates on the same candidates. Since this is a group task, you can consult within the group or interact in whatever manner you choose.

* Instructions and prompts will guide you at every stage.
* If you have a question, ask the experimenter.

YOU ARE NOW READY TO PROCEED WITH THE TASK

A3.2 Further Clarifications provided by experimenter

The instructions were repeated orally and questions answered. Specifically, subjects were told:

(a) that they would be given forty cases, over four blocks, and,

(b) that they should try to emulate the committee's decision rule (that they have inferred from the training) as closely as possible (as opposed to stating their own opinions).

A3.3 Instructions on use of Cognitive Feedback
   (Given orally, only to cognitive feedback groups)

Cognitive feedback groups were given a brief training on the meaning and use of the different components of feedback. Specifically, they were shown the following transparencies through an overhead projector:

(i) The screen (figure 3.7).
(ii) Consistency feedback, i.e., displays of the consistency index and scores (figure 3.3).

(iii) Policy feedback, stacked bars denoting weights given to the cues (figure 3.2).

(iv) Feedback on agreement, i.e., bar charts denoting the matching indices (figure 3.4).

Each slide was accompanied by a brief explanation on the meaning and use of a particular display.

A3.4 Questions answered by subjects in the cognitive feedback condition only (after completing the individual task in block 1)

The subject was shown a display representing a particular display, followed by a set of questions. Four displays were shown, in the following order:

(i) Policy feedback (figure 3.2),

(ii) Consistency index (figure 3.3),

(iii) Matching index (figure 3.4), and,

(iv) Consistency scores (figure 3.3).

Each display was followed by these questions:

1. Please refer to the display shown above. It represents:

(a) an overall index of consistency, based on your estimates,

(b) agreement between the group members,

(c) importance accorded to the three criteria by your estimates.

(d) consistency scores, based on your estimates.
2. Did you use the information provided by the display in formulating your estimates?

Yes  No

If Yes, how?

________________________________________________________

Thus, subjects in the cognitive feedback condition answered four sets of questions, each pertaining to a particular display.

A3.5 Questions answered by all subjects (after completing part 2)

1. Describe (in words, equation, etc) what decision rule you followed in making your own individual estimates.

   ______________________________________________________
   ______________________________________________________
   ______________________________________________________

2. Describe (in words, equation, etc) what decision rule the group followed in making the group estimates.

   ______________________________________________________
   ______________________________________________________
   ______________________________________________________

3. To what extent, generally, do you agree with the group decision:

(a) With respect to overall estimates?

<table>
<thead>
<tr>
<th>Not at all</th>
<th>Somewhat</th>
<th>Very Much</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
</tbody>
</table>
(b) With respect to the importance attached to the following criteria:

**WORK EXPERIENCE:**

<table>
<thead>
<tr>
<th>Not at all</th>
<th>Somewhat</th>
<th>Very Much</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
</tbody>
</table>

**TEST SCORES?**

<table>
<thead>
<tr>
<th>Not at all</th>
<th>Somewhat</th>
<th>Very Much</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
</tbody>
</table>

**EDUCATION?**

<table>
<thead>
<tr>
<th>Not at all</th>
<th>Somewhat</th>
<th>Very Much</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
</tbody>
</table>

4. Was a group decision rule explicitly stated upfront (Y/N)? ____

5. If so, how did it change over time?

____________________________________________________________________

____________________________________________________________________

6. Who (if anyone) was the group leader, in your opinion? ____

7. Who, in your opinion, was the most influential member of the group? ____

8. Who, in your opinion, was the best member of the group? _____
9. Which of the following phrases best describe the atmosphere of the group (circle)?

* Tense
* Competitive
* Relaxed
* Cooperative
* Other phrases _______________________

10. To what extent did the computer:

(a) help you (individually) in making your own decision?

<table>
<thead>
<tr>
<th>Not at all</th>
<th>Somewhat</th>
<th>Very Much</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
</tbody>
</table>

(b) help the group (as a whole) reach a group decision?

<table>
<thead>
<tr>
<th>Not at all</th>
<th>Somewhat</th>
<th>Very Much</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
</tbody>
</table>

11. How clear were the instructions regarding the task in this part of the experiment?

<table>
<thead>
<tr>
<th>Not at all</th>
<th>Clear</th>
<th>Very Clear</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
</tbody>
</table>

12. Could the quality of the group decision have been improved? If so, how?

__________________________________________________________
A4 Post-experiment Debriefing

The debriefing was conducted at the end of the experiment. In two cases, the subject had to leave and returned the completed questionnaire a few hours later.

A4.1 Instructions

* This part is for the individual only. Given the following set of 10 cases, please enter your estimates. This is to be done alone. Please do not consult the other members.

A4.2 Judgment Task

How would you rate the following candidates (on 1-9)?

<table>
<thead>
<tr>
<th>Work Experience</th>
<th>Test Scores</th>
<th>Education</th>
<th>Overall rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>9</td>
<td>3</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>4</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>3</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>5</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>2</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>8</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>4</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>6</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>4</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>7</td>
<td>8</td>
<td></td>
</tr>
</tbody>
</table>
A4.3 Questions answered by the subject at the end of the session

1. What other criteria (besides the ones considered here) would you consider important in a task like this?

2. Have you, in the past, been associated with personnel selection at any level (Y/N)?

3. How interesting was the task of selecting personnel?

<table>
<thead>
<tr>
<th>Not at all Interesting</th>
<th>Somewhat Interesting</th>
<th>Very Interesting</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>5</td>
<td></td>
</tr>
</tbody>
</table>

4. How realistic, in your opinion, was the task?

<table>
<thead>
<tr>
<th>Not at all Realistic</th>
<th>Somewhat Realistic</th>
<th>Very Realistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>5</td>
<td></td>
</tr>
</tbody>
</table>

Please comment:

5. How serious were you in performing the task?

<table>
<thead>
<tr>
<th>Not at all Serious</th>
<th>Somewhat Serious</th>
<th>Very Serious</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>5</td>
<td></td>
</tr>
</tbody>
</table>

6. How serious was the group in performing the task?

<table>
<thead>
<tr>
<th>Not at all Serious</th>
<th>Somewhat Serious</th>
<th>Very Serious</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>5</td>
<td></td>
</tr>
</tbody>
</table>
7. How clear were the instructions generally?

<table>
<thead>
<tr>
<th>Not at all Clear</th>
<th>Clear</th>
<th>Very Clear</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5</td>
</tr>
</tbody>
</table>

8. How easy was the system to use?

<table>
<thead>
<tr>
<th>Not at all Easy</th>
<th>Easy</th>
<th>Very Easy</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>5</td>
</tr>
</tbody>
</table>

9. Please give us some information about yourself (in absolute confidence). At no time will your name appear in the results. The data will only be used in an aggregate statistical sense).

(a) Program enrolled in

Undergraduate: Freshman Sophomore Junior Senior

Graduate (MBA): Full time Part time

Graduate (other): M.S. Ph.D. Other (Specify)

(b) Sex

(c) Age

(d) Full time work experience (in years)

(e) How familiar are you with computers in general?

<table>
<thead>
<tr>
<th>Not at all Familiar</th>
<th>Somewhat Familiar</th>
<th>Very Familiar</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>5</td>
</tr>
</tbody>
</table>

(f) How many hours (per week) do you use computers?

10. Your general comments regarding the experiment:

__________________________________________________________________________________
__________________________________________________________________________________
REFERENCES


