

800

T  
AMW  
Kle  
604,583  
Nov. 82

DECISION-MAKING AND RESEARCH GAMING

Jonathan Hilary Klein

Royal Holloway College  
University of London

A thesis submitted for the degree of PhD

1982

RHC

604583 6



a30214 006045836b

ProQuest Number: 10097513

All rights reserved

INFORMATION TO ALL USERS

The quality of this reproduction is dependent upon the quality of the copy submitted.

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



ProQuest 10097513

Published by ProQuest LLC(2016). Copyright of the Dissertation is held by the Author.

All rights reserved.

This work is protected against unauthorized copying under Title 17, United States Code.  
Microform Edition © ProQuest LLC.

ProQuest LLC  
789 East Eisenhower Parkway  
P.O. Box 1346  
Ann Arbor, MI 48106-1346

## ABSTRACT

The thesis is concerned with research in the field of human decision-making, concentrating on techniques of gaming for the pursuit of this research.

Following an introduction to the work and a statement of the research programme as it was initially conceived, some current ideas in gaming are investigated. The Superior Commander system of game control is introduced. The content of research games is discussed, and the Organisational Control Game, a board war game designed for research, is described. It is shown that the Organisational Control Game and Superior Commander system successfully meet the requirements for a useful research game and gaming methodology.

A detailed literature survey of the psychological secondary task technique for assessing mental processing load is presented. It is noted that the technique might be extended to the study of tasks which have a large problem-solving component. A secondary task experiment on such a task, a chess problem task, is described. It is demonstrated that the secondary task approach can provide techniques for the investigation of complex problem-solving and decision-making tasks.

A series of plays of the Organisational Control Game, in which the players had had previous military experience, is described. These games are compared with an earlier series of games, in which the players were students. Certain differences in playing style are identified.

R.N.C.  
LIBRARY

The research programme is re-examined, and modifications to it are described. The need for a technique for elucidation and examination of an individual decision-maker's perceptions of his decision-making environment is identified. The technique of cognitive mapping is shown to be suitable for this purpose. A cognitive map analysis of a series of games in which the players were serving army officers is presented.

### NOTE

Most of the work described in this thesis was undertaken while I was a postgraduate student in a research group at Royal Holloway College. The results presented here represent the joint efforts of the members of the group. Four papers describing this work have been published:

1. D.F. COOPER & J.H. KLEIN (1980) Board wargames for decision making research. European Journal of Operational Research, 5, 36-41.
2. D.F. COOPER, J.H. KLEIN, M.R.C. McDowell & P.V. JOHNSON (1980) The development of a research game. Journal of the Operational Research Society, 31, 191-193.
3. J.H. KLEIN & D.F. COOPER (1981) Assessing problem complexity. European Journal of Operational Research, 6, 243-247.
4. J.H. KLEIN & D.F. COOPER (1982) Cognitive maps of decision-makers in a complex game. Journal of the Operational Research Society, 33, 63-71.

These papers are reproduced in an appendix at the end of the thesis.

In addition, work described here has also been presented at a number of scientific conferences:

1. International Council of Psychologists, Regional Conference on "Resolution of Conflict: the Psychologist's Role", University of Surrey, 16-19 July 1978.
2. Third European Congress on Operations Research, Amsterdam, Netherlands, 9-11 April 1979.
3. Operational Research Society Annual Conference, University of Stirling, 4-7 September 1979.
4. Fourth European Congress on Operations Research, University of Cambridge, 22-25 July 1980.
5. Young O.R. Conference, Coventry, 2-4 December 1980.

### ACKNOWLEDGEMENTS

Much of the work described in this thesis is experimental. I should therefore like to thank the experimental subjects: the students at Royal Holloway College who participated in the secondary task experiments and in the first series of games; the local government officers who participated in the second series of games; and the army officers who volunteered for the third series of games, not forgetting their commanding officer, Lieutenant Colonel P.A. Horwood, who volunteered them.

I should like to thank, for their advice and encouragement, colleagues at Royal Holloway College, particularly Dr Yannis Darzentas, Dr Godwin Ovuworie (now at the University of Benin), and Mr Peter Storr (now with North Thames Gas). I should also like to thank friends at the University of Sussex, for, in particular, first interesting me in this area of research: Dr Malcolm Dando (now at the University of Bradford); Dr Chris Huxham (now at the University of Aston); and Dr Richard Sharp (now with the British Airports Authority).

I should like to thank Ms Sue Allcock, who typed the thesis, and Ms Margaret Dixon, for, in particular, preparing the transcripts of the third series of games to a uniformly high standard from recordings of rather variable quality.

I should like to thank Dr Philip Johnson (now of the Admiralty Surface Weapons Establishment, Portsmouth) who was a member of the research group at Royal Holloway College for a year, and contributed much to the running of the experimental games.

I wish to thank, in particular, three colleagues whose high

involvement with my work and whose complementary viewpoints of it must account in large part for any worth it may possess. Professor Ken Bowen, as Head of Research at the Defence Operational Analysis Establishment, West Byfleet, was responsible for the funding of the work; later, as an associate at Royal Holloway College, his advice and judgement was always valuable. Dr Dale Cooper (now at the University of Southampton) has been closely associated with every step of the work, both as a senior colleague and as an "unofficial" supervisor; I should particularly like to acknowledge, with gratitude, all that he has taught me about how to do research. Professor Coulter McDowell, of the Department of Mathematics, Royal Holloway College, was my supervisor throughout the project; I am extremely grateful for his encouragement, advice, strategic guidance and faith.

The work described in this thesis was supported in part by the Ministry of Defence (Defence Operational Analysis Establishment, West Byfleet) under Contract Number ER3/9/4/2088/012 to Professor M.R.C. McDowell.



## TABLE OF CONTENTS

Abstract		2
Note		4
Acknowledgements		6
Table of contents		8
List of tables		13
List of figures		15
<u>Chapter 1. Introduction</u>		18
1.1	Introduction	18
1.2	Decision-Making Research	19
	Decision-Making in Crisis	21
	Definition of Crisis	23
	Hypotheses Concerning Crisis Behaviour	25
1.3	A Gaming Approach	27
	Processing Load Imposed by a Game Scenario	29
1.4	Crisis and Stress	30
1.5	A Research Programme	32
1.6	Plan of the Thesis	34
<u>Chapter 2. The Use of Games</u>		37
2.1	Introduction	37
2.2	Classification of Games by Purpose	39
2.3	Classification of Games by Structure	46
2.4	The Intelligence Man System of Game Control	52
2.5	The Superior Commander System of Game Control	56

<u>Chapter 3. Board War Games for Decision-Making Research</u>	63
3.1 Introduction	63
3.2 Experiments with a Tactical War Game	65
Recording a Tactical Game	66
Intelligence Man System Applied to Panzer Leader	68
3.3 Requirements for a Research Game	70
Requirements for a Game	71
Requirements for a Methodology	73
3.4 A Board War Gaming System for Research	76
The Organisational Control Game	76
Use of the Superior Commander System	80
Initial Experience with the Superior Commander System	83
3.5 War Game Design	84
<u>Chapter 4. The Organisational Control Game</u>	86
4.1 Introduction	86
4.2 Overall Structure of the Game	86
4.3 Mode of Play	90
Rules of Play	91
Dummy Task	92
Research Phase	93
4.4 Scenarios	94
Scenario 1 - A Defence Scenario	96
Scenario 2 - An Advance-to-Contact Scenario	100
4.5 Playing the Game	104
4.6 Notes on the Development of the Game	109
Design of Scenarios	110

	Modifications to Scenario 1	111
	Alterations to Game Rules	112
	Dummy Task	113
	Other Notes	114
<b>Chapter 5.</b>	<b><u>Secondary Task Methods for Assessing Mental Processing Load - A Literature Survey</u></b>	<b>117</b>
5.1	Introduction	117
5.2	Secondary Tasks	118
5.3	The Capacity Model of Human Information Processing	121
	The Single-Channel Limited Capacity Model	122
	Modifications to the Single-Channel Model	124
5.4	Secondary Tasks and Limited Capacity	125
5.5	Loading Task Experiments	128
	Memory	129
	Learning	132
	Interference Effects	133
5.6	Subsidiary Measuring Task Experiments	135
	Vehicle Driving	135
	Mental Processes	137
	Other Examples	139
5.7	Discussion	141
5.8	Extension of Secondary Task Application	146
<b>Chapter 6.</b>	<b><u>A Secondary Task Experiment to Assess Problem Complexity</u></b>	<b>151</b>
6.1	Introduction	151

6.2	The Secondary Task Technique	153
	The Subsidiary Measuring Task Technique	153
6.3	The Nature of Complex Decision Problems	154
6.4	A Secondary Task Experiment	158
	The Primary Task	158
	The Secondary Task	162
	Results	163
6.5	Conclusions and Discussion	168
	Improvements in Sensitivity	169
	A Foot-Tapping Secondary Task	171
	Concluding Remark	172
<u>Chapter 7. Experience with the Organisational Control Game</u>		174
7.1	Introduction	174
7.2	Games Played with Emergency Planning Officers	175
	Experimental Procedure	175
	A Counter Replacement Test	176
	Perception of Options	177
	Experienced and Inexperienced Commanders	181
7.3	A Re-Examination of Research Strategy	187
7.4	Games Played with Army Majors	193
	Game Communications	195
	Initial Impressions of Communication Transcripts	196
<u>Chapter 8. A Cognitive Map Analysis of the Organisational Control Game</u>		200
8.1	Introduction	200

8.2	Cognitive Mapping	202
	Cognitive Map Literature	208
8.3	Coding Cognitive Maps	212
8.4	Results	217
	Concepts and Relationships	217
	Differences in Detailed Interpretation	225
	Confidence of the Players	225
	Sequence Effects	229
	Dynamic Aspects	229
8.5	Discussion	233
 <u>Chapter 9. Summary and Conclusions</u>		 237
9.1	Introduction	237
9.2	A Summary	237
	Gaming Methodology and Game Design	237
	Secondary Task Methods	241
	Decision-Making Research	242
9.3	Final Discussion	246
 References		 252
 <u>Appendix : Published Material</u>		 271
1.	Board Wargames for Decision Making Research	272
2.	The Development of a Research Game	278
3.	Assessing Problem Complexity	281
4.	Cognitive Maps of Decision-Makers in a Complex Game	286

## LIST OF TABLES

- 2.1. Classification of games by structure.
- 4.1. Specimen game timetable.
- 6.1. Problems ranked in order of increasing difficulty.
- 7.1. Player's perceptions of number of options open to  
Blue and Red forces.
- 7.2. Ratio of Blue to Red options.
- 7.3. Ratios of Blue options in different rounds.
- 7.4. Ratios of Red options in different rounds.
- 7.5. Experimental design for the "Major" series of  
games.
- 7.6. Transcript of communications for Player C in the  
first round of Scenario 1.
- 8.1. Text for Player C, in the first round of Scenario 1,  
with concepts underlined and labelled.
- 8.2. Concepts and links between concepts for Player C,  
Scenario 1.

- 8.3. Concepts for Player A, Scenario 2.
- 8.4. Concepts for Player C, Scenario 2.
- 8.5. Concepts, links and densities of the cognitive maps.
- 8.6. Subjective ranks, goal-increasing scores and ranks, and description-speculation scores and ranks for the two scenarios.
- 8.7. Sequence effects on subjective estimates of the players' confidence.
- 8.8. Concepts for Player C, Scenario 1, second round.

## LIST OF FIGURES

- 2.1. Classification of games by purpose.
- 2.2. The three-room war game.
- 2.3. The situation which the players in the Intelligence Man "Diplomacy" games perceived.
- 2.4. The real Intelligence Man "Diplomacy" game.
- 2.5. The basic hierarchy of decision-makers required for the Superior Commander system.
- 2.6. The Superior Commander system, in full, as seen by the experimental player.
- 2.7. The Superior Commander system, in full, as seen by the experimenter.
- 3.1. The Superior Commander system applied to the hierarchy of the Organisational Control Game.
- 4.1. Structure of the friendly hierarchy in which the experimental subject considers himself to be playing.
- 4.2. The fictitious Republic of Arachide.



- 4.3. The setting for Scenario 1.
- 4.4. The subject's Divisional area in Scenario 1.
- 4.5. The setting for Scenario 2.
- 4.6. The subject's Divisional area in Scenario 2.
- 5.1. The capacity model and secondary task methods.
- 6.1. The loads due to two primary tasks.
- 6.2. Solution processes for simple and complex problems.
- 6.3. Chess problem 6 in the experimental set.
- 6.4. Chess problem 12 in the experimental set.
- 6.5. Total proportional secondary task error rate plotted  
against the number of incorrect answers.
- 8.1. Two concepts linked by a positive relationship.
- 8.2. A "fictional" cognitive map.
- 8.3. Cognitive map of Player C in Scenario 1.
- 8.4. Cognitive map of Player A in Scenario 2.

- 8.5 Cognitive map of Player C in Scenario 2.
- 8.6. Alternative views of the same objective situation.
- 8.7. Cognitive map of Player C in Scenario 1, second  
round.

## CHAPTER 1

### INTRODUCTION

#### 1.1 Introduction

This thesis is concerned with research in the field of human decision-making, concentrating on techniques of gaming for the pursuit of this research. It describes work carried out in the course of a programme of research intended, firstly, to develop appropriate gaming techniques, and, secondly, to apply these techniques to the investigation of human decision-making. These two aspects of the programme are reflected in the contents of this thesis. Thus, many of the discussions and conclusions presented here are concerned with gaming techniques and methodology. However, I have also discussions and conclusions to present concerning human decision-making, particularly in the later chapters of the thesis.

Most of the work to be described here (including all of the experimental work) was carried out in a two-year period during which I was a postgraduate student in the Department of Mathematics at Royal Holloway College (University of London), working on a "Crisis Gaming" project funded by the Ministry of Defence. The progress of this project had been described in three progress reports (Cooper et al, 1978; McDowell et al, 1979a, 1979b), and, in addition, we have published a number of papers describing the work in the Operational Research literature (Cooper, 1978, 1979; Cooper & Klein, 1980; Cooper et al, 1980; Klein & Cooper, 1981, 1982). The four papers with which I was associated are reproduced in the Appendix. In this thesis I have been able to substantially develop and explore much of what was presented in these papers.

This chapter introduces the subject matter and provides a suitable background to the work to be described in subsequent chapters. In particular, I want to explain as fully as possible the rationale of the work, so that it will be clear why our research programme took the course that it did. Research programmes do not come into existence spontaneously, or, at least, they should not: a number of decisions as to areas of interest and direction of research should be taken, and, if necessary, retaken. A programme of research into decision-making should be no exception to this rule.

I shall begin by consideration of the general area of decision-making research, and proceed to describe our initial interest in decision-making in crisis. This will lead us to an appreciation of the usefulness of games as research tools in this area. From here I shall go on to consider how the stress content of crisis might be simulated in a laboratory environment, and thus arrive at a statement of our research programme as we conceived of it initially. This programme, although subsequently modified, led us to start investigating particular areas of research gaming and decision-making science. I shall show how we intended to operationalise this programme. What happened when this operationalisation was started forms the contents of the subsequent chapters, which are briefly set out and described at the conclusion of this chapter.

## 1.2 Decision-Making Research

Although precise definitions of Operational Research vary greatly, most workers in the field would agree that its main aim must be to aid decision-makers. Traditionally, the main area of attention has

been the definition and articulation of techniques directed towards the resolution of reasonably well defined problems. There is little doubt that Operational Research has been extremely successful in this area. Problems that only a few decades ago might have been considered to be extremely taxing to resolve satisfactorily are today often regarded as relatively routine.

Yet, in the last few years it has been increasingly recognised that the service provided by Operational Research to decision-makers is not as full as is desirable. While great progress has been made in techniques for the resolution of well-defined problems, comparatively little has been done to provide concrete assistance for decision-makers faced with more complex and ill-defined problem situations. Ackoff (1979) has appropriately described such situations as "messes", and Radford (1978) has referred to the unpredictability and rapid change inherent in the environments of such problems as "turbulence".

Techniques for the resolution of well-defined problems have largely been provided by mathematics, often inspired by mathematical applications in the physical or life sciences. Indeed, a predominantly mathematical approach is always likely to be fruitful when a problem can be well-defined, because well-defined problems can generally be formulated in mathematical terms.

In dealing with ill-defined messes, the situation is somewhat different. Here, mathematical descriptions are, by definition, poor

fits, and, when tried, appear crude and uncomfortable. To assist decision-makers under such circumstances, it is necessary to search for frameworks other than models with a high mathematical content, predominantly inspired by methods appropriate to the "harder" sciences.

In fact, what is required is an Operational Research based on a "science" of decision-making (as argued by, for example, Dando et al, 1977). Since this state of affairs does not at present exist, it is impossible to state precisely what form this science would take, but it seems highly probable that it would be a "human" science, deriving much of its content from such disciplines as psychology and sociology. This, of course, does not preclude anything else that might seem appropriate.

It seems clear that to develop a decision-making science, some study of decision-making is required. It was in this spirit that the work to be described in this thesis was undertaken. We were interested in a practical study of decision-making, under experimental (that is, suitably controlled) conditions. Further, our general aim in such a study would be to identify ways in which human decision-making might be improved.

### Decision-Making in Crisis

An area in which human decision-making is at its most frail and susceptible to failure is decision-making under conditions of

crisis. Much work has been done (particularly in the field of political science) on the behaviour of decision-makers in crisis. This work ranges from detailed studies of actions in specific crises of various types to broader generalisations which describe the type of effects that might be caused by crisis. However, the linking of these two areas of research tends to be tenuous. The derivation of the broader perspectives from the specific instances tends often to be vaguer and less rigorous than is ideal. Consequently, the impact of the broader perspectives is weakened. Reviewing concepts and theories pertaining to behaviour in crisis, Robinson (1972) states (p.27): "there is no such thing as a theory of crisis or even theories of crisis". Later (p.35) he continues: "students of crisis have relied heavily on ad hoc theorizing and few have placed their investigations in any fundamentally theoretical context. Until efforts are launched toward systematizing and integrating theories, knowledge will remain relatively superficial and anecdotal". In similar vein, Milburn (1972) notes a tendency to formulate "descriptive and empirical hypotheses about crisis" as opposed to "prescriptive hypotheses" - a tendency from which he intends to move away (p.262).

A linking of the ideas met in the "case-study" type of work and the "general description" type of work might be attained by experimental investigation of crisis behaviour. Comparatively little has been done, however, to investigate decision-making in crisis in controlled experiments. Given the initial assumption that crisis situations do influence decision-making in certain definable ways, an experimental approach to the study of models of decision-making

which attempt to account for such phenomena would appear to be valid and useful.

Thus, we initially formulated the aim of our work as being: to develop a suitable methodology and we use it to test models of decision-making. The models would take the form of descriptions of decision-making behaviour, particularly under conditions of crisis. More specifically, we would hope to study a decision-maker's perceptions of situations and his assessment of the importance of various factors connected with his decisions. Useful models of decision-making should be able to account for changes in many aspects of behaviour due to crisis, and it should be possible to examine the models by analysis of the changes.

#### Definition of Crisis

It will be useful to have in mind a working definition of crisis. I shall present two complementary pictures of crisis, which each capture a different flavour of the area of interest. The first is due to Harris (1974, p.3), who defines a system to be in a state of crisis "when it, or one of its subsystems, is unable to cope with its intended sphere of activity". This "systems" definition can be cast in a form more appropriate to a single human decision-maker. Effectively, crisis can arise under one or more of the following conditions (Cooper, 1977):



1. There is no way of reconciling any of the options considered by the decision-maker with all of his aims.
2. The quantity of information is too great for the decision-maker to process adequately in the time allowed.
3. The decision-maker does not have appropriate information processing strategies available.

These three aspects of crisis can be seen to correspond roughly to the three dimensions of threat, time, and surprise that Hermann (1972a) identifies in our second definition of crisis, drawn from the field of political science. In full, Hermann defines (p.13) a crisis to be a situation that:

1. Threatens high-priority goals of the decision-making unit.
2. Restricts the amount of time available for response before the decision is transformed.
3. Surprises the members of the decision-making unit by its occurrence.

Clearly, the two definitions are by no means identical, but both emphasise the threat to aims imposed by the situation, the limited time available to achieve resolution, and the unpreparedness of decision-makers to deal with the situation. With the picture of

crisis provided by these two definitions, we can proceed to consider the behaviour of decision-makers in crisis.

### Hypotheses Concerning Crisis Behaviour

As a source of testable hypotheses concerning decision-making in crisis, we turn to Milburn (1972), who has collected a set of 24 such hypotheses. Milburn is not the only worker who has attempted to formulate his ideas in this way - indeed, the book in which Milburn's hypotheses appear (Hermann, 1972b) boasts a total of 311 "empirically testable propositions" (p. 304) by various authors. However, inspection shows many of Milburn's to be oriented towards the behaviour of individuals (rather than groups) and, generally, to be formulated in terms which appear to offer relatively simple, operational testing measures. I should emphasise, though, that it is not simply a question of picking the "easiest" hypotheses first, but rather of picking those which offer a framework upon which future work could be based. Note, also, that in focussing attention on Milburn's work we follow a course suggested by Sharp and Dando (1977).

Milburn drew a distinction between "research findings relevant to scientific theory and specific knowledge applicable to policy problems" (p. 271). He explains: "in the world of the scientist, descriptive hypotheses help to approximate certain theoretical variables in concrete operational terms which the scientist may observe in the laboratory or field to note what happens. On the other hand, a decision-maker employs prescriptive hypotheses as guides to his actions, as recipes for success or failure". Consequently, "hypotheses concerning crisis management are not like conventional scientific statements; rather, they are imperatives - recipes for action". Milburn formulates his hypotheses by first of all asking a

question "which an experienced crisis manager might reasonably and legitimately seek to answer before, or at the start of, a crisis" (p. 272). This is followed by a hypothesis "from empirical research". Finally, a decision rule is presented based on the (presumably valid) hypothesis. I present, in full, two of Milburn's hypotheses, chosen for their particular amenability to testing.

Hypothesis 13 (p.274)

- Question : Are our perceptions of the situation highly definite and fixed?
- Hypothesis : Crises increase a tendency toward rigidity of perception and thought.
- Decision Rule : Avoid simplistic renditions of the problem, such as those cast in terms of the capacity of the other side without reference to their attitude or intentions.

Hypothesis 17 (p.275)

- Question : Does the crisis appear to give the other side more flexibility and alternatives, while restricting our own? As compared to the other side, do we seem to be losing our ability to control events?
- Hypothesis : In a crisis one's own alternatives appear to contract while the other side's options seem to grow.

Decision Rule : Recognise that the other side is probably experiencing similar feelings of relative loss of control and limited alternatives. By your own action give the other side a range of choices; do not force them into a corner. Attempt to think through or simulate what the opponents may regard as limitations on their behaviour. In order that the other side may seek and review alternatives, do not demand responses in substantially less time than they normally require for processing decisions.

As well as presenting us with hypotheses which might be tested, Milburn's formulations provide an obvious extension. Having validated (or otherwise) a given hypothesis, we could then go on to test the usefulness of the appropriate decision rule, by running similar tests on actors who have been "indoctrinated" with the rule.

### 1.3 A Gaming Approach

As I have already noted, comparatively little controlled experimentation has been carried out on decision-making in crisis, or, indeed, on decision-making in general. Most research has been based on the study of real-life decision-making. This work is unquestionably valuable, but such a research methodology quite clearly does not have the scientific power of a controlled experiment.

The experimental approach to decision-making is research gaming. A research game sets up an artificial decision-making environment in

which the decision-maker may act: thus he becomes an experimental subject. By suitable design of the game the decision-making environment may be controlled to the required degree, and the decision-maker's behaviour recorded in such detail as required. Although when stated so concisely, the design and running of a research game sounds extremely straightforward, in practice it is not. In fact several issues have to be dealt with, such as: just how, in practice, an experimenter achieves a suitable degree of control over the game environment; how an experimenter can be sure that his subject is behaving as a real decision-maker, rather than merely as a subject in an experiment; and what aspects of behaviour should be studied. Discussion of these issues, and others, will form much of the content of subsequent chapters in this thesis.

Having decided to adopt a gaming approach, there is still a wide choice as to precisely what type of game might be appropriate. One dimension along which games might be measured is the degree of complexity of the decision-making environment. This can range from extremely simple to extremely complex. At the simple end of the spectrum are matrix games such as those described by Pruitt and Kimmel (1977). Matrix games are characterised by a very small number of options open to the experimental subject (typically two), and a very small number of possible outcomes to the game (typically four), the whole game being presented in simple and rather clinical mathematical form. Complex games are typified by war games such as described by Shephard (1963), and are characterised by large numbers of options and outcomes (often not really enumerable objectively). In between the two extremes can be found a few interesting hybrids (see, for example, Enshoff, 1971), but it is probably true to say that most research gaming has tended towards the extremes rather than

deliberately setting out to occupy the middle ground.

We decided that since we were interested in real decision-making, the complex type of game, with a rich decision-making environment, was more suited to our requirements. It is by no means clear how one would, in general, extrapolate the results of a simple matrix-type game to decision-making in the real world (but see Emshoff, 1971). A complex game would, ideally, provide a rich artificial world similar enough to the real world to permit "real" decision-making in the game (though the actual situation is far less straightforward than this argument suggests, as will be seen in subsequent chapters).

Having decided on a complex research game approach, it was then decided that the game should be a war game. Other types of complex reality-simulating game would have been equally appropriate: for example, business gaming. We chose war gaming because one of us (Dr. Dale F. Cooper) had previous military experience, and also because, due to our funding by the Ministry of Defence, we had contact with workers with some war gaming experience and also the option of using serving military officers as players if necessary (as we subsequently did).

#### Processing Load Imposed by a Game Scenario

Considering once again the definition of crisis due to Harris (1974), we note that attention is drawn to the information processing load imposed on the decision-maker by the crisis situation. The idea that a decision-maker has only a limited processing capacity seems reasonable (as will be shown in later chapters, there is considerable hard evidence for a capacity model of human information processing);

it has received a recent exposition in the Operational Research literature by Sharp and Dando (1979) who use what they call a "decision resource" model to account for the effects of crisis on decision-making. They regard crises as situations in which the limited processing capacity is exceeded, and thus new processing strategies must be adopted. However, the analysis provided by the new strategies will necessarily be cruder than that provided by the normal strategies, giving rise to effects such as those hypothesised by Milburn (1972).

We decided that the information processing load imposed on our experimental players by our game might well be one of the variables we would need to control. In particular, we envisaged that at some point in our experiments we might wish to run different game scenarios which we knew to impose the same processing load, or which could be ranked according to increasing processing load. Accordingly, we decided to investigate the use of a psychological technique, the secondary task technique, as a means of measuring processing load associated with a task such as playing a game scenario.

#### 1.4 Crisis and Stress

A different model of the decision-maker to the limited capacity model that may also account for crisis behaviour is the arousal model. The argument goes that a crisis induces psychological stress in a decision-maker, and this stress will cause many of the observed crisis effects. Note that this model in no way invalidates the limited capacity model - both effects may occur in crisis, and might well be roughly additive.

There is good evidence to suggest that stress will affect the

performance of a complex task such as decision-making. Poulton (1971) discusses stress effects on skilled performance and decision-making. Essentially, stress can be directly related to the level of psychological arousal. For a given task, the relationship between quality of task performance and arousal levels follows an inverted U-shaped curve - that is, there is an optimum arousal level for a given task below or beyond which the quality of performance diminishes.

In crisis, we would be interested in hyperarousal caused by stress due in particular to the aspects of crisis associated with threat to high-priority goals. To bring about this type of emotional stress in a gaming situation would be difficult to carry out ethically - in any normal game, goals relating to the game decision-making environment would not be sufficiently important to induce severe hyperarousal. Some kind of deception in which "real" goals (external to the game decision-making environment) were apparently severely threatened would probably be necessary. We were not anxious to become involved in this type of experimental work.

However, arousal can be artificially increased by various techniques, one of which is exposure of the subject to various types of noise. It was our intention to use loud, continuous "white" noise as a stressor, and to compare decision-making behaviour in stressed and unstressed conditions, looking once again for effects of the type described by Milburn (1972). We carried out a considerable amount of work in preparing equipment for experiments using noise as a stressor, including some consideration of methods of actually measuring physiological arousal in our subjects. However, we did not eventually conduct any experiments using noise, and consequently no further



discussion of this aspect of the work will be presented in this thesis. For a study of the relationship between decision-making and stress I would refer the reader to Broadbent (1971); a more physiological approach to arousal is provided by Duffy, (1962). Broadbent (1957) has presented a brief overview of the effects of noise on behaviour, and some interesting comments on the measurement of stress have been made by Parrot (1971).

### 1.5 A Research Programme

I have now briefly described our initial thoughts concerning the research in which we were interested. It is now possible to give a concise statement of our initial research programme.

Our major aim was to develop a research game suitable for testing hypotheses concerning decision-making in crisis. Having developed such a game, we would proceed with hypothesis testing in the following manner:

1. Select a hypothesis.
2. Test hypothesis using game.
3. Formulate appropriate decision-making rule.
4. Test decision-making rule using game.

Measures of decision-making would vary according to the hypothesis under consideration, but it was felt that such measures as the number of options perceived by the player as being open to both himself and his opponent would be of particular importance.

Against this major programme, two subsidiary programmes would be developed for later insertion into the main programme. The first was the development of a secondary task method of measuring the processing load imposed by a game scenario. This was required because we envisaged, in particular, running individual players through different scenarios of similar processing load under stressed and unstressed conditions, and, more generally, we felt that some method of ranking scenarios in terms of increasing processing load would be enlightening as far as scenario design methods were concerned. The second subsidiary programme was the development of an understanding of the stress aspect of crisis, and in particular the preparation of a technique for using noise as a stressor.

This, then, was how we viewed the programme at the outset; however, the programme did not work out this way. As we progressed we found it more sensible to revise our aims and follow a different course in our research. As will become clear in later chapters, many of our ideas concerning the nature of decision-making needed re-examination, and ultimately we found it more profitable to take our research in a slightly different direction.

Specifically, having designed and developed a working research game, we found it more appropriate to use it to develop methods of describing decision-making rather than follow our original intention of hypothesis testing. Consequently the secondary task method of measuring the processing load imposed by game scenarios was not required in our main programme, though, as will be seen, the secondary task work was, in itself, extremely interesting and provided results of some worth. The induction of arousal by noise, and the associated

testing of decision-making behaviour under stressed and unstressed conditions, was not ultimately pursued.

#### 1.6 Plan of the Thesis

I conclude this chapter by outlining the plan of the rest of this thesis, considering each subsequent chapter briefly.

Chapter 2, "The Use of Games", is a tour of some current ideas in gaming. It is not intended to be a review of gaming literature. It simply identifies points that will be of relevance in later chapters. Considerable use is made of classifications of games by purpose and structure as formulated by Bowen (1978), and I consider the Intelligence Man system of game control (Sharp & Dando, 1977) before introducing the Superior Commander system (Cooper, 1978, 1979) which we used.

Chapter 3, "Board War Games for Decision-Making Research", is concerned with the content of research games. It is derived largely from our early experience with commercial board war games and the development of our own Organisational Control Game. I derive a set of requirements for a useful research game and gaming methodology, and show how the Organisational Control Game and Superior Commander system meet these requirements. Much of the content of this chapter has been published (Cooper & Klein, 1980).

Chapter 4, "The Organisational Control Game", describes

in detail the Organisational Control game, and should give the reader a thorough idea of the nature of the game and the complexity of the decision-making environment.

Chapter 5, "Secondary Task Methods for Assessing Mental Processing Load - A Literature Survey", is a review of literature describing development of the psychological secondary task technique of mental processing load assessment. I describe the capacity model of human information processing, and show how a secondary task may be used either as a loading task or a subsidiary measuring task. Finally, I indicate possible extensions of the technique to study tasks which have a large problem-solving component.

Chapter 6, "A Secondary Task Experiment to Assess Problem Complexity", describes a secondary task experiment we conducted on a task with a large problem-solving component, namely a chess problem task. I demonstrate that the secondary task approach can provide techniques for the investigation of complex problem-solving and decision-making tasks. An account of this experiment has been published (Klein & Cooper, 1981).

Chapter 7, "Experience with the Organisational Control Game", returns to our main programme and describes a series of games played with players who had had previous military experience. I comment on their decision-making and compare it with the behaviour of student players (an account of this comparison, by Cooper et al, 1980, has been published). I then proceed to re-examine our research programme, and show how we found it necessary to alter our research aims. The chapter concludes

with a description of games played with a number of serving army Majors.

Chapter 8, "A Cognitive Map Analysis of the Organisational Control Game", uses the psychological technique of cognitive mapping to analyse the games described at the end of the previous chapter. It is shown that cognitive mapping is a suitable technique for elucidation and examination of the individual decision-maker's perceptions of his decision-making environment, and may begin to provide the framework of a language for discussion of decision-making. Much of the content of this chapter has been published (Klein & Cooper, 1982).

Chapter 9, "Summary and Conclusions", presents concisely the results and conclusions of the work described in the previous chapters. This chapter emphasises in particular the use of techniques such as cognitive mapping as a basis for developing a decision-making science appropriate to Operational Research.

The reader may find it useful to bear in mind a convenient classification of the chapters in the thesis. Chapters 1,2,3 and 4 are primarily concerned with gaming methodology and game design. Chapters 5 and 6 are devoted entirely to secondary task methods of measuring mental processing load. Chapters 7 and 8 are primarily concerned with decision-making research. These three areas of work are summarised in Chapter 9.

## CHAPTER 2

### THE USE OF GAMES

#### 2.1 Introduction

In the previous Chapter I outlined the aims of our research programme as we initially formulated them and noted the integral part played by research gaming in the programme. In this chapter I shall consider the use of games in research and related fields and discuss various points that will be of relevance in later chapters.

This chapter is not intended to be a complete, or even partial, review of gaming literature. Such a review is far beyond the scope of this thesis, and would be of little relevance. A comprehensive survey of gaming in the United States of America, with particular emphasis on military gaming, has been undertaken (Shubik & Brewer, 1972a, 1972b; Shubik et al, 1972; Shubik, 1975a, 1975b; Brewer & Shubik, 1979); as far as I know no similar survey on such a large scale exists for United Kingdom games, though Gibbs (1974) provides a British-based register of games and simulations. A United Kingdom gaming survey might well fill a useful "directory" function, but, since games are used in such a wide variety of contexts and for such disparate purposes, it does not seem to me that it would be able to offer much beyond that.

A working definition of a game will be useful in what follows. As I shall have occasion to do often in the chapter, I follow Bowen (1978, p.3-4) in stressing that a game will be played for a purpose. The purpose will be "to develop some understanding of the way in which purposeful behaviour can effect the situations occurring in the game",

where the situations in some sense model real situations in the real world. A game is thus some type of simulation of the real world where the aspect that is of particular interest is the interaction between the purposive behaviour of the player or players and the rest of the game world. As a rule, the purposive behaviour is supplied by real, purposive human beings. Situations in which the behaviour is supplied by mechanical models of deterministic or probabilistic nature can be regarded as special cases. (This contrasts with the view often conveyed in Operational Research texts of games with real players being special cases of entirely mechanical simulations.)

It is worth making two points here. Firstly, gaming as defined above has little to do with that branch of mathematics known as "game theory", and, as far as the work presented in this thesis is concerned, there is no connection at all. It was hoped at one time that game theory would advance in some way the understanding of human behaviour, but this hope was by and large unfulfilled, as modern game theory texts might reluctantly concede (for example, Jones, 1980, p.16). Extensions and developments of game theory may be valuable, however, in the construction of worthwhile decision aids - two such developments which I regard as extremely worthwhile are hypergames (Bennett, 1980) and metagames (Radford, 1975). It is interesting to note, in passing, that game theory sometimes appears appropriate to the study of (non-human) animals (see for example, Maynard Smith, 1974). A game theoretic approach to modelling behaviour may be appropriate wherever the organism in question is unable to consciously reflect about its behaviour. Conscious reflection, which implies, for example, being able to conceive of and compare alternative futures, is an ability which most creatures most of the time seem not to possess.

The second point is that, despite the possibly frivolous overtones that the word "game" implies, games are to be taken seriously. Unfortunately, this sometimes does not happen - most of us involved in gaming have met people who refuse to believe that anything that could be called a game could have anything worthwhile to offer. At least some of these people should know better. In some instances, of course, the misunderstanding may be deliberate, for rhetorical purposes - for example, a report (Time Out, 1981) on a civil Defence game with which I have been involved.

## 2.2 Classification of Games by Purpose

My tour of the gaming world begins with a consideration of the classification of games by purpose offered by Bowen (1978). Two criteria are used for this classification (see Figure 2.1): "control of players" and "selection of players".

As far as the selection of players is concerned, two states are possible: the game may be designed to meet the needs of a specific group of people, or suitable people may be found to play the game. In the latter case is included the research game, where, to put it simply, a game is designed to research some situation, and subjects are then found whose abilities and skills will be appropriate to the research. Also in this regime can be found the game for fun, where once again anyone can play provided they meet the fundamental requirement that they should enjoy it. On the other hand, games designed to meet the needs of a specific group of people are clearly intended to educate those people in some way.

Bowen uses the criterion of control of players to separate the



		SELECTION OF PLAYERS	
		Game designed to meet the needs of a specific group of people	Suitable people found to play the game
CONTROL OF PLAYERS	Play controlled to make the players' behaviour explicit	TEACHING	RESEARCH
	Play unrestricted subject only to game rules and format	LEARNING	FUN

Figure 2.1. Classification of games by purpose, after Bowen (1978).

categories of fun games and research games. In research games, play is controlled to a high degree so that player behaviour in specific situations is made explicit; however, in games for fun, such control is clearly superfluous. When the criterion of control is applied to the educational games, Bowen finds that he can identify two distinct types of educational game: a teaching game, in which play is highly controlled to guide players through the game in order to put across a specific lesson; and a learning game, in which such control is not exercised and the players are free to extract from the game environment whatever experience they may generate.

If one has never been involved in practical gaming it is easy to dismiss such considerations as those above as trite and pointless. In fact, though, it seems that an understanding of the purpose of a game, and, in particular, an understanding of what the game is not supposed to achieve, is a valuable component in directing the design of the game. Thus, it is important to realise that if a research game is not suitably controlled it will "degenerate" into a fun-type game (though, of course, it need not actually be enjoyable to anyone) - and there is the related point, that a fun game (for example, a commercial board war game) is not really suited, as it stands, to research.

Of particular interest to those involved in educational gaming should be the distinction between teaching and learning games. In practice the distinction may be subtle, and it is quite possible for one type of game to degenerate into the other without anyone realising. A teaching game, it will be recalled, is intended to convey specific lessons (presumably about the real world, of which it provides some type of model) whereas a learning game offers a far less explicit

experience (which nevertheless should be relevant, in some sense, to the real world). Clearly the two types of game are appropriate in different situations.

Learning games are particularly appropriate when the object of the exercise is to supply the players with general experience of working in an unfamiliar environment. This is the intention of a game such as the civil defence game HOT SEAT (R.M.C.S. reports, 1977). This game, which has been described in the O.R. literature (Hartley et al, 1979), is a large scale exercise designed as part of a training course for senior officers in local government with designated wartime roles. It is in fact a simulation of the administration of a typical county in the period following a widespread nuclear attack on the United Kingdom. Hartley et al state (p.862): "it was intended that the players should be led to draw valid conclusions for themselves rather than be taught "correct" procedures: for who knows now for certain what the correct procedures would be". Thus the game is clearly intended to be a learning game, and rightly so.

Dando (1981) has recently criticised HOT SEAT, as it is currently played, on grounds related to the fact that the game has, in his opinion, degenerated into a teaching game. However, he does not offer much practical advice on the prevention of such degeneration in the context of the design of a useful game. Replies to Dando's criticisms from representatives of the designers (Clayton, 1981; Johnson & Hartley, 1981) have been published.

I have myself been involved in the design of a similar game, HOT SPOT (R.H.C. reports, 1980; Klein & McDowell, 1981), which is a smaller scale home defence game with similar general purposes as

HOT SEAT, save that it is designed for use specifically in a metropolitan borough environment. I cannot honestly claim that we have in any way eliminated the degeneration problem, beyond possibly making the danger more explicit. In the three years intervening between HOT SEAT and HOT SPOT the issue has become clearer, and the type of gaming-consciousness promoted by Bowen (1978) has become more widespread, at least among game designers. If there is any advance in this aspect of the games it will almost certainly be due to this general progress in understanding rather than any radically new approach to game design. Sounding an optimistic note, at the time of completion of HOT SPOT our clients, the Emergency Planning Division of the Greater London Council, appeared to have a very good understanding of the issue, and seemed determined to avoid allowing the game to teach untested doctrine. Whether the same will be the case in, say, five years time, however, is impossible to predict.

If considerable effort has to be expended in ensuring that a learning game functions properly, it should not be imagined that the design of a teaching game is any easier. A teaching game is supposed to put across to its players a specific point or specific set of points: therefore the designer should be absolutely clear what exactly he is trying to teach, and should ensure that no other message is accidentally conveyed. This point is emphasised in a description of an industrial relations teaching game designed by Dando and Brown (1981): the game was intended to show shop stewards that the content of communications from management may often serve to disguise the substance of communications. The designers comment (p.627): "our aim was to produce an industrial relations teaching game which illustrated this point so well that no-one who took part was likely to forget it - even

in a critical negotiation". The designers assert that their results "overwhelmingly confirmed that the point had got home and that its real-world consequences were understood".

As Bowen points out (Bowen, 1978, p.14) the distinction between teaching and learning games is not rigid: "a game could be a deliberate cross between a teaching and a learning game". Probably most educational games fall somewhere between the two extremes. This is quite permissible provided the game designer has no illusions about the nature of the messages the game is likely to convey to players. In an introduction to business games, Lloyd (1978) seems to be making this point when he exhorts the game administrator to identify his objective or objectives for the game (p.14). He offers a number of possible objectives, ranging from "teaching of specific material" to promotion of "better understanding between departments or disciplines", which seem to correspond roughly to Bowen's teaching and learning regimes.

Clearly the hybridisation of game types can be extended, with suitable precautions, to all four categories of game. An entertaining example is provided by Laver's political games (Laver, 1979) which are published as fun games, yet have a very explicit learning element in them. "Each game is an attempt to capture one of the essential puzzles of the political process" (p.9).

More worrying, to me, at least, are the political games described by Mandel (1977). These games presented their players, who were often as familiar with the political situation under consideration as "real" decision-makers might be expected to be, with scenarios depicting foreign policy crises in which they had to act. Mandel suggests that

decision-makers in a game exhibit greater "imagination, flexibility and attentiveness" (p.615) than they would under standard operating procedures. At the same time, "certain distortions in international perceptions during crises are reflected in political gaming". Mandel uses the first point to argue that gaming is a useful exercise for decision-makers (a learning mode), but the second point, he argues, shows that these same games also have value as research games. It seems to me that if political gaming really reflects distortions, then it may be dangerous to assume that it can at the same time truly fulfill its purpose as a learning game, since it is possible that the distorted perceptions may be learnt. Political games can be research games or learning games, but in either case more care must be taken to attain the intended aims of the games than is reported by Mandel (although, of course, his descriptions of the games in his short paper may have suffered for the sake of brevity).

It requires little effort to identify a pure fun game - chess is a fine example, of course. Research games, in the Bowen sense, are more difficult to find: as he comments (Bowen, 1978, p.83), "games have all too often been called research games despite their having no stated purpose nor any coherent method of control in order that any research purpose might be met". Work with simple matrix games springs to mind as an example of gaming used as a decision-making research tool. In this context, the "cautionary tale" of Huxham et al (1981) makes entertaining reading, throwing some considerable doubt on the validity of much previous experimental work using matrix games.

Two recent research games of a more complex nature than matrix

games are the doctored fun board game of Sharp and Dando (1977), and the battle game described by Daniel (1980). Both these games will be met again later in this chapter.

### 2.3 Classification of Games by Structure

In addition to his classification of games by purpose, Bowen (1978) has also developed a classification of games by structure. Cooper (1979) has briefly described this structural classification, and I follow his description closely in this section.

The classification is summarised in Table 2.1. Two-player games are considered (though the classification is easily extended to n-player games) between a friendly Blue player (B) and an opponent Red player (R), who may be Nature. Each player may exhibit different kinds of behaviour, effect on the game, and knowledge of the game.

Behaviour may either be person-like (P) or automaton-like (A). Classification "A" describes any player who always follows a set of rigid rules - this would clearly include a computer as a player. If this condition is not strictly adhered to, then the player is classified "P".

Regarding effect on the game, a player may be able to interact with the game situation, and thereby modify it, in which case his situation is modifiable (M). Alternatively, the game may be so designed that no modification is possible (N).

Regarding knowledge of the game, a player may have access to a limited or closed information set (C), or he may have open access

Players	B Blue, the friendly player R Red, the opponent, or Nature
Behaviour	P Person-like A Automaton-like
Effect on game	M Modifiable state of the game N Non-modifiable state of the game
Information access	C Closed or limited access O Open access O' Extended open access to intentions

Table 2.1. Classification of games by structure, after Ewens (1978).



to all information about the physical state of the game (O). If, as occasionally is found, a player's open information is extended to the thoughts and intentions of his opponent, then he is said to be in an extended open category (O').

Each player's state in the game, then, is described by three parameters. In this scheme, the game of chess would be classified B(P,M,O) . R(P,M,O): both players are person-like, can modify the state of the game, and have open access to all information concerning the physical state of the game. As with the classification of games by purpose, this structural scheme may seem rather pointless - unless one has actually been involved in practical gaming. Then it can be seen as a useful framework upon which to develop a game structure appropriate to one's needs.

Much research has been carried out using the three-room game system described by Shephard (1963) and illustrated in Figure 2.2. The idea is that Blue player, Red player and Game Control each occupy a separate room: only Game Control has access to all game data. Blue and Red communicate all their moves to Game Control, from whom they receive information or misinformation about the state of the game. Thus a considerable degree of control is exercised over precisely what data is accessible to the players. In Bowen's structural classification, a three-room game typically has the structure B(P,M,C) . R(P,M,C).

This structure may be in at least one way inappropriate to a research game. The inappropriateness centres around the fact that both players are able to modify the state of the game. Consequently

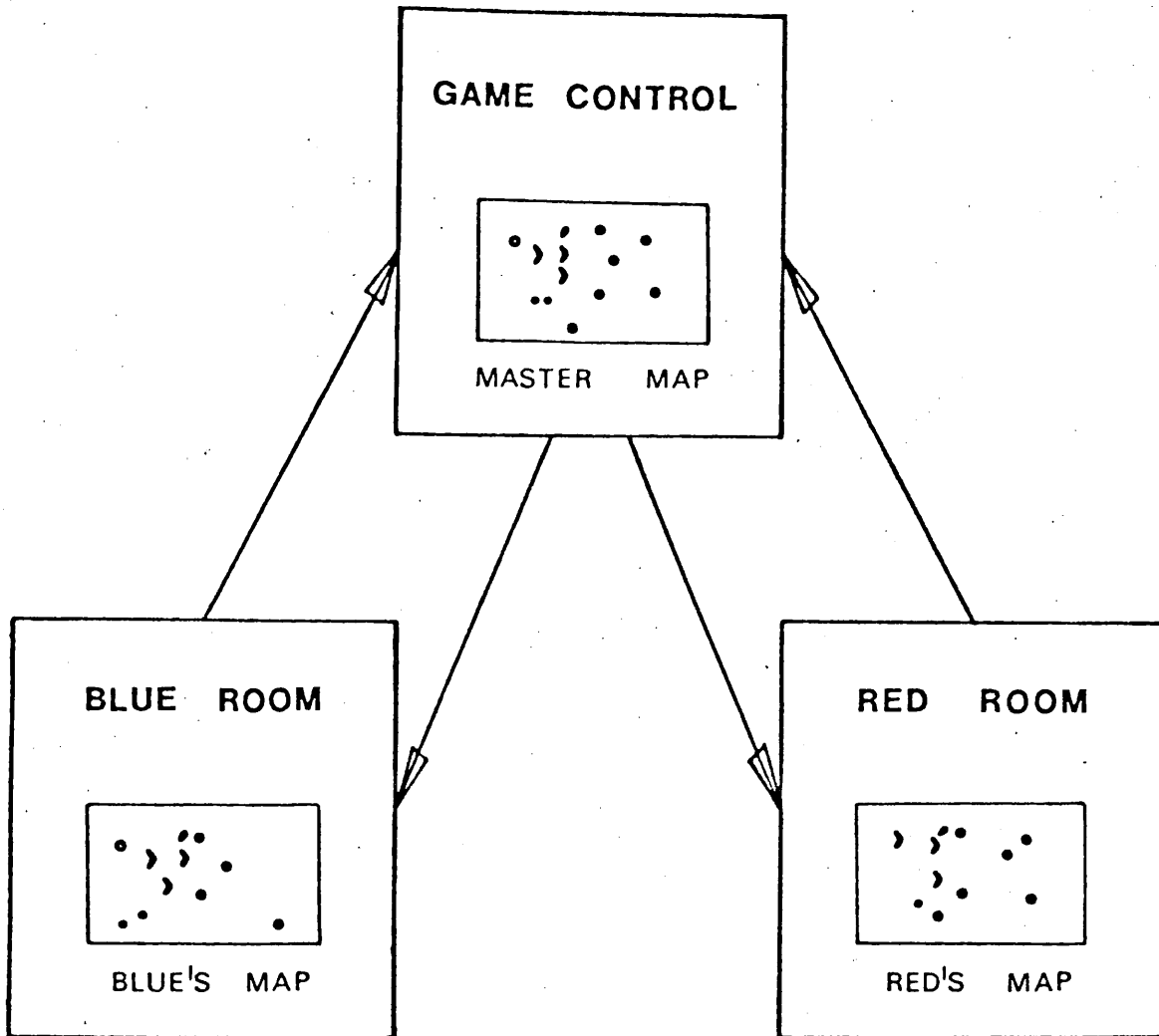


Figure 2.2. The three-room wargame, adapted from Shephard (1963). Game Control sends information and intelligence to Blue and Red, who communicate moves and intentions to Game Control. Blue and Red have no contact with each other except through Game Control.

the experimenter has comparatively little control over the state of the game: he will find it difficult to direct the game towards the specific states in which he is interested. If the experimenter's interests are sufficiently general, this is acceptable. However, if the experimenter wishes to run a replicable experiment - that is, if he wishes to subject several different players to identical game states, so that he may test behavioural hypotheses statistically - then his structure is of little value. This is unfortunate, because the richness of the interaction and detail that is possible in a three-room game provides a game environment which, other things being equal, would provide a highly desirable environment for decision-making research.

To meet the requirement of strict replicability, what clearly is needed is a game which the player or players cannot modify, so that the game state is always under the control of the experimenter alone. This non-modification condition can of course be relaxed at the end of the game, when it is no longer essential that the game state is not disturbed by a player. Thus we would have a single-decision game in which a player would spend most of the game simply observing the changing states and at the end of which he would be able to make just one decision. The structure of a simple single-decision game would be  $B(P,N,C) \cdot R(A,N,C)$ . Note that only one of the players in this structure (Blue) is an experimental subject; Red has been designated automaton-like (this is not essential, but it is practical).

Unfortunately, the single-decision game loses by its very nature the richness of interaction that is a key feature of the three-room game. For most of the duration of the single-decision

game, a player will do nothing - he will just watch. Cooper (1979, p. 53?) comments: "players are likely to feel restricted and rather helpless, particularly under conditions of crisis, and they may be encouraged to become less committed to their roles as decision-makers". He continues: "changes in observed behaviour are likely to reflect the players' increased role separation"; clearly, the players must be made to feel they are active participants in the game.

What clearly would be desirable is a game structure which combines the rich interaction property of the three-room game with the replicability property of the single-decision game. This amounts to saying that we require a game which simultaneously has both the structure  $B(P,M,C) \cdot R(P,M,C)$  and the structure  $B(P,N,C) \cdot R(A,N,C)$ . As it stands, this is clearly impossible.

If we resort to deception, however, a solution may be available. The experimenter requires a  $B(P,N,C) \cdot R(A,N,C)$  game. If we can persuade the player that he is really playing a  $B(P,M,C) \cdot R(P,M,C)$  game, though, it should be possible to design a game which meets both the requirements of interaction and of replicability. The player will be under the illusion that he can modify the game; in reality game states are strictly controlled, and he cannot.

I conclude this chapter by describing two gaming systems which are attempts to carry out this superimposition of game structures by deception. The first system, the Intelligence Man system, goes part of the way towards this aim; the second system, the Superior Commander system, appears to meet the requirements entirely, on paper, and it is the system which was used in the experiments I shall describe in subsequent chapters.

#### 2.4 The Intelligence Man System of Game Control

The Intelligence Man system for the control of games was developed by Sharp and Dando (1977) as part of their programme of decision-making research at the University of Sussex. They had chosen to adopt a commercial board game, "Diplomacy" (Avalon Hill, 1976) which they recognised as having features which might make it particularly relevant to their interests. The game involves a great deal of complex and potentially rich negotiation between players, which is generally regarded as being of far more central importance than the comparatively mundane complementary task of moving counters on the board. The game is at its richest when played by seven players - the maximum possible number. In their initial attempts to study the behaviour of players in this rich environment, Sharp and Dando ran their games using a "three-room" methodology: players were separated, and communications between them were in writing, transmitted by the game controller. Thus a complete record of the game, and in particular the negotiations between players, was produced for study.

However, Sharp and Dando wished to introduce replicability to their games. They wanted to be able to present as many players as they required with identically evolving games, so they could compare the behaviour of subjects in a strictly controlled game environment. The Intelligence Man system utilised a deception to bring this about.

All experimental players were told that they were each to be part of a two-person decision-making team, consisting of a decision-maker and an adviser (or "Intelligence Man"). The decision-maker would be an experienced player located in another room; the experimental

players would act as advisers to their decision-makers, although they would take over as active decision-makers later. An adviser would be party to all information available to his decision-maker (communications sent and received, and the situation on the game board). His only role in the game, however, would be to send advice and information to his decision-maker. There would be no guarantee that any advice would be followed, though of course it would be considered by the decision-maker in formulating his next moves.

In reality, all experimental players were advising an imaginary player, and all were receiving precisely the same data about the state of the game. The game was entirely predetermined, except at the end when the players took over as decision-makers and made their own final move (believing that the game would continue for some moves yet). The differences between the players' view of the game and the real situation are illustrated in Figures 2.3 and 2.4.

Thus the game structure is in reality of the single-decision type,  $B(P,N,C) \cdot R(A,N,C)$ , though the players do not see it as such. Cooper (1979) has found it convenient to extend the Bowen classification in order to describe the situation the players see. A player may not only either be able to modify (M) or not modify (N) the game: there is also an intermediate state where a player has no direct effect on the game situation, but may be able to influence it (I), by advice, for example. The structure of an Intelligence Man game as seen by a player is then  $B(P,I,C) \cdot R(P,M,C)$ .

Now, the ideal situation would be where a player believes himself to be in a three-room structure  $B(P,M,C) \cdot R(P,M,C)$ . It can be seen that the Intelligence Man system gets close to this: the only

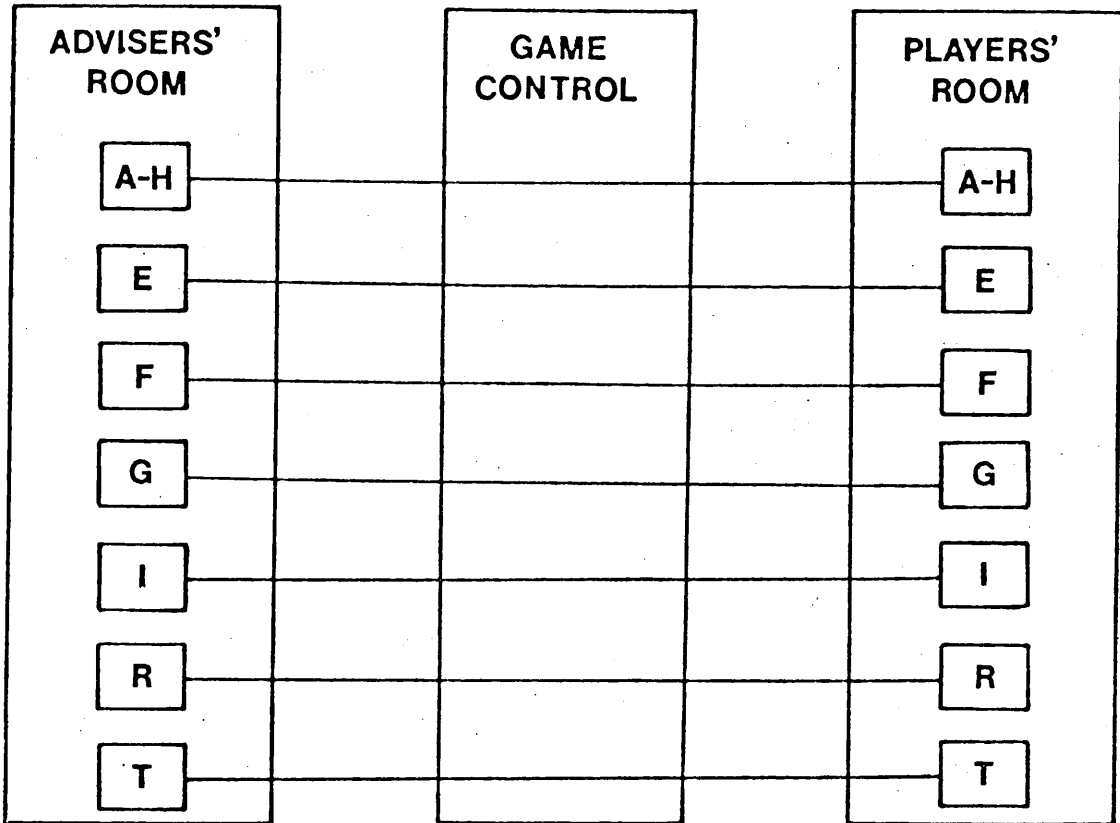


Figure 2.3. The situation which the players in the Intelligence Man "Diplomacy" games perceived. Players represent different countries: A-H = Austro-Hungary; E = England; F = France; G = Germany; I = Italy; R = Russia; T = Turkey. (After Sharp & Dando, 1977.)

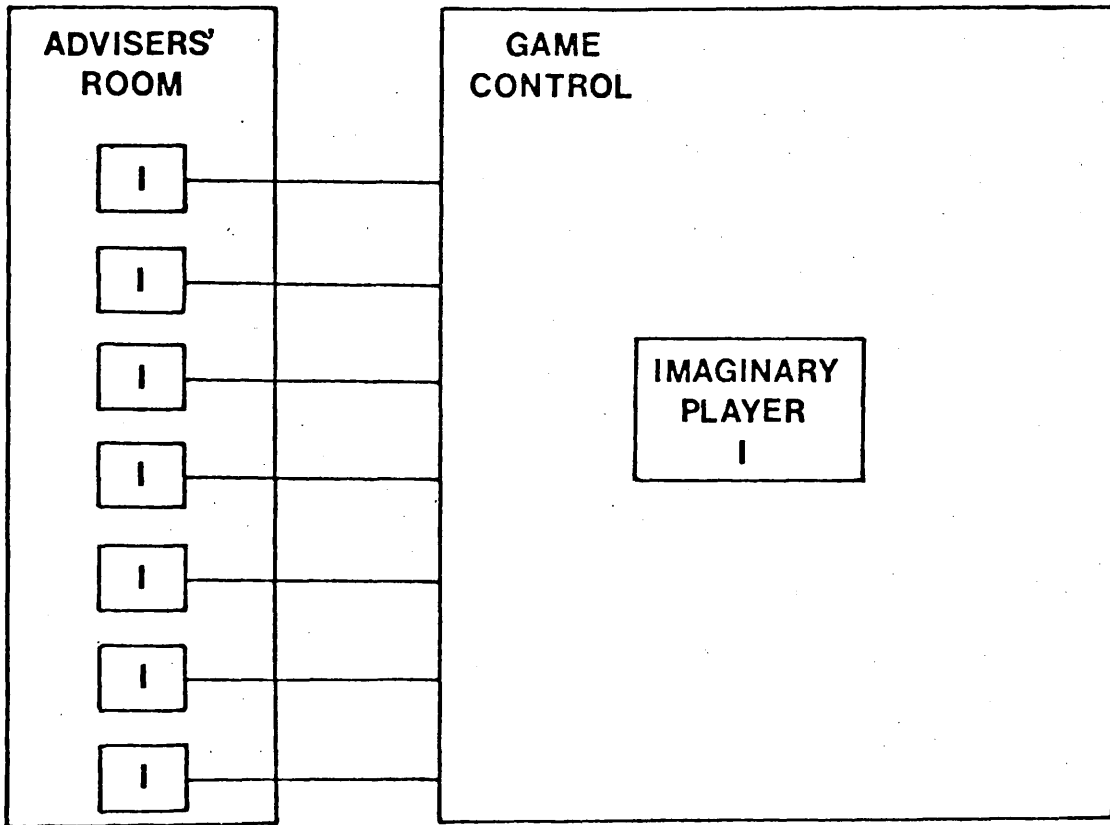


Figure 2.4. The real Intelligence Man "Diplomacy" game. All players were advisers to Italy, an imaginary player directed by Game Control. (After Sharp & Dando, 1977.)



difference is the influence/modify disparity in the players' perceived views.

How critical is this disparity? Sharp and Dando (1977, p.287) report that they were able "to produce a testing game for theories about perception, which was complex (and therefore rich in behavioural alternatives) but controlled". They felt that they had demonstrated "that complex games in general can be played in a controlled manner to achieve useful results", and emphasise that their system's virtue is in "providing interactive responses in a non-interactive situation".

Despite this, there is still a difference between being a decision-maker and a mere adviser, and the separation of game-playing and experimental subject roles that was noted in the context of single-decision games may still occur, albeit less extremely. This may be why Daniel (1980), in comparing the performance of players in a research battle game which was first run using a type of single-decision structure (in which the player decided when his decision should be made) and then using the Intelligence Man system, feels (p.417) that "it is possible ... that changing from the original game to the "Intelligence Man" game has reduced the players' performances." Cooper (1979, p.533) comments: "built into the Intelligence Man system itself is just that separation of roles which stress and crisis will exacerbate". To avoid this, we need a gaming system "in which the player's main role is perceived as Decision-Maker rather than Intelligence Man".

## 2.5 The Superior Commander System of Game Control

The Superior Commander system (Cooper, 1978, 1979) is a gaming

system which permits a single-decision game to be played in which the experimental player believes he is a real decision-maker, able to modify the game, throughout the game. In this chapter, only a brief description of the system will be given; the system in practice will be described more fully in later chapters.

The Superior Commander system works by embedding the experimental player in a hierarchy of decision-makers, as shown in Figure 2.5. The experimental player is told that he will be supplied with data concerning the state of the game from two (or more) subordinate decision-makers. On the basis of this data he can make his decisions, which he will communicate to his subordinates, who will implement them. The subordinate decision-makers are separated from the experimental player; all communications are carried from player to player by the game controllers. The subordinate decision-makers will necessarily each be playing in a subsection of their superior's game environment.

In reality, the subordinate decision-makers do not exist: the game is entirely predetermined. Often, therefore, the subordinate commanders will appear to deviate from the commands given by the experimental player. This can be explained to the experimental player quite easily: he is told that the details of the situation in the subordinate's area of interest made his superior's instructions inappropriate or even impossible to carry out.

One problem with such a system might be that the experimental player's necessary distance from the low level, where the implementation of play is going on, may cause him to lose interest in the game. To deal with this problem, the player must be assigned a

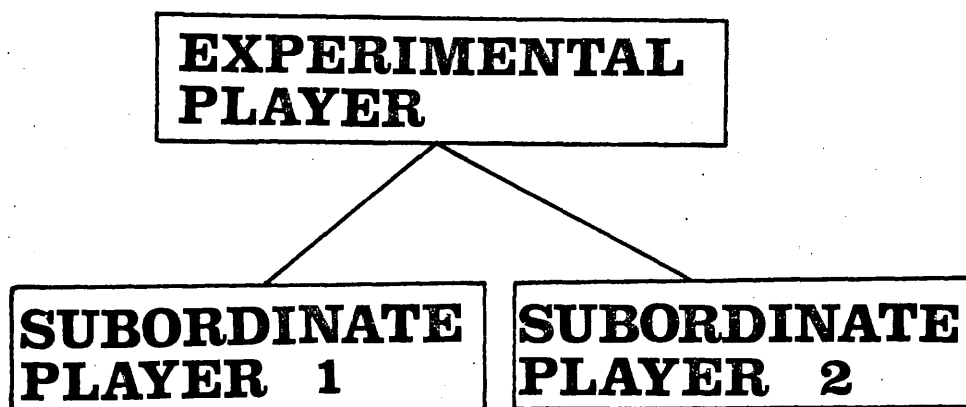


Figure 2.5. The basic hierarchy of decision-makers required for the Superior Commander system.

dummy task. The dummy task may be any task that appears to permit the player direct interaction with the game environment (as opposed to interaction via subordinate decision-makers). The task should be distinct from the main decision-making duty, but should appear to the player to be important to the outcome of the game. Needless to say, in reality the dummy task will have no effect on the game at all.

I have described the absolute essentials of the Superior Commander system. In fact, it seems appropriate to expand the command hierarchy in various ways. This expansion is illustrated in Figure 2.6, which shows the player's view of the game, and Figure 2.7, which shows the reality of the situation.

Probably the most important expansion is the addition of a third level to the decision-making hierarchy superior to the experimental player. The experimental player believes this superior to be a real player, though in fact he is not. The superior's purpose is to direct the experimental player, and assist in defining for him what his objectives in the game should be. This goes a long way towards ensuring that all experimental players have some common view of the aim that they and their decision-making team is attempting to achieve. In many games, this may not be at all obvious.

Fictitious decision-makers at the lower level (and, indeed, at the level of the experimental player) may be introduced as the game situation demands. There is also no real reason why all data on the game state should be supplied (apparently) by low level decision-makers. Additional data may quite legitimately be supplied by the game controllers direct to the experimental player.

As can be seen, a game played using the Superior Commander

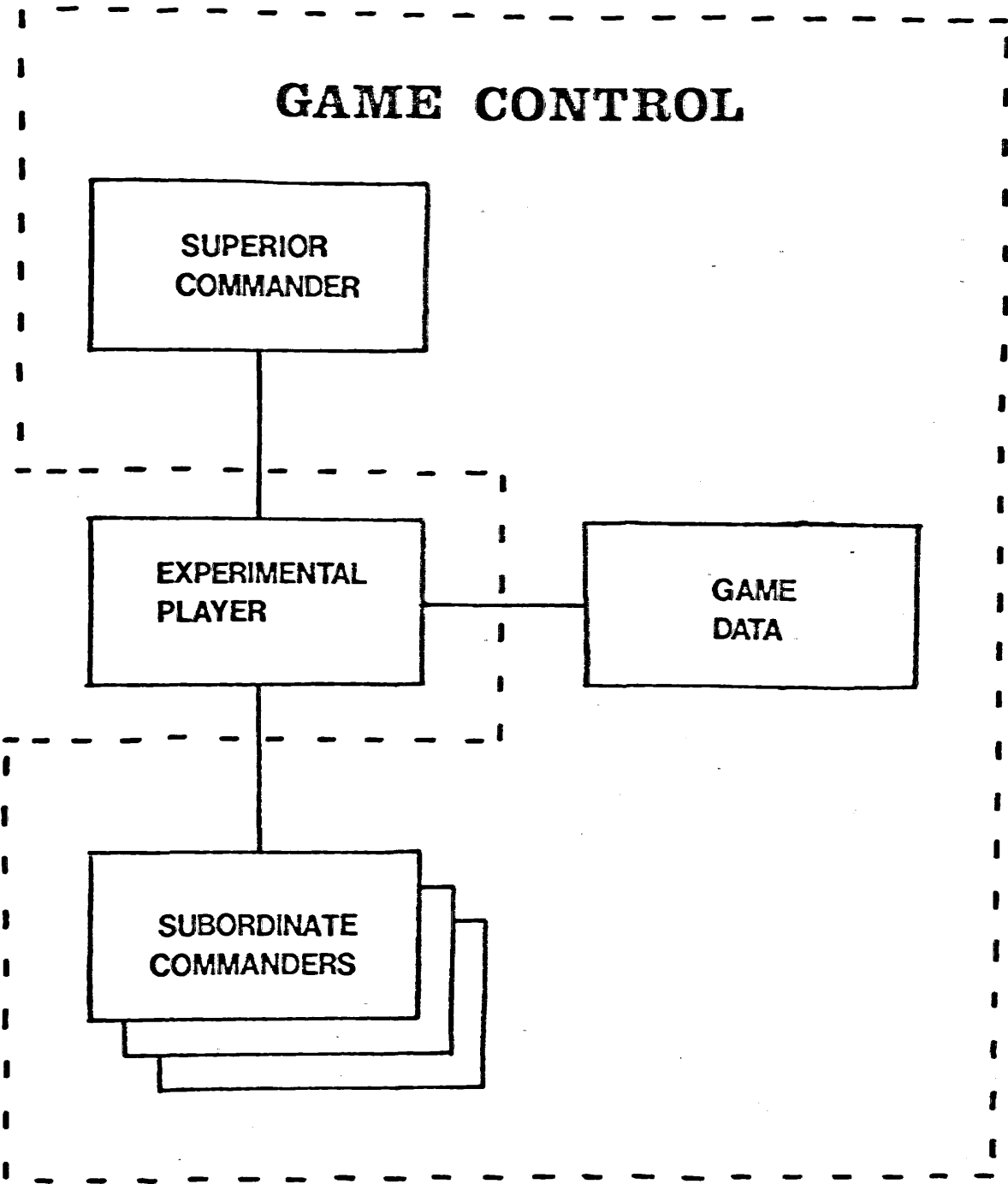


Figure 2.6. The Superior Commander system, in full, as seen by the experimental player.

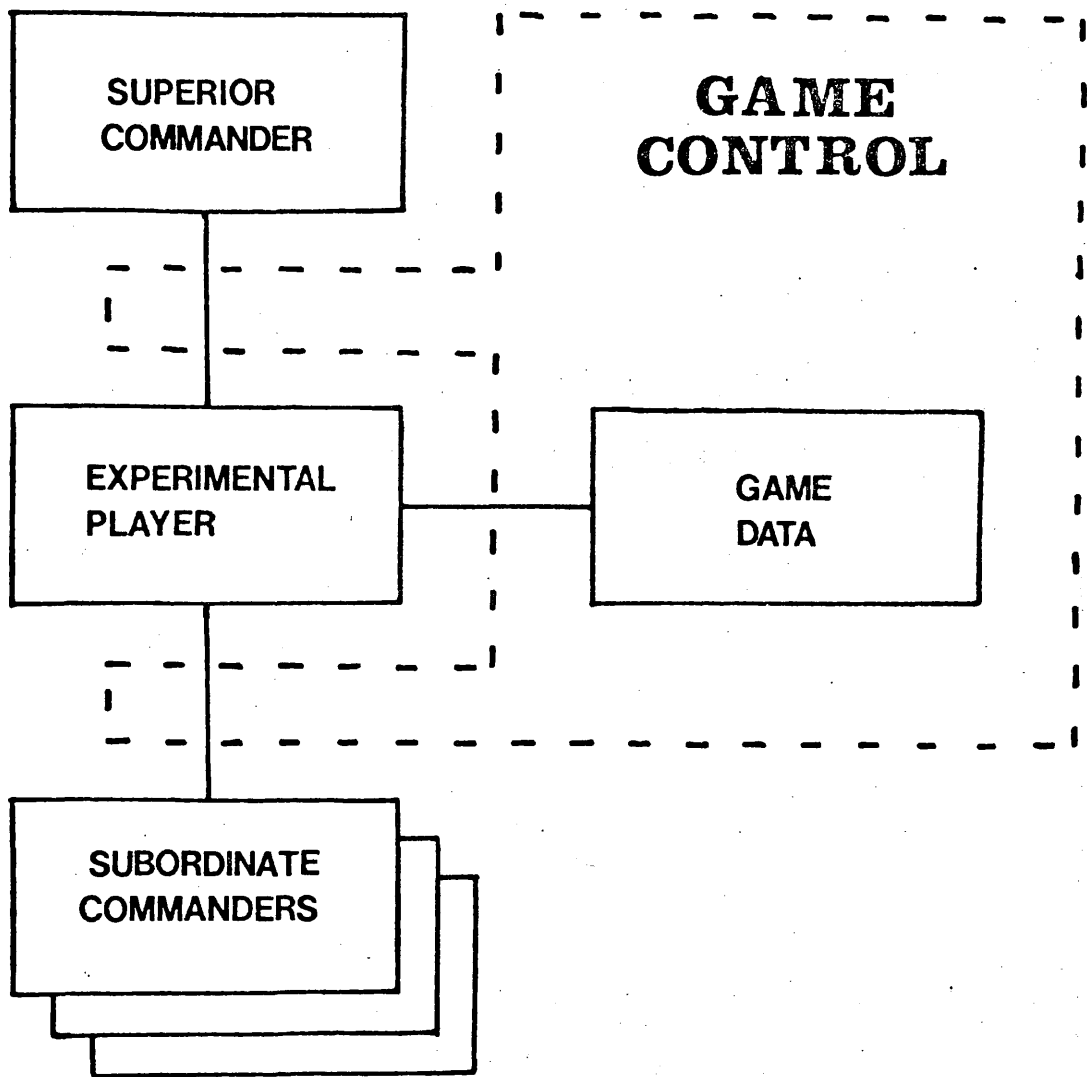


Figure 2.7. The Superior Commander system, in full, as seen by the experimenter.

system has the single-decision structure  $B(P,N,C) \cdot R(A,N,C)$ . As far as the player is concerned, when playing against some kind of enemy (which may be Nature) he is very definitely a decision-maker who is able to modify the game, and therefore its apparent structure is the three-room  $B(P,M,C) \cdot R(P,M,C)$ . Thus the Superior Commander system satisfies our requirement that a single-decision game can be made to seem like a three-room game.

Cooper (1979, p.536) notes three minor difficulties associated with the system. First, deception is required, with attendant special demands concerning game secrecy. Second, the system imposes some restrictions on scenario design. Finally, the system does not entirely eliminate the effect of a player having two roles in a game, as experimental subject and as decision-maker, although the effects should be reduced considerably. However, "any gaming approach must have this problem".

Cooper concludes (p.536): "the Superior Commander methodology is ... a general methodology for research games that can be used for a range of purposes for which realism and good experimental control are both required". In the following chapters I shall describe the development and use of the Organisational Control Game, a research game developed around the Superior Commander system, which convincingly demonstrates the effectiveness of the system.

## CHAPTER 3

### BOARD WAR GAMES FOR DECISION-MAKING RESEARCH

#### 3.1 Introduction

The previous chapter was concerned mainly with developing ideas about the structure of a research game. This ultimately led to a description of Cooper's Superior Commander System (Cooper, 1978, 1979). The description was quite deliberately couched in general terms, to emphasise the point that the structure of the system is largely independent of the specific subject matter of the game. The Superior Commander System would be applicable to any situation in which some kind of hierarchical command or administrative structure could be legitimately used - for example, war gaming, as will be described in this chapter, or business gaming.

By contrast, this chapter will be concerned mainly with developing ideas about the content of a research game. Many of the points to be made here have already been aired in the literature (Cooper & Klein, 1980). I shall develop the ideas in the context of a description of the early work of game development which we carried out at Royal Holloway College. It should be realised, therefore, that although the problems that are identified in this chapter are, I feel, quite general in the field of research gaming, the specific solutions that I offer are by no means considered to be unique. In other words, the research game which we developed is not the only possible game which meets the requirements which were specified. Other games could be developed, and may well be.

The initial aim of our work was to develop a suitable methodology



to test hypotheses about decision-making behaviour. More generally, we wished experimentally to test models of the decision-maker. It is not really possible to study, in a controlled fashion, decision-makers in their natural decision-making environments. In such environments, of course, important variables are not subject to control in the strict way we might like for hypothesis-testing. We wished to study decision-making in controlled environments, where variables could be manipulated to enable specific aspects to be examined; a gaming approach seemed suitable for this purpose (Bowen, 1978). Our initial experimental work was directed towards the development of a game in which a subject would be required to make decisions under experimentally controlled conditions, and in which it would be possible to analyse his decision-making in detail.

War gaming has been used for the analysis of military situations and the training of commanders since the early nineteenth century (Wilson, 1970). Although many of these games were played for recreation by military officers, it is only in the last twenty-five years that war games have become available to the public (Strategy and Tactics, 1977). A wide variety of commercial board war games are now produced (Palmer, 1977).

Our interest centred on war gaming for mainly practical reasons. One of us (Dr. Dale F. Cooper) had some military experience which would be valuable in designing game scenarios. In addition, we had a source of military decision-makers + military officers- to be supplied by our sponsors to play the game. It was therefore natural to prepare for them a military decision-making environment. A board war game is an economic and relatively simple depiction of a military situation.

In the next section I describe our initial experiments with a commercially available board war game, and with the Intelligence Man System for the control of research games. Both the game and the gaming system proved to have severe limitations for our purposes. In the following section I discuss in detail the properties which a board war game should possess if it is to be used for research. Next I consider the game which we developed, and its control by the Superior Commander System. In the final section I comment briefly on war game design.

### 3.2 Experiments with a Tactical War game

The first stage of our experimental programme was to assess whether a commercial board war game could be adapted to our research purposes. The game we chose to study was "Panzer Leader", a war game manufactured by Avalon Hill (1974), depicting tactical armoured combat on the Western Front in the Second World War. The game is appreciably complex. It features many different types of fighting unit, each with its own characteristics, and a wide variety of terrain which influences combat in many ways. This complexity was one of the most attractive features of the game, and that which suggested to us that it might have potential for research. It was envisaged that scenarios could be designed which involved decisions concerning a wide range of possible courses of action, offering a wide range of options to players. Palmer (1977) regards "Panzer Leader", and its companion game "Panzer Blitz" (Avalon Hill, 1970) as being exciting and requiring a high level of skill. It was hoped that these features would encourage experimental subjects to become involved with the scenarios and induce them to play as well as they possibly could.

It was recognised that in preparing an experimental war game of this kind the "Panzer Leader" rules might have to be simplified considerably, since few subjects are likely to have had much previous experience in the playing of commercial war games of this type. Even if playing only an advisory role, a subject would require a reasonable degree of competence before he could begin to make meaningful judgments of game situations.

Our initial choice of a gaming methodology was a variant of the Intelligence Man system developed by Sharp and Dando (1977), which was discussed in the previous chapter. In this system, it will be recalled, subjects, believing themselves to be acting as advisers to real players, are led through identical series of predetermined moves. Subjects supply the experimenter with information about their views of the game in the form of messages to the players whom they believe they are advising, and, possibly, by a real move which they may be asked to make to conclude the game. Sharp and Dando used their Intelligence Man system successfully with an adaption of the commercial board game "Diplomacy" (Avalon Hill, 1976).

#### Recording a Tactical Game

Our preliminary investigations were concerned with the feasibility of recording "Panzer Leader" games. It seemed to us that a desirable feature of an experimental game would be that it should be possible to record the entire course of the game accurately, unambiguously, quickly and with a minimum of disturbance to the players.

"Panzer Leader", in common with most commercial war games, is played on a hexagonal map-board, each hexagon being individually coded.

Counters representing units are approximately  $\frac{1}{2}$ " X  $\frac{1}{2}$ " and are identified by a series of numbers representing combat strengths and movement capabilities (sufficient to identify the type of unit) as well as a number identifying each unit individually. In even simple scenarios as many as twenty or more units may be moved or engage in combat at each round of the game.

A straightforward game using what we judged to be one of the simplest scenarios supplied with the game (with a small number of units and few types of unit) was played between players who had had previous experience with "Panzer Leader" and other war games. In fact, the players were members of the Royal Holloway College Games Club. The experimenters attempted to record the whole course of play, while it was going on, on specially prepared forms.

The results of this first trial indicated convincingly that recording "Panzer Leader" and similar games is not a feasible proposition. It is not possible to get close enough to the board while a move is taking place to be able to record the move without interfering considerably with the players. This means that either play must be held up frequently, or there must be a long pause at the end of each round in which the recorder notes in detail the outcome of the round. Neither alternative is satisfactory. Additionally, there are several operations in "Panzer Leader" and other games which require the player to explain to the recorder exactly what he is doing; virtually all combat requires such a commentary.

Even when the time spent on the recording process exceeded tolerable limits, and, interference to the players was considerable, recording errors still occurred. It does not seem possible to

record "Panzer Leader" reliably without severely disrupting the flow of the game.

#### Intelligence Man System Applied to Panzer Leader

Our next experiments were to assess whether the Intelligence Man system could be applied to games of the style of "Panzer Leader". As was described in the previous chapter, Intelligence Man has worked admirably for "Diplomacy" (Sharp & Dando, 1977), but the kind of war games we hoped to use differ from "Diplomacy" in many ways and the applicability of Intelligence Man could not be taken for granted. The output of an Intelligence Man experiment is to a large extent obtained in the messages the adviser sends to the person he believes is playing the game; therefore, it is important to find out exactly what the subject, in his role as adviser, makes of his task. He may see himself as being in a similar role to a real player, and his messages may provide detailed suggestions concerning the microstructure of play; alternatively, he may take a wider view, and, freed of the constraints of organising the details of movement and combat, may point out strategies and options that might otherwise have gone unnoticed. It is in this latter mode that we felt that the output of the game would be most interesting, at least for our research purposes; in any case, it is doubtful whether an adviser with little or no practical experience of the game could supply much worthwhile advice in the former mode.

Four subjects were required for an Intelligence Man evaluation applied to "Panzer Leader". Two subjects played a normal game, and each was assigned one of the other subjects as an adviser. The advisers were separated from the players (all communications being

in writing, and carried by the experimenters) and they were presented with an identical game. There were, needless to say, several hitches in transmitting the details of play to the advisers' game, because of recording problems already discussed. The subjects were, once again, members of the Royal Holloway College Games Club.

It is important to note that Intelligence Man, as used here, differs from conventional Intelligence Man experiments in that there was a real game going on - no deception was involved. Our purpose was to study the nature of the output of the advisers and their general attitude to the game. Their reactions to the specific scenario were not being tested.

Despite repeated discouragement from the experimenters, the advisers tended to become extremely involved with the detailed mechanics of the game. Much advice consisted of details of movement and deployment of individual units; the game was viewed from a tactical rather than strategic viewpoint and there was little sign that the advisers had studied and assessed alternative courses of action. They were, in effect, thinking as frustrated players rather than as advisers. This was not the result for which we had hoped.

It seems likely that it is an intrinsic quality of "Panzer Leader" and similar tactical games to encourage this mode of thinking. "Panzer Leader" was designed as a game of tactical level combat, and to expect subjects to seek comparatively abstract strategies when it is clear to them that the outcome of the game will be decided at a tactical level is perhaps unfair. Everything about the format of "Panzer Leader", the wealth of terrain detail, the diversity of units, and the range of combat strengths, demands attention at a tactical

level. At this stage, we decided that this type of game had little to offer us in our experimental work.

It was becoming clear that we were ourselves going to have to design our own, strategic level, research game. The system of game control that we would use would undoubtedly be of central importance. The Intelligence Man system appears to work reasonably well - in the experiment I have described we satisfied ourselves that the role of adviser is one that an experimental subject can reasonably be expected to adopt. However, as I shall argue later in this chapter, this may not be enough - we have certain reservations concerning the validity of results obtained with the Intelligence Man system, and this led us to incorporate the Superior Commander system (Cooper, 1978, 1979) in our game.

In order to describe the considerations which we felt it necessary to bear in mind when designing our game, I shall, in the next section, set out the requirements which we felt an experimental game and gaming methodology would need to have in order to be of value to us for research. In the section after this, I shall describe the way in which we attempted to meet these requirements.

### 3.3 Requirements for a Research Game

I identify six major requirements for an experimental game, and three major requirements for a gaming methodology. The latter have already been, in part, considered in the previous chapter; here, they are considered in detail, in the context of the experimental work described in the previous section.

## Requirements for a Game

G1: Simple, easily learnt, rules. Detailed analysis of "Panzer Leader" games shows that even experienced players sometimes unintentionally break the game rules without their opponent or any observers noticing. Inexperienced players have virtually no chance of playing a faultless game. Even if subjects are only to be advisers, they must nevertheless understand fully the potential of every unit on their game board, and the effect of every piece of terrain. Bearing in mind that we may wish to use subjects who have little or no war game experience, and that we would hope an experiment would last no longer than one working day, including a minimum of training time, it is clear that the game rules must be much simpler than those of "Panzer Leader". They must be simpler, indeed, than those of all but the most elementary commercial war games.

Not only must the rules be simple in practice, but they must be seen to be simple: that is, they must be easy to learn. This requires that they are concise and short, do not require any specialised knowledge, and, of course, are well-written.

G2: Speed of Play. An experimental war game is rarely shorter than five rounds long, and it is essential that the play moves as quickly as possible, otherwise the subject will become fatigued, bored, and lose interest in the game towards the end. (Usually it is towards the end of a game that we would hope to extract the most interesting information.)

We feel that an experimental war game should not last longer than



five or six hours. For our purposes we regard sixty minutes as the maximum acceptable time for a round, although it would be preferable if this could be cut by as much as half. The length of a war game round is usually roughly proportional to the complexity of the game, and it is felt that if, as advocated in G1, a simple game is developed, this problem will cease to be as serious as it was with the "Panzer Leader" games, in which the players' spirits definitely lagged towards the end.

G3: Ease of recording. It should be possible, if necessary, to take down an accurate record of an experimental game. The format of the game, therefore, should allow easy, swift recording, with the minimum of interference to the players. "Panzer Leader" was not suitable in this respect.

G4: Strategic aspect. From the point of view of our research, a strategic orientation is the most important property of the game, because strategies and options form the material which we would like the subject to consider. The fact that "Panzer Leader" is in practice almost entirely tactical has already been discussed. We would like a game which is entirely strategic, presenting no tactical aspects to the subject at all. He should adopt a strategic frame of mind naturally, prompted only by the format of the game; it should not appear to be a rather unconventional and contrived way of looking at the situation.

Scenarios must be rich in strategic options, and there must be potential decisions for the subject to make. To some extent preparation of these rich scenarios could clash with requirement G1, that the game be simple. Careful design of scenarios is necessary to reconcile these conflicting requirements.

G5: Interest. To be successful, an experimental game needs to be interesting enough to hold the attention of the subject for the duration of play. Thus, in designing a simple game, we must not allow the game to become too abstract and unrealistic. The excitement and interest of "Panzer Leader" is attributed in no small measure to its complexity and its realism (Palmer, 1977, Chapter 7). A strategically-rich game, as advocated in G4, seems likely to hold the interest of subjects.

G6: Natural Extraction of Experimental Data. As much as possible of the data that we wish to extract from the experiment should be provided by the subject in the natural course of playing the game. Intelligence Man achieves this elegantly: the data is supplied by the advisers' communications to their (non-existent) players, and also by their move at the end of the game. This has the advantage of not encouraging the subjects' tendency, often prominent when experimenters try to extract data from them directly, to supply the sort of data they feel the experimenters expect (see, for example, Silverman, 1977). We would hope for a similarly "natural" data extraction process in any other gaming methodology.

#### Requirements for a Methodology

M1: Replicability. The fundamental property of an experimental methodology that we require is replicability: that is, the capability to put many subjects into a scenario which develops in a predetermined way each time it is played. Only in this way can we hope to derive general results about decision-making behaviour, rather than a commentary on the behaviour of individual subjects in unique, non-replicable situations.

In our case the problem is complicated by the constraint that subjects must not realise that scenarios are predetermined. They must believe that they are modifying, or at least influencing, the course of events, when in fact they are not. The Intelligence Man system has managed this deception admirably (Sharp and Dando, 1977), and as far as I know it was the first methodology developed to allow the replication of complex games. Any other methodology must perform as well in this respect as Intelligence Man.

M2: Complexity. A game must contain sufficient complexity and detail to be accepted by the player as realistic. If a player is expected to play in a sensible fashion, he must be convinced that the game environment is a reasonable representation of reality, and that his role in the game is realistic. In particular, the game must offer the player scope for complex interaction with a detailed game environment. Such scope is available in three-room war games (Shephard, 1963), but is lacking in single-decision games, of which Cooper (1979, p.532) has commented: "it is not obvious that the players would regard them as realistic".

M3: Role of the Subject. I come now to what I feel is an important shortcoming of the Intelligence Man system, especially in situations in which subjects are stressed: situations in which we would be interested. A subject in a gaming experiment can be regarded as having two roles: as a Participant in the game; and as a Subject in the experiment. The experimenter is interested in the subject's Participant role. However the subject is constantly aware, to some extent, of his Subject role. Effects due to this awareness are not wanted by the experimenter, and should be minimised.

The ideal situation, from the point of view of the experimenter, is one in which the subject is entirely involved with the game - he is, effectively, entirely Participant. Under stress, however, the Participant role may become, in some sense, unpleasant. An escape from this role is therefore sought, and is conveniently provided in the form of the Subject role. In stressful circumstances the subject is more likely to become particularly aware of his Subject role, and his Participant role will weaken: he will become less involved in the game. If we are interested in the effects of stress on his Participant role behaviour, this shift in awareness of roles should clearly be avoided as much as possible.

It is felt that the Intelligence Man system, in which the subject can only hope to influence the game indirectly as an adviser, is likely to exacerbate this problem. An adviser is intrinsically less involved in a game than a player; other things being equal, he will be less committed to the game. In particular, under stress he is likely to lose interest and involvement to a greater degree than he would if he felt he played a more important, central part (see Cooper, 1979). Such effects may account, in part, for the results of experiments by Daniel (1980) using the Intelligence Man system.

We require a gaming methodology that gives a subject a central part: one which, in the subject's view, is not merely useful but essential to the play of the game. At the same time, we must not infringe the requirement of replicability, which gives a methodology its strength as a scientific means of examining behaviour.

### 3.4 A Board War Gaming System for Research

In this section I describe the Organisational Control Game, which, in conjunction with the Superior Commander system, has been developed to meet the requirements outlined in the previous section. I believe that these requirements have been satisfied and that we have a gaming system which has the potential for a wide range of applications. Initial trials of the system and the results and conclusions from these trials will be discussed.

#### The Organisational Control Game

The Organisational Control Game is a hierarchical war game with three different levels of play, involving many players. For our purposes we need only consider the structure of the Friendly side.

The highest level of play involves a Corps under the command of a single Corps Commander. The Corps is split into three Divisions, each under the command of a Divisional Commander, and each Division is further split into a number of Task Forces, each under the command of a Task Force Commander. During play, all the various Commanders are separated from each other, and can communicate only in writing.

Each of the Task Force Commanders engage in combat with the enemy in a tactical war game not unlike "Panzer Leader", played on a detailed mapboard representing only the area of terrain within the Task Force's sphere of influence. The only major differences, from a Task Force Commander's point of view, are that he has to report periodically on the state of the battle to his Divisional Commander and that the Divisional Commander will send him advice, information and directives at regular intervals.

The Divisional Commander receives reports from the Task Force Commanders which furnish most of his knowledge of the state of the lower level, tactical battle. He plays on a mapboard covering the Divisional area, and on it the Task Forces are represented, each subdivided into a few battalions, although their detailed structure is not visible at this level. The Divisional Commander's view of the game is far less detailed than those of his Task Force Commanders, but covers a wider area. Additional information is supplied to the Divisional Commander by intelligence reports from Local Reconnaissance, and by information sent from the Corps Commander, to whom he is required to report from time to time.

The Divisional Commander's role is to examine the information that comes in to him and to use it to develop strategies which he can attempt to implement by sending appropriate directives to his Task Force Commanders. He is playing a strategic game. In addition to this main task, he is also involved with an additional task: the maintenance of supply lines to his Task Forces.

At the top of the chain of command the Corps Commander receives reports from his Divisional Commanders, as well as his own intelligence units. On the basis of these he forms his overall view of the game and sends information, advice and directives to his Divisional Commanders accordingly.

In general, experimental subjects will play Divisional Commanders, and I now wish to show that, as far as the Divisional Commanders are concerned, the Organisational Control Game meets the six requirements for an experimental game set out in the previous section.

G1: Simple, Easily Learnt Rules. The design of the game is such that all combat with the enemy is handled by the Task Force Commanders. The Divisional Commander, in order to make maximum use of the representation of the battle on his gameboard, need only be supplied with general combat rules. By removing the responsibility for, or even the possibility of commenting upon, the detailed tactical aspects of the game, many of the problems associated with learning the rules necessary in a complex war game are eliminated. The Divisional Commander's rules are, in comparison to those of a Task Force Commander, extremely simple indeed. For example Divisional Commanders are given the following combat guidelines (from the "Rules for Divisional Commanders" as reproduced by McDowell et al, 1979b, p.A-4):

Combat is resolved in detail at the Task Force level, and Divisional Commanders require no detailed knowledge of the combat rules.

A friendly Task Force consists of four battalions and an enemy regiment consists of three battalions. The battalions are likely to be of roughly equal size. In an engagement between two battalions, that which is defending will have an advantage over the other. However, various other factors will alter the strengths of units.

The defensive strength of a unit is increased if the unit did not move in the previous turn; this represents the ability of a unit to 'dig in'. Both the offensive and defensive strengths of units are severely reduced if they are not kept in supply. The strengths of units are increased if they occupy hill hexes or town hexes.

G2: Speed of Play. Because individual units are represented as aggregates on the Divisional Commander's map, there is very much less work to do in assessing them and organising them in each round. This reduces considerably the time taken for each round.

G3: Ease of Recording. The relatively few units counters on the map at Divisional Commander level make recording a far easier task than with a more complex game. In fact, the Divisional Commander cannot move units directly, only indirectly through his orders to the Task Force Commanders, and detailed combat evaluation only occurs at the Task Force level. Very little work is therefore required to prepare a record of a game at the Divisional Commander level.

G4: Strategic Aspect. From the strategy and tactics point of view, the Organisational Control Game at Divisional Commander level is very different from "Panzer Leader". Any details concerning the tactical battle are entirely eliminated from the Divisional Commander's map; his view is in all respects strategic. Thus he will generally give strategic rather than tactical advice to his Task Force Commanders, and he is concerned almost entirely with strategies and possible courses of action, both of his own forces and of the enemy.

G5: Interest. It is felt that the Divisional Commander's task in the Organisational Control Game, while conceptually simple and unencumbered by detailed rules, is potentially interesting for a player, even one who has had no previous experience with war games. Interest is of course closely related to the specific game scenarios to be played. The scenarios we developed will be described in the next chapter. For the time being, I note that the Organisational Control Game allows a fair amount of flexibility in scenario design -



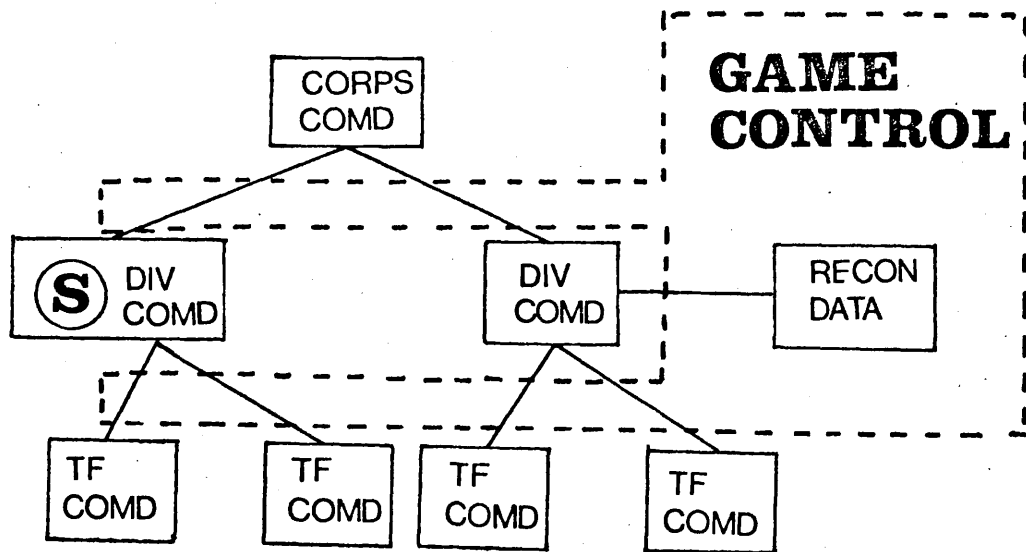
thus we were able to develop interesting and exciting game scenarios depicting entirely different situations.

G6: Natural Extraction of Experimental Data. Data is extracted from the subject in the natural course of the game in two ways. Firstly, at regular intervals in the game the Divisional Commander sends information, advice and directives to his Task Force Commanders; analysis of these communications provides some information about his view of the game. Secondly, from time to time the Corps Commander requests a report from the Divisional Commander; the Divisional Commander outlines his thoughts concerning topics in which the experimenters are interested, and he gives his assessment of the strategic situation. These communications, too, may be analysed.

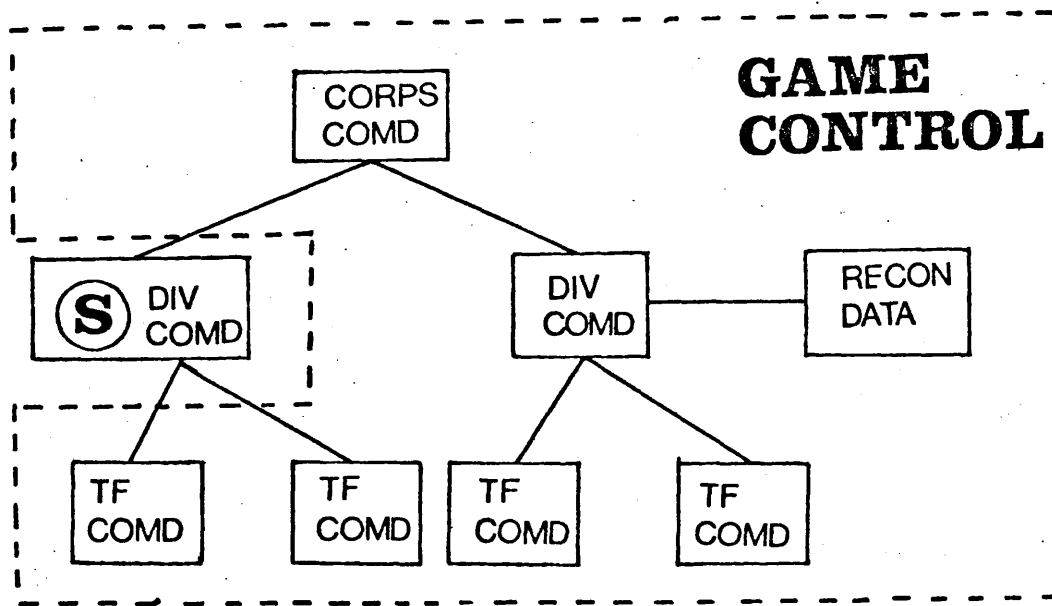
#### Use of the Superior Commander System

The Organisational Control Game, with the experimental subject in the role of one of the Divisional Commanders, meets our requirements for a research game. I shall now show how the Superior Commander system provides a technique for using the Organisational Control Game while satisfying the three requirements for a research gaming methodology.

A Divisional Commander player believes himself to be playing in a hierarchical system composed of several other players (Figure 3.1a). In fact he is not: all other players are played by Game Control, and the course of the game is entirely predetermined (Figure 3.1b). All communications from the Task Force Commanders, whose games constitute the action of the battle, are written beforehand, and the entire development of the scenario is controlled by the experimenters.



**a**



**b**

Figure 3.1. The Superior Commander system applied to the hierarchy of the Organisational Control Game. (a) From the viewpoint of the experimental subject ( **S** ), who believes Game Control to be responsible only for communication links. (b) from the viewpoint of the experimenter, who knows all players but the subject to be played and controlled by Game Control. The number of Divisional Commanders and Task Force Commanders need not be precisely as shown in the above diagrams.

Thus the condition of replicability (M1), is met: the same game may be played with different subjects as many times as the experimenters wish. The obvious problem is: if the Divisional Commander is sending directives to his Task Force Commanders, will he not become suspicious when he finds that there is no apparent reaction to these directives? With careful scenario design this is not a major problem. Scenarios can be designed so that the directives the Divisional Commanders give, provided that they are strategically sound, can reasonably produce the predetermined outcome. Discrepancies, even quite major ones, can be accounted for by explaining to the Divisional Commander that, given the state of the low-level tactical battle not apparent on his map, his directives could not be implemented.

The important point is that the subject believes he is modifying the game as a player, rather than influencing it as an adviser. In addition to this, the apparent importance of his part in the game is increased by a dummy supply line maintenance task. This dummy task, as was explained in Chapter 2, is designed so that, although it seems to be important to the game, it has no effect on combat results whatsoever.

In this way condition M3, requiring that a subject's involvement with the game be maintained, even under stress, is met, and an answer to our major reservation concerning the Intelligence Man system is found. We do not claim that this methodology will entirely eliminate role separation effects (it is impossible to achieve this entirely unless the subject is unaware that he is playing a game at all), but it is likely to reduce them considerably.

This leaves requirement M2, that the gaming system should permit a substantial level of complexity and detail. Clearly, this is the case. The Superior Commander system imposes no restriction on the detail of the game environment, nor on the complexity of apparent player interaction with it.

#### Initial Experience with the Superior Commander System

The first trials of the Organisational Control Game and Superior Commander system involved members of the Royal Holloway College Games Club. A scenario comprising six rounds was devised depicting two Task Forces defending against an advancing enemy. (This scenario, which later was designated "Scenario 1", will be described in detail in the next chapter; here, I wish only to comment on the performance of the methodology.) Subjects, in the roles of Divisional Commanders, had to advise two subordinate Task Forces and also decide on the deployment of a third Task Force which became available in Round Five (when it would be too late to affect the course of the game, which, unknown to the subjects, ended on the next round).

The deception was complete and worked well. Subjects were convinced that the other commanders were real players and they felt that they were positively influencing the game. We felt confident that if the deception worked well with these experienced players then it would be likely to work with most subjects, and subsequent experience, as will be seen, seems to have confirmed this prediction.

Subjects felt that they were fulfilling a useful function in their communications with their Task Force Commanders, and they did not feel helpless and peripheral in their roles. The dummy supply line

task was seen as important, and successfully gave subjects the illusion that the game was affected by their handling of the task, even though it was soon realised that it was straightforward. The problems with game involvement that are inherent in Intelligence Man seem to have been reduced quite considerably.

Subjects reported enjoying the game. For experienced war gamers, it was a novel experience to be able to forget entirely about game mechanics and concentrate entirely on strategies. The piecing together of information from various sources to make a coherent picture was mentioned as being particularly interesting, and, in general, the idea of a hierarchical game was found to be very appealing.

### 3.5 War Game Design

Our early experiences with board war games led us to state nine requirements which should be met if these games are to be used for research. Six requirements related to the form of the game itself; the other three were concerned with the methodology to be employed to control and replicate the game. Subsequently, we developed a game, the Organisational Control Game, and a methodology, the Superior Commander system, which satisfied these requirements.

In this chapter the only commercial wargame I have discussed in detail has been "Panzer Leader". However, we looked at and played a variety of board war games, exhibiting many different features, in our preliminary work. From the range of games we studied, we extracted those characteristics of physical design, layout and rules, which were most suited to our needs. I recommend that anyone seriously

considering using a board war game for research, whether one that is commercially available or one, like ours, which is specially developed, should play a variety of different commercial games to obtain an understanding of their capabilities.

To assist in selecting appropriate games, I recommend Palmer (1977) and the Strategy and Tactics Staff Study (Strategy and Tactics, 1977). Palmer provides detailed notes on nearly 300 professionally-produced war games, and is, as far as I know, the only comprehensive book on board war games, their characteristics, and how to play them. The Strategy and Tactics Staff Study gives notes on a basic library of about 40 war games, and lists many others; the book is concerned mainly with conflict simulation games.

There are at least two methodologies for the control of research games, the Intelligence Man system of Sharp and Dando (1977) and the Superior Commander system of Cooper (1978, 1979) discussed here. With an appropriate system for their control, board war games can be used to examine decision-making in a variety of research contexts.

This chapter has been concerned with the practicalities of board war game design. I have essentially been attempting to specify the nature of the content of a research war game. In the following chapter I shall add substance to these specifications by describing the details of our Organisational Control Game. Some idea of these is necessary to fully appreciate the results to be described later.

## CHAPTER 4

### THE ORGANISATIONAL CONTROL GAME

#### 4.1 Introduction

In this chapter I shall describe in detail the Organisational Control Game, which was developed by myself and Dr. Dale F. Cooper as a research game which implemented the Superior Commander system. A complete description of the game, including copies of all administrative material, is given in our final project report (McDowell et al, 1979b), and it would, in theory, be possible to implement and run the game entirely from the specifications given in that report. Some notes on the development of the game are included in a previous report (McDowell et al, 1979a). I shall not reproduce the entire contents of these reports here. Rather, I intend only to describe the game in sufficient detail to make the context of the results to be presented in subsequent chapters clear.

In the previous two chapters, the Superior Commander gaming system was introduced, and it was shown how this system was implemented in the Organisational Control Game. In this context, many ideas about the content of a research game were developed. This material should be borne in mind when reading the description of the Organisational Control Game presented in this chapter.

#### 4.2 Overall Structure of the Game

The experimental subject in the game takes the role of a Divisional Commander in a hierarchy of players consisting of a Corps Commander, three Divisional Commanders, and six Task Force Commanders

(Figure 4.1). As far as the experimental subject is concerned, he has as a superior the Corps Commander, and as subordinates two Task Force Commanders; no direct lines of command link him to any of the other players ("peripheral" players).

It will be recalled that, in reality, none of the other players exist, and the course of the game is entirely predetermined. The Divisional Commander is therefore told, in the course of an extensive briefing before play (McDowell et al, 1979b, Annex B), that he will remain isolated from the other players during play - all communications will be written, or tape-recorded - and, in fact, every attempt will be made to prevent him ever meeting the other players. This, it is explained (p.B-3), is "to keep personal interactions to a minimum". The Divisional Commander is also told that his two subordinate Task Force Commanders will be played by "experienced graduate students", and that all other roles (Corps Commander, peripheral players, and the enemy) are played by members of the Game Control staff. No explicit information is supplied about the structure, if any, of the enemy team.

The "cover story" to this rather elaborate organisation is that the game is designed to aid the study of the effect of "information flow within hierarchies" on decision-making (p.B-2). It is pointed out that "the information received by the decision-maker on the higher level is not raw data taken direct from the environment, but has been collected and assessed at the lower level, and should therefore be in a form more immediately useful to the high-level decision-maker". Further, "because this information is reaching the higher level decision-maker second hand, it may present an incorrect or at least biased view of the environment, and the



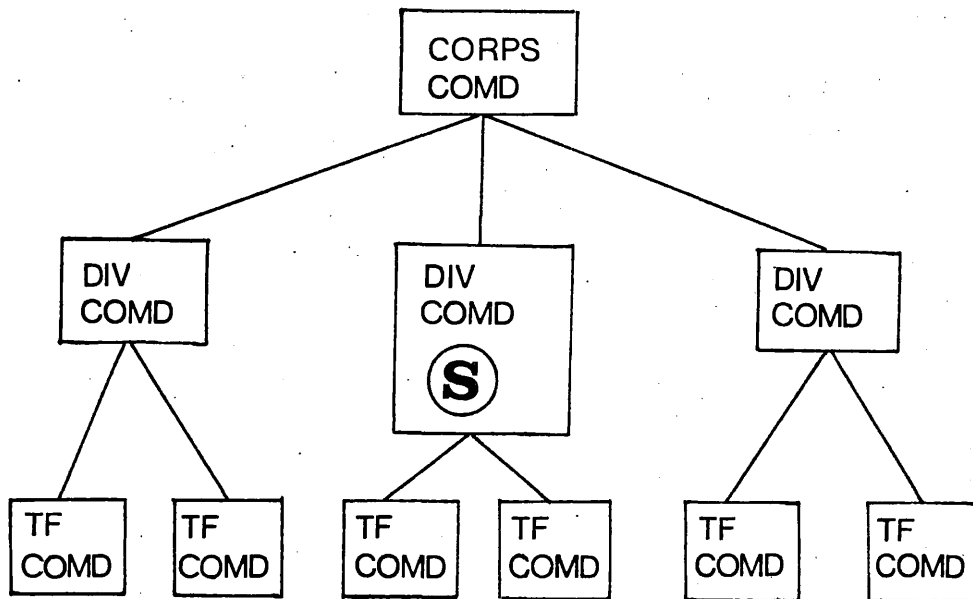


Figure 4.1. Structure of the friendly hierarchy in which the experimental subject ( **(S)** ) considers himself to be playing.

decision-maker has no way of testing this information by direct comparison with the environment."

The game is a time-stepping game - that is, it proceeds in a series of rounds representing successive game times (as opposed to real times). Each round represents a game time two hours ahead of that represented in the previous round. In a simple time-stepping game, players would make their moves for a period, and then the Game Controllers would assess the outcomes of those moves, eventually presenting the consequent situation at the beginning of the next period. However, because of the structure of the Organisational Control Game, this simple system must be adapted somewhat. Each round is therefore split into two phases: the Divisional Commander is told that these are his Task Force Communication Phase and Corps Communication Phase.

In the Task Force Communication Phase, the subject receives reports from his Task Forces, giving details of their situations. Reconnaissance information is also supplied. On the basis of these reports, the Divisional Commander issues directives to his Task Forces.

In the Corps Communication Phase, the Divisional Commander receives a communication from his Corps Commander, to which he is expected to reply. While this is going on, he is told, his Task Force Commanders are carrying out his directives and engaging in combat with the enemy, so that, at the end of the phase, it is possible to move to the beginning of the next round.

This two-phase system is further elaborated by the addition of

two other phases, which, though not essential to the "logic" of the game, contribute importantly to the game in other ways. These phases will be described shortly.

#### 4.3 Mode of Play

The Organisational Control Game is a board war game. The Divisional Commander is presented with a map board which covers his area of interest in the game - the Divisional area. The map board shows broad details of terrain features such as rivers, woodlands and hills, as well as towns, roads and railways. It represents an area of about 26 x 13 miles, on a scale of about 1:40,000.

As is conventional in board war games of this type, the map board is overlaid by a hexagonal grid, the hexagons (known to war gamers as "hexes") being used to specify positions of units. Each hex represents an area of approximately one ninth of a square mile.

Units are represented by means of small, square counters with identifying marks. Blue counters represent friendly fighting units, red counters represent enemy fighting units. Friendly units of supply, which, as will be explained, have an important role in the game, are green. Fighting units are represented at battalion level. Since just four battalions make up a friendly Task Force, it will be realised that combat cannot be represented in great detail at the Divisional Commander level.

However, the Divisional Commander is told that the Task Force Commanders have a different view of the game: they are supposedly presented with detailed map boards showing only that area of terrain

immediately concerning their own units. Details of combat at the Task Force level are not supplied to the Divisional Commander; however, he is left in no doubt that it takes place at a significantly greater level of detail than that at which he perceives the game (for example, at company level or lower).

A Divisional Commander is also supplied with maps which show his Divisional area in relation to surrounding Divisional areas, and in relation to the area of country in which the conflict is taking place. These aspects of the game environment are emphasised during the briefing prior to the start of the game.

#### Rules of Play

As explained in the previous chapter, the Divisional Commander need not concern himself with the details of combat, since ostensibly this is carried out by his Task Force Commanders at the lowest level of the command hierarchy. The Divisional Commander is therefore given only rules of play pertaining to the units as they are portrayed on his map board - that is, at battalion level.

The Divisional Commander is told that his Task Forces are predominantly infantry. In one time period, a battalion can move up to a maximum of six hexes on a road, but this movement capability is cut back on rough ground or when the unit is engaged in combat, or under enemy fire. The fighting strength of a unit is increased if it remains stationary for sufficiently long to "dig in", and its fighting capabilities are also enhanced if it occupies a town or high ground. It is made clear to the Divisional Commander that the capabilities of enemy units are not precisely known, although they are not likely to

differ radically from those of the friendly units.

The course of the game is predetermined. We had decided that scenarios would be designed in which the component battalions of a Task Force would remain in close support of one another. It occurred to us that if a Divisional Commander were to issue a directive to the effect that a Task Force be split, failure to comply would be particularly difficult to account for in terms of obstructions at the Task Force level. To discourage a Divisional Commander from attempting to split either of his Task Forces, therefore, it is explained that if a Task Force is dispersed beyond a certain limit, its fighting strength becomes severely reduced, due to difficulties with communication and coordination of activities.

#### Dummy Task

It will be recalled from the previous chapters that an important part of the Organisational Control Game is a "dummy" supply line maintenance task which, although it has no real affect on the game, gives the illusion of being essential and thus increases the player's involvement in the game. We decided that we wanted to keep this task separate in time from other parts of the game, and so a Supply Phase was introduced between the Task Force Communication Phase and the Corps Communication Phase in each round. The Supply phase is devoted solely to administration of the supply task.

Rules governing the administration of the supply task are simple. Combat units expend a certain amount of supply in each round; this amount increases in heavy combat. Supply is represented by supply unit counters on the map board. The Divisional Commander is told

that the combat effectiveness of fighting units will drop if they exhaust all their supply. He therefore has to keep his Task Forces supplied with units originating at a supply depot in the rear of his Divisional area. For this task he is required to move supply units on his map board - this is the only part of the game in which he is permitted to affect the position of counters on the board. The Supply Phase is relatively short - five minutes is generally more than adequate.

The supply task is straightforward: scenarios are designed so that there is very little danger of Task Forces running out of supply due to cutting of lines of communication or unexpectedly heavy combat. Consequently, the Divisional Commander is, in practice, never faced with a situation in which his forces have no supplies. This is essential, because, due to the predetermined nature of the game, even if forces did run out of supplies, no effect on the game would be detectable.

### Research Phase

As a game, the Organisational Control Game is complete in its three-phase form, each round comprising a Task Force Communication Phase, followed by a Supply Phase, followed by a Corps Communication Phase. The communications which the Divisional Commander sends out to both Task Force and Corps provide a considerable amount of output suitable for analysis.

However, we found that we required output in addition to that supplied in the course of the game. We wanted to ask the Divisional Commander specific questions about his decision-making, the answers to

which would not, in general, be made clear in the "natural" output of the game.

To allow for this additional questioning, a Research Phase was added at the end of some rounds, following the Corps Communication Phase. In this phase the Divisional Commander is required to reply to direct questions from the experiments. It is made clear, of course, that this phase has no effect on the course of the game.

#### 4.4 Scenarios

Two different combat scenarios have been designed for use with the Organisational Control Game. They are both set within the context of a broad general political scenario. I shall briefly outline this general scenario before considering the two combat scenarios in detail.

The general scenario is based in the Republic of Arachide, a fictitious developing country with a strong agricultural economy (Figure 4.2). Arachide has two neighbours: the States of Basile and Cive. Relationships between Arachide and its neighbours have been relatively cordial until recently, when the Seurat region of Arachide, the subject of a long-standing territorial dispute between Arachide and Basile, was found to contain a potentially large oilfield. What had previously been a "largely academic argument" (from the game briefing as reproduced in McDowell et al, 1979b, Annex B, p.B-5) suddenly "flared into a major political dispute".

The dispute escalated, diplomatic relations between the two countries ceased, and Basile mobilised its forces, called up reserves, and requested foreign aid. However, Arachide reacted to this activity

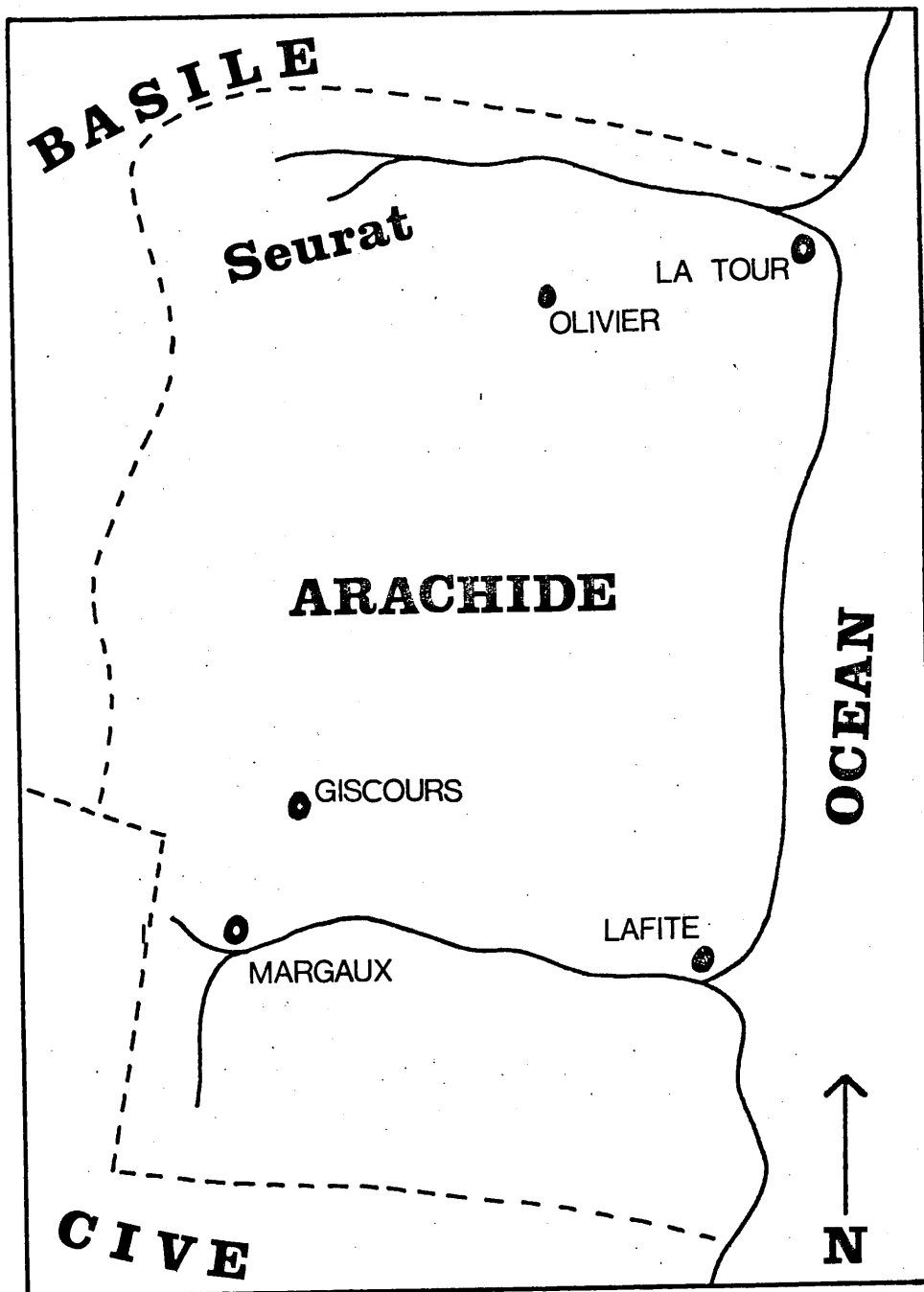


Figure 4.2. The fictitious Republic of Arachide.



rather slowly, and consequently was caught with little defence when, a few days later, Basile launched a large-scale invasion.

The combat scenarios are set one week after the outbreak of hostilities, at which time Basile has made considerable progress into Arachide territory, having met little resistance in some areas, but in others having run into considerable opposition from hastily deployed defensive forces. It seems that Arachide is now beginning to successfully check the advance of Basile's forces, although this is by no means certain.

The two combat scenarios are located in widely separated regions of Arachide. Subjects are allocated the roles of Divisional Commanders in the Arachide army, taking part in defensive activities one week after the invasion by Basile. Despite the commonality of their aims, however, the players in the two scenarios find themselves in completely different situations.

#### Scenario 1 - A Defence Scenario

Scenario 1 takes place in an area to the West of Arachide, to the North of the town of Giscours (Figure 4.3). The area is predominantly hilly, and many small towns are scattered among the hills. At the time the Divisional Commander starts play, the area is subject to a strong enemy thrust from the North-West, part of the enemy's overall push to the East. The thrust immediately threatens Giscours and important rail links to the North and West. The friendly Corps has taken up defensive positions in the area, its aim being to prevent the enemy achieving its objectives for as long as possible - hopefully until friendly reinforcements arrive.

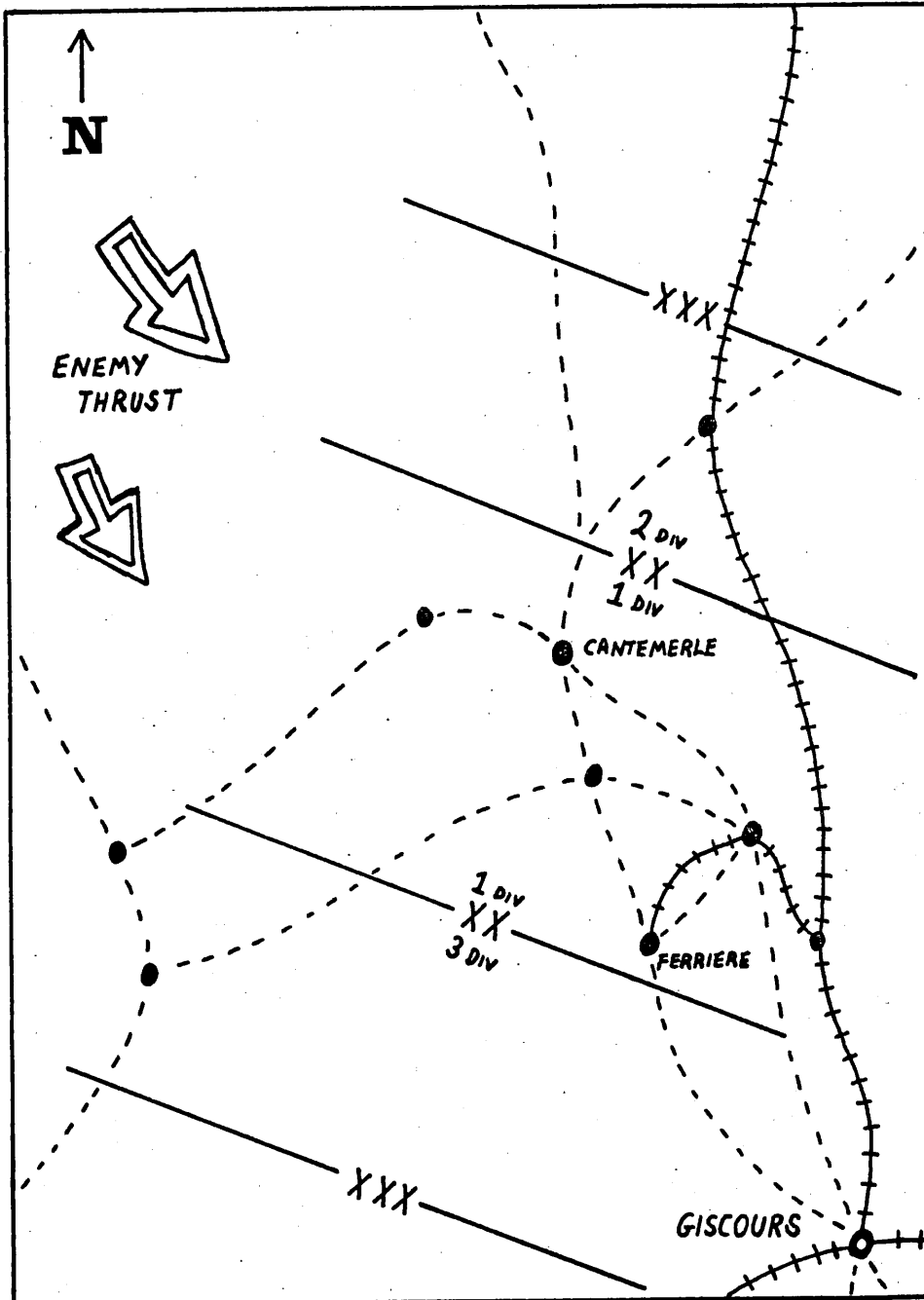


Figure 4.3. The setting for Scenario 1.

The Divisional Commander's division (Figure 4.4) is located in a central position, with a friendly division located on each flank. The Division consists of three Task Forces: two are deployed in forward defensive positions and are under the command of Task Force Commanders; the third is grouping, but is expected to become available later as a strategic reserve, with deployment under the direct control of the Divisional Commander.

At the beginning of the game, no contact has occurred between the Division and the enemy, though it is clear that this will occur very shortly. The Divisional Commander is issued with a directive (McDowell et al, 1979b, p.B-11) instructing him to "delay the enemy advance ... for as long as possible", preventing "any advance towards Giscours"; to "hold a defensive position" in the area which his Task Forces currently occupy, denying the enemy access to towns West of the area; and to "prevent the enemy from reaching and blocking the railway link to the North".

As the game progresses, the forward Task Forces make contact with the enemy and battle starts. The Task Force in the South comes under pressure, but it is the Northerly Task Force which suffers the heavier attack, with one of its battalions completely wiped out. Both Task Forces attempt to break contact and withdraw in order to stronger positions but find this a difficult operation to carry out. While the Task Forces are being kept engaged, fast-moving enemy units slip past their positions and begin to pose a direct threat to locations in the West of the Divisional area. By the time the game ends, and the Divisional Commander can deploy his reserve Task Force, as well as an additional Task Force which he has meanwhile been allocated, his forward Task Forces, considerably reduced in size, are

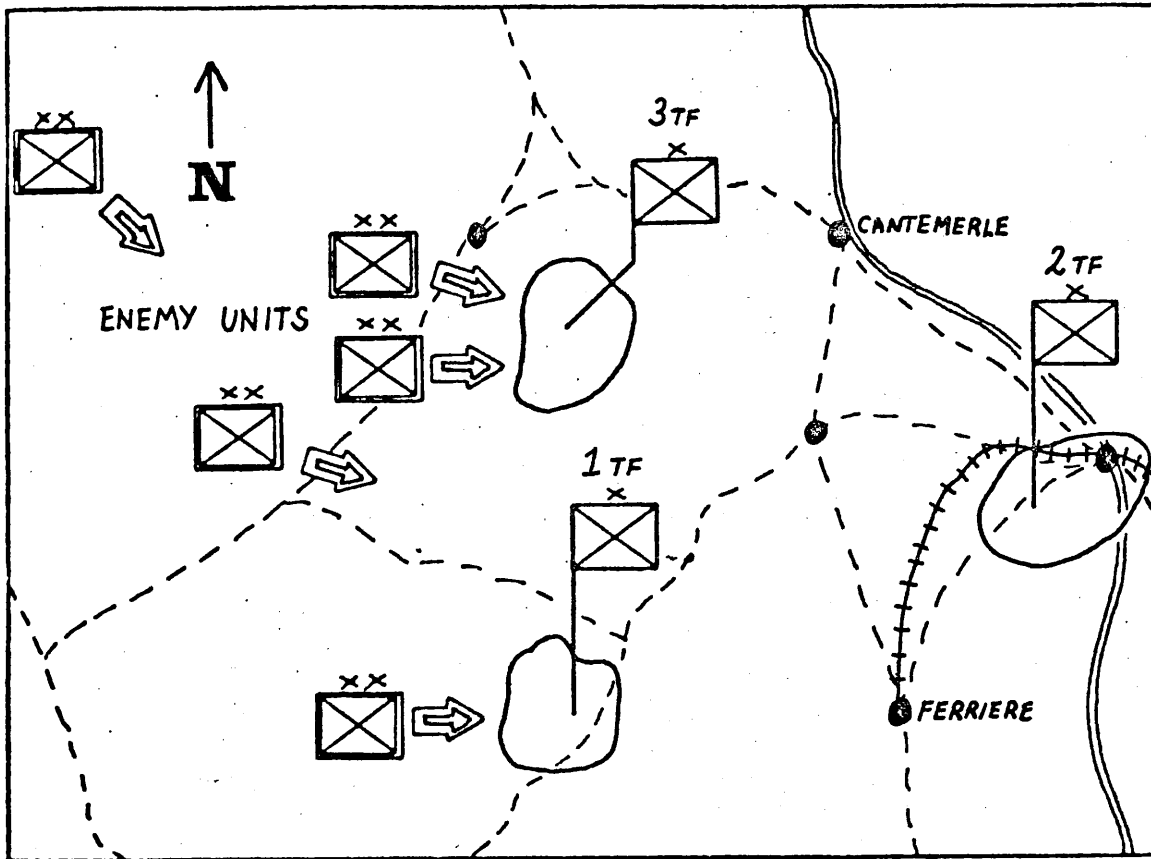


Figure 4.4. The subject's Divisional area in Scenario 1.

defending positions against enemy units some miles to the west of their original locations, while other enemy units are still moving West without opposition and threatening to cut the forward Task Forces' supply lines.

The overall picture in Scenario 1, therefore, is of a defence against an enemy force which is clearly too strong to repulse - the best that can realistically be done is to delay the enemy advance as long as possible. The true strength of the enemy, and the corresponding weakness of the friendly position, only becomes clear as the game proceeds.

#### Scenario 2 - An Advance-to-Contact Scenario

Scenario 2 takes place in the North of Arachide, in an area centred on the town of Olivier (Figure 4.5). The area is dominated by the valleys of the Harris River and its tributary, the Jackson. At the time the Divisional Commander starts play, the enemy is advancing South-West down the Jackson Valley towards Olivier, and, indeed, its advance units have already moved virtually unhindered into Olivier. The friendly Corps is in the process of moving into the area from the South, its general objectives being to free Olivier and keep the railway line through the area open.

The Divisional Commander's division (Figure 4.6) is once again located in a central position. Moving in from the hills to the West of Olivier, one Task Force is heading for Olivier with the intention of freeing it, while another is advancing towards the town of Bouscaut, from where it intends to move East to meet the approaching enemy. These two Task Forces are under the control of Task Force Commanders; a third Task Force is grouping in the hills West of

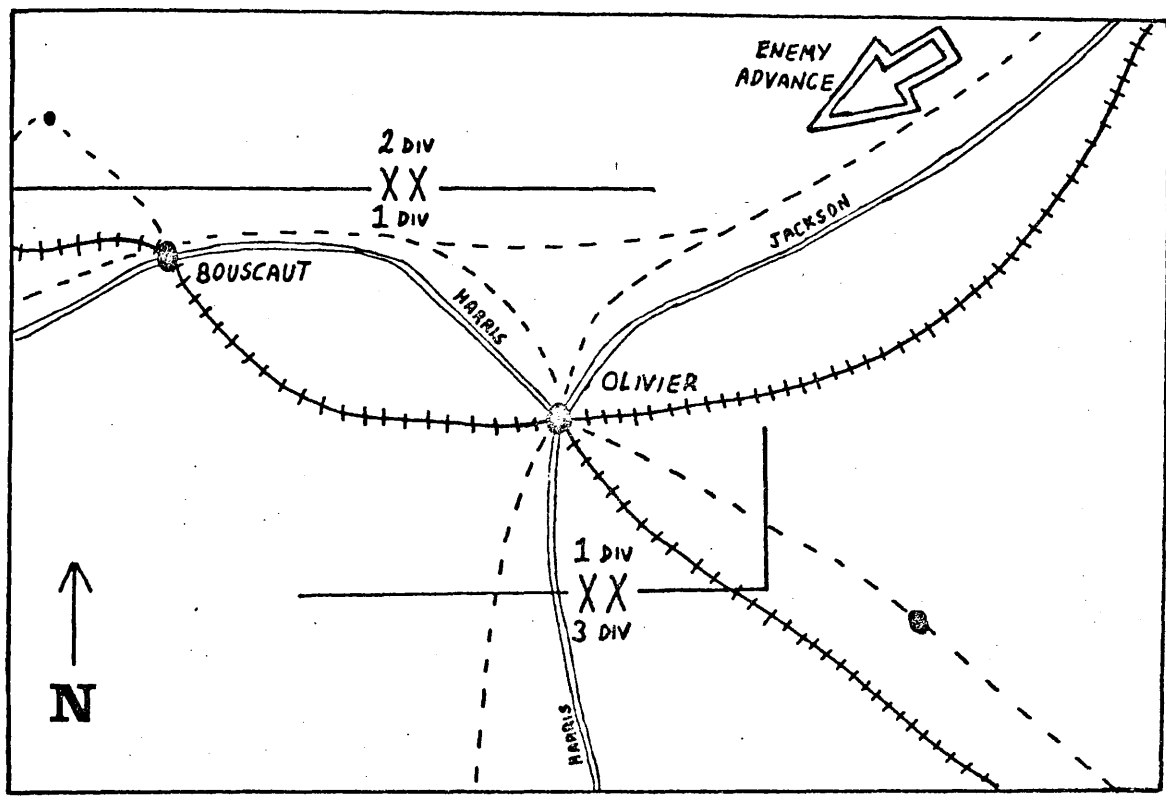


Figure 4.5. The setting for Scenario 2.

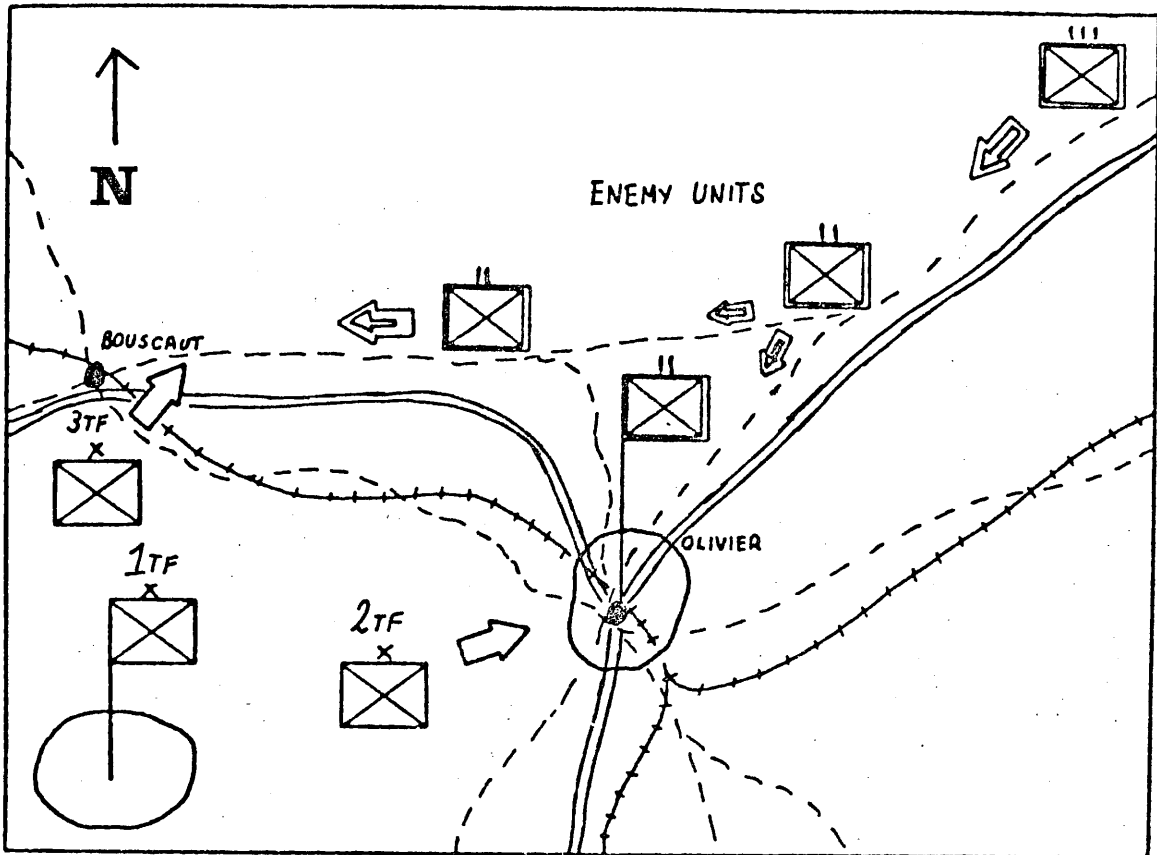


Figure 4.6. The subject's Divisional area in Scenario 2.

Olivier, but is expected to become available as a strategic reserve, with movement under the direct control of the Divisional Commander, later in the game.

At the beginning of the game, no contact has occurred between the Division and the enemy. The directive to the Divisional Commander (McDowell et al, 1979b, p.E-11) instructs him to "establish control of Olivier and the area immediately around it, making sure that the railway line between Olivier and Bouscaut is kept clear".

As the game progresses, the Task Force assigned to Olivier approaches the town. Enemy forces in Olivier turn out to be greater than expected, but, after a fierce fight, the Task Force manages to occupy a considerable proportion of the town. However, even by the end of the game, Olivier is not entirely free of the enemy, and the battle continues. Meanwhile the other forward Task Force, having reached and passed through Bouscaut, has met with enemy opposition a few miles to the East, and by the end of the game is making some impression on the enemy's strong hill position. The overall situation is made more serious by the fact that further enemy units are rapidly approaching the area. Once again, by the end of the game, the reserve Task Force, as well as an additional Task Force, can be deployed on the orders of the Divisional Commander.

Thus Scenario 2 is an advance-to-contact, in which what might at first seem a rather straightforward operation is complicated when it becomes apparent that the enemy are stronger than at first they appeared. Once again, the true situation only becomes apparent as the game proceeds.



#### 4.5 Playing the Game

In this section I describe briefly the practical arrangements for play of the Organisational Control Game. We had decided that the whole game, including briefing, debriefing, and appropriate refreshments, should fit into one working day. We had facilities prepared for two subjects to play simultaneously, one for each scenario - as a rule, all our games were played in this "parallel" fashion. Such an arrangement can be fully administered by just two Game Controllers: each with responsibility for one scenario, and both sharing responsibility for other duties. However, since it is intended to make the players believe that quite a large team is involved in playing the game, short "walk-on" appearances by any convenient colleagues are a useful aid. Indeed, such incidental events are probably quite important in maintaining the Superior Commander deception successfully.

A few days prior to the game, a player would have been sent a set of documents introducing the game which he would have been asked to read before coming to play. The set of documents consists of:

1. Introduction to the Organisational Control Game, describing briefly the hierarchical control system.
2. Brief summary of rules for Divisional Commanders in the game.
3. Background to the general scenario - a brief description of the Republic of Arachide and its recent history.

4. General background to the player's combat scenario.
5. Notes on the player's combat scenario.

The entire set of documents covers ten sheets of A4 paper (typed and single-spaced), including maps. The documents are reproduced in McDowell et al (1979b), Annex A.

A typical timetable for the game is presented in Table 4.1. As can be seen, the game lasts for six rounds, and Research Phases are appended to the first, third and fifth rounds. The total time required according to the schedule is 4 hours and 40 minutes. A half hour coffee break is usually taken after the general briefing, and a break of an hour for lunch may be taken after the second or third round. If these breaks are taken, total running time will be 6 hours and 10 minutes. Thus, if the game begins at 10.00, it should be over by 16.10.

Players from both scenarios attend the general briefing together. This briefing (McDowell et al, 1979b, Annex B) first discusses the hierarchical control system. It then goes on to describe in detail the Republic of Arachide, and gives details of the escalation of the conflict with neighbouring Basile. The two areas depicted in the combat scenarios are briefly discussed before the rules and mechanics of the game are described. A number of maps and diagrams are shown in the course of the briefing.

Immediately before play begins, a player receives a short briefing in his own game room. This briefing is intended to familiarise him with the use of the equipment of the game room.

		<u>Duration (Minutes)</u>
General briefing		30
Individual player briefing		10
Round 1:	Board update	5
	Task Force Communication Phase	15
	Supply Phase	5
	Corps Communication Phase	10
	Research Phase	10
Round 2:	Board update	5
	Task Force Communication Phase	10
	Supply Phase	5
	Corps Communication Phase	10
Round 3:	Board update	5
	Task Force Communication Phase	10
	Supply Phase	5
	Corps Communication Phase	10
	Research Phase	10
Round 4:	Board update	5
	Task Force Communication Phase	10
	Supply Phase	5
	Corps Communication Phase	10
Round 5:	Board update	5
	Task Force Communication Phase	10
	Supply Phase	5
	Corps Communication Phase	10
	Research Phase	10
Round 6:	Board update	5
	Task Force Communication Phase	10
	Supply Phase	5
	Corps Communication Phase	10
Debriefing		25

Total duration: 4 hours, 40 minutes

Table 4.1. Specimen game timetable, with Research Phases in Rounds 1, 3 and 5.

The equipment in the game room consists of : a detailed map board of the player's Divisional area, with counters representing combat and supply units; a wall map of the Corps area and a wall map of the whole of Arachide; and a tape recorder and several cassette tapes with which to record all messages to Task Forces and to the Corps (and also any reports made in the Research Phase). Small scale reproductions of the map boards, and a list of counters, are presented in Annex D of McDowell et al (1979b). The map boards for each scenario cover the areas shown in Figure 4.4 (Scenario 1) and Figure 4.6 (Scenario 2).

Play proceeds in a series of rounds. Each round begins with a board update by a Game Controller, during which counters are moved on the map board to correspond with combat at the Task Force level in the preceding round. Then the Task Force Communication Phase begins. Written messages are brought, by the Game Controller, from the Task Forces, describing their situations, and the Divisional Commander responds by issuing a directive to the Task Force Commanders which he records on tape. A report also arrives from reconnaissance units in this phase. At the end of the phase, the tape which the Divisional Commander has prepared is taken away.

The Supply Phase occurs next, during which movement of supply by the player should be noted by the Game Controller. It is important that this is done, since the player must believe that the supply task is affecting the game at the Task Force level, though of course it is not.

In the Corps Communication Phase, a written message is brought by the Game Controller from the Corps Commander. Once again the player

records a response on tape, which is taken away at the end of the phase. (However, if a Research Phase is played at the end of the round, the same tape can be used for both the report to Corps and the research response.) Overlays depicting movements of units in the Corps area are supplied for the Corps area map.

In the Research Phase, the player would be asked to respond to questions such as the following:

1. What different plans did you actively consider before choosing the one that you did?
2. Were there any plans that you felt were not worth serious consideration? What were they?
3. How did you select the plan that you did? What factors in the scenario influenced your decision?
4. What different courses of action do you see as open to the enemy at this stage? Which do you think he will choose? What factors influenced your appraisal?

All communications to the subject (including Corps area map overlays) are prepared before the game begins. The Communications are hidden in a control room until required. Messages from Task Force Commanders are handwritten; messages from the Corps Commander and reconnaissance units are typed. Copies of all messages, as well as instructions for movements of unit counters, are included in Annex D of McDowell et al (1979b).

The debriefing has no formal form. Players are not always immediately told of the Superior Commander deception, since they may be required to play again, and their colleagues may be future subjects. Essentially, therefore, most of the debriefing is concerned with asking them for their views and opinions of the game. The type of questions asked might include:

1. How adequate were the data coming from the Task Forces?
2. How adequate were the instructions from Corps?
3. How well do you think the Task Force Commanders played? Will they achieve their objectives in a reasonable number of rounds?

Notes for the debriefing are presented in Annex E of McDowell et al (1979b).

#### 4.6 Notes on the Development of the Game

Development of the Organisational Control Game into the form in which it has been described in this chapter took place in three series of games. The first were the "Student" games in which the experimental players were students at Royal Holloway College. These games were intended to test the viability of the Superior Commander system and the overall plan of the Organisational Control Game and were not envisaged primarily as decision-making research games. These games were referred to in the previous chapter.

The second series of games were the "Emergency Planning Officer" games, which were intended partially as development games, but were also envisaged as providing some research output. The experimental players in these games were Emergency Planning Officers from various counties. The research aspects of these games will be discussed in Chapter 7.

The third series of games were the "Major" games which were intended entirely for research purposes. These games were played by serving army officers with the rank of major, and will be discussed in full in Chapters 7 and 8.

The game as it has been described in this chapter is essentially that played in the final "Major" series. In this section I shall briefly mention some of the developments that took place in the first two series.

#### Design of Scenarios

The first combat scenario we designed was Scenario 1, the defence scenario. Scenario 1 was used for most of our "Student" games, but we soon realised that, if our research results were not to be "scenario-specific" we would have to have the ability to test our results across different scenarios. Consequently we designed Scenario 2, the advance-to-contact, as a deliberate contrast to Scenario 1.

We were concerned that in Scenario 2, where the friendly side have apparently greater freedom of action than in Scenario 1, there would be increased danger of friendly movements differing inexcusably

from that specified in players' directives. In fact, this does not seem to be a great problem. We found that in the advance-to-contact, provided the tasks of the Task Forces are made sufficiently clear at the beginning of the game, difficulties arising from players developing totally different overall strategies from that which actually is played are small; indeed, it turned out that there was less variation in strategies than in the defence scenario. This seems to be because the instructions to the player can be far more precise in the advance-to-contact scenario. Objectives can be specified and time limits set in some detail. All this is acceptable when the enemy do not have the initiative. In the defence case, where the enemy do have the initiative, instructions must be more general. There is a higher possibility of a player's strategy differing from that which occurs in the game.

#### Modifications to Scenario 1

We found it necessary to make some modifications to Scenario 1 during our first two series of games. Specifically, we found that, although virtually all players recognised the need for a withdrawal of forces during the defence their opinions differed as to when the withdrawal should be initiated, its speed, and its precise direction. This became particularly clear in our second series of games. To meet this problem we modified the scenario in a way which seemed to conform most readily to the apparent expectations of players. Thus the scenario was developed from game to game using both the behaviour of players during the game and their comments afterwards. This method of game development - that of feeding back information supplied directly and indirectly by players into the scenario - is essentially that described by Brewer (1978).



Since players differed considerably in their styles of withdrawal, it was clear that any one type of withdrawal would not satisfy all players equally. It would conform more to some players' ideas than others. We had to make sure, therefore, that in the prepared communications from Task Force to Division, Task Force Commanders explicitly stated why they had behaved in the way they had. In cases where players' directives were not carried out, these reports by the Task Forces were to serve as the only explanation the player would get about why this had occurred. Since the communications were written without any knowledge of an individual player's intentions, every possible set of intentions had to be covered.

The present form of Scenario 1 differs very little from the original; it was only at the lowest level - the movement of individual battalions - that changes were made. Nevertheless, Scenario 1 in this form is very much more satisfactory to players than before, and events are questioned far less than in the original version.

We found that no modifications were necessary to the initial design of Scenario 2. This is probably due to the experience we gained during the design of Scenario 1.

#### Alterations to Game Rules

We found that we had to make the rules of combat in the game more explicit, and make unit movement rules more complex, than we originally intended. As far as these factors are concerned, it would seem that our experience with student players in the first series of games positively misled us - student players accepted grossly simplified rules, and, indeed, commented that they were thus enabled

to concentrate more on game strategies rather than mechanics. We found this encouraging. However, players in the second series of games, who had military experience, did not accept that these rules provided anything like an accurate representation of the type of combat we were gaming. Consequently, the rules were altered to their present form, reflecting in more detail such features as the effect of terrain on combat. We found that players with military staff experience had no difficulty in assimilating these more complex rules. We believe this to be because of the clear correspondence between the rules of game and the natural "rules of reality".

#### Dummy Task

In both scenarios as they now are, the dummy task is a supply task, involving the maintenance of supply links from a depot to the Task Forces. This task has always been used in Scenario 1. In Scenario 2 we originally developed an artillery allocation dummy task: players were required to allocate medium artillery fire support to locations within their Divisional area.

There were two main reasons why the artillery allocation task proved unsuitable. Firstly, since the task could have no real effect on the course of the game, we were unable to provide any feedback to the players concerning the success of their fire support. They would have grown suspicious and frustrated had we repeatedly told them that the artillery was having no effect - that they were, in fact, consistently missing their targets. Secondly, we found that the artillery task was interfering with the players' perception of the battle that was being gamed. The decision-making involved in the artillery allocation became associated with that involved in

controlling the Task Forces. The artillery task was effectively an uncontrolled variable in the game.

Some brief experimentation with an air strike allocation dummy task indicated that this too was unsuitable, for similar reasons to those mentioned above in connection with the artillery task. The supply dummy task has, however, proved suitable for use in both scenarios; an additional advantage of it is that it requires the Divisional Commander to think about lines of communication, future plans and enemy intentions, and encourages a strategic view of the game.

#### Other Notes

It will have been noticed, in the descriptions of the combat scenarios, that both scenarios end with the deployment of reserve combat forces by the Divisional Commander. At first sight this might seem to contradict a fundamental principle of the Superior Commander system: that nothing the Divisional Commander does can really affect the game. However, the scenarios are so designed that the reserve forces cannot engage in combat before the game ends. Thus, no real effect can occur. At the same time, the reserve deployment decision may be an interesting decision to examine. This final "free" decision is the single decision of the single-decision game structure being used (see Chapter 2).

The timing of phases as presented in Table 4.1 was arrived at as a result of the early game trials. These timings were found to give players adequate time to complete their tasks in each phase, without leaving them with unacceptably long periods of inactivity. Thus the

timings in Table 4.1 are not likely to stress or disturb players and so affect their decision-making behaviour.

Tape recorders to record all communications sent out by experimental subjects were only used in the third "Major" series of games. In the other two series, all communications were written. It was felt, particularly in the "Emergency Planning Officer" games, that some detail was being lost by requiring players to give written directives and reports - these communications were sketchier and less thorough than we had hoped for. The tape-recorded messages were undoubtedly a major improvement in this respect, and the players in the third series of games apparently felt unrestrained regarding the level of detail incorporated in their messages.

It should be noted that in the third series of games two alternative sets of reports from the Task Forces to Divisional Commanders were prepared. One set was as described in this chapter - handwritten reports from the Task Force Commanders. The other set contained the same data, but was presented in the form of short handwritten reports from the individual battalions (a Task Force consists of four battalions). Thus data from the subordinate level could be presented in one of two conditions: an aggregated condition (reports from Task Forces), and a disaggregated condition (reports from battalions). The reasoning behind the preparation of these two conditions, and their use, will be discussed in Chapter 7. Both aggregate and disaggregate communications are presented in Annex D of McDowell et al (1979b).

I deliberately chose to describe the Organisational Control Game and its development separately rather than include these descriptions in the accounts of decision-making research which are to follow in

Chapters 7 and 8. This split emphasises the fact that the material in this thesis is concerned not only with decision-making research but also some important aspects of the design and development of research games. So far it has been the game development aspect of my work with which I have been mainly concerned.

In Chapters 7 and 8, I shall direct my attention to the use of the Organisational Control Game for decision-making research. Before this, however, I wish to turn to another part of our research programme. The next two chapters will be concerned with our work on secondary task methods for assessing mental processing load.

## CHAPTER 5

### SECONDARY TASK METHODS FOR ASSESSING MENTAL PROCESSING LOAD - A LITERATURE SURVEY

#### 5.1 Introduction

This chapter and the chapter which follows describe our subsidiary research programme to develop a method of measuring the mental processing load imposed by complex problem-solving tasks. It will be recalled, from Chapter 1, that when we began our gaming research we felt that we might wish to measure the processing load imposed by various game scenarios. Accordingly, we decided to investigate the use of a psychological technique, the secondary task technique, as a means of doing this.

Our gaming research developed in a different direction to that which we had initially envisaged, and consequently we did not find it appropriate to apply the secondary task technique to our game as we had originally intended. However, we carried out a substantial amount of work in developing a secondary task method suitable for our requirements, and, as will be shown, we have achieved an encouraging degree of success.

Secondary task methods have been used frequently in ergonomic and psychological research for over twenty years, and the majority of this chapter is concerned with reviewing the literature in this field. In Chapter 6 I shall describe an experiment we carried out at Royal Holloway College to adapt the secondary task methodology to the assessment of the mental processing load imposed by complex problem-solving tasks.

## 5.2 Secondary Tasks

Secondary task experiments are characterised as experiments in which human subjects are required to perform two tasks - primary and secondary - concurrently. In general, secondary task experiments have been directed towards the investigation and measurement of the mental processing load imposed on subjects performing specific primary tasks (our own work is no exception to this general rule). "Mental processing load" is a concept which we shall shortly consider in more detail; for the time being it may simply be regarded as a measure of the amount of "brain power" required by a task. Thus a task is said to impose a certain mental processing load upon its performer.

The reasons for measuring the mental processing load imposed by tasks have varied considerably. At one extreme, ergonomists have been concerned with studying the load on subjects operating complex machinery, in which the positioning of controls and displays and the overall design may be crucial to satisfactory operation. In other experiments, the load has been due almost entirely to purely cognitive processes; psychologists have used such experiments to examine mental mechanisms in their subjects.

All secondary task experiments have been based on the assumption of a relationship between the imposed processing load and the standard of task performance. The nature of the relationship is derived from a capacity model of human information processing, which provides a theoretical basis for using task performance as a means of evaluating load. The capacity model postulates some "natural limit" to mental processing capacity: provided this limit is not exceeded, little decrement in task performance should be expected. Thus more often

than not there may be no observable change in the performance of a task as its complexity, and hence the mental processing load associated with it, is varied. For example, differences in the quality of car driving under conditions of light and heavy traffic are subtle and difficult to detect unambiguously, yet it seems obvious that the processing load on the driver must differ considerably in the two situations.

The use of a secondary task offers two different approaches to the problem of evaluating the processing load in cases where the performance of the task which is of primary interest does not vary under normal conditions. The first approach is to use the secondary task to increase the load on a subject, causing his performance on the primary task to deteriorate. This type of experiment, where the secondary task is a loading task, can be used to compare the difficulty of various similar primary tasks: the more difficult the task, the greater the deterioration that results when the secondary task is imposed. We would expect the more difficult aspects of the primary task to be the most sensitive to deterioration, and these aspects may be pinpointed using a loading task of appropriate load.

The alternative approach is to use the secondary task as a subsidiary measuring task. Here, the error rate in the secondary task is assumed to reflect the load due to the primary task, to which the subject is instructed to assign priority and which he should perform to the best of his ability at all times. This approach presents more experimental and theoretical problems than the loading task approach, but it appears to have a greater potential for studying the mechanics of processing associated with a wide variety of tasks.



In all the secondary task experiments I shall consider in this survey (which I believe constitutes a fair proportion of published work to date) the primary tasks always tend to be tasks in which the subjects have at their disposal some experience of responses appropriate to the task stimuli with which they are presented. In general, such tasks are either very simple or they are well-learned, and in these circumstances the processing of data is essentially directed towards the selection of a response rather than towards the construction of an appropriate "response model". The algorithms required to generate a response, if they do not already exist, can be created with little effort. In this respect the tasks differ considerably from complex problem solving or resolving tasks, in which solution paradigms do not exist already, and in which complex models must be created before a suitable response can be selected.

Reviews of secondary task experiments have been published at various times. Knowles (1963) described several secondary task experiments related to ergonomics and the operation of man-machine systems. Brown (1964) provided a more extensive survey of the ergonomic literature, emphasising the distinction between loading tasks and subsidiary measuring tasks, and highlighting the important differences between them. Rolfe (1971), in an extensive critique of secondary task experiments, reviewed a wider range of work, reflecting the emergent interest of experimental psychologists in the techniques, although he, too, was mainly concerned with the ergonomic literature. Kerr (1973) reviewed the psychological literature relating to the use of secondary tasks in evaluating processing demands during mental operations.

The present survey includes much of the literature covered by

these authors as well as more recent work. My approach is to consider individually the fields in which secondary task techniques have been applied. In the next section, I begin by describing the capacity model, which provides the theoretical basis for secondary task methods, and I outline briefly some recent developments to the model. Next I review experiments involving loading tasks, and then those using subsidiary measuring tasks. Finally, I discuss the interpretation of the results of secondary task experiments, and consider possible developments. This will lead into a consideration of our own secondary task work at Royal Holloway College, to be described in the following chapter.

### 5.3 The Capacity Model of Human Information Processing

Much of current thinking in the field of human information processing is based on some form of limited capacity model of the data and information processing mechanisms of the human brain. Broadly speaking, current versions of the model indicate that the concept of limited capacity is a valid and useful one, providing a framework for both theoretical development and experimental investigation of human information processing.

The capacity model is founded on the idea that effectively there exists in the human brain some kind of "central" information processing mechanism that is the location of some or all mental operations. The model is not concerned with physically locating the mechanism, or, indeed, with forming any hypotheses concerning its physical nature, but rather with elucidating the organisation of the processes involved in mental operations. In its present form, this mechanistic, computer-like view of thought processes is intended to serve only as a useful

integrated description of observed phenomena; the nature of the relationship between processes in the model and physical processes in the brain is not specified.

#### The Single-Channel Limited Capacity Model

Welford (1952) proposed a model of human mental processing as a single-channel system of limited processing capacity in order to account for phenomena associated with responses to stimuli closely spaced in time. He noted that when responses were required to two separate but similar stimuli, the response time for the second was longer than that for the first if the second followed the first sufficiently closely. Welford proposed that the central mechanisms of the brain required a finite "organising time" to deal with stimulus information, and that no two central organising times could overlap. As a result, information sometimes had to be held in store. This would cause an observable delay in the response to one stimulus which followed another sufficiently closely. A delay might also result due to feedback from the previous response, which itself would require central organising time.

Welford's model is a single-channel one in that no parallel processing of information can take place. The central mechanism can handle only one information processing operation at a time, and the finite organising time required by each operation effectively limits the capacity of the system, which becomes "blocked" for that duration.

Welford's model was reinforced by Davis (1957), who showed that delays in response to the second of two closely-spaced stimuli also occur when two sense modalities (in his work, auditory and visual) rather than

one are used for the adjacent stimuli, indicating that the single-channel bottleneck is associated with central processing mechanisms rather than sense receptors: information from both sense modalities must pass through the same channel. Later Davis (196?) showed that it was possible for information from a first stimulus to be stored and not processed until after a second stimulus had been processed, and his work also indicated that blocking of the single channel did not occur if a response to the stimulus could be so well learnt as to become automatic.

Broadbent (1958, 1971) modified the model by introducing a filter which would select only stimuli possessing some arbitrary common feature. These stimuli alone would be permitted to pass on to further processing. In this way selective attention phenomena would be accounted for: for example, the "cocktail party" phenomenon - the ability to attend to one individual's speech in the midst of several others. Broadbent's theory, and subsequent developments of it, are discussed by Kahneman (1973).

The simple single-channel model was considered unsatisfactory by Elithorn and Lawrence (1955), who argued that the concept of blocking of the channel was "nonproven and uneconomic" (p.116). They felt that evidence for Welford's theory was far from conclusive. Work by Adams and Chambers (196?) indicated that the single-channel model could not account for results obtained when the stimuli to which subjects were to respond were known by the subjects to be certain to occur. Creamer (1963) suggested that human data processing mechanisms operated as a single-channel system only when subjects were uncertain as to the type of event to which they were expected to respond.

## Modifications to the Single-Channel Model

Since the single-channel model was proposed, it has been developed in directions that have resulted in models which concur to a greater degree with experimental observation. Such models are generally based on hypotheses which relax the stringent requirement of strict serial information processing while retaining the idea of limited capacity. Kerr (1973) identifies two such alternative hypotheses: that of Undifferentiated Capacity and that of Limited Capacity Central Mechanism.

The Undifferentiated Capacity hypothesis has been used by Kahneman (1973) as a basis for a study of attention and effort. Processing capacity is not fixed, according to Kahneman, but determined by variables such as level of arousal, and it may be allocated to tasks in parallel up to the maximum level available at that time. Kahneman distinguishes two kinds of interference, manifested as a deterioration or decrement in performance, between concurrent tasks. Structural interference is caused by the inherent structural demands of the tasks: for example, it would occur when tasks require motor responses that are incompatible to some degree. Capacity interference depends on the total processing load imposed by the tasks and is not related to their nature: rather, it is a consequence of attempting more than one task at one time. Kahneman argues that even such disparate tasks as walking and mental arithmetic can be observed to interfere.

In the Limited Capacity Central Mechanism hypothesis, processing capacity within the central mechanism is allocated according to the specific processing demands of the tasks. Not all

mental operations require the mechanism. Keele (1973) suggests that operations such as rehearsal, response initiation, and movement correction do require the mechanism, whereas memory retrieval, for example, does not. Only if the operations related to two tasks simultaneously require capacity in the mechanism will interference occur.

Although the various processing models display different features, all have in common the idea that there is some definite limit to available processing capacity (although this limit may vary from time to time). It is this feature that is central to the use of secondary task techniques.

#### 5.4 Secondary Tasks and Limited Capacity

The capacity model as it relates to secondary task techniques is illustrated in Figure 5.1. We consider a processing mechanism with a limited capacity  $C$  (Figure 5.1a). A primary task requiring a capacity  $P$  is now imposed on the mechanism. Provided the available capacity is not exceeded (that is, provided  $P$  is less than  $C$ ) the primary task will be performed well (Figure 5.1b). The difference between  $C$  and  $P$  is known as the reserve capacity  $R$  (sometimes called the spare capacity). As long as  $P$  is less than  $C$  there will be little relationship between the processing capacity required by the task and its standard of performance, and it would not be possible to distinguish between the task requiring capacity  $P$  and another requiring some different capacity  $Q$  (Figure 5.1c) in terms of performance. Both tasks lie well within the capacity of the mechanism.

Now consider a secondary task requiring capacity  $S$  (Figure 5.1d),

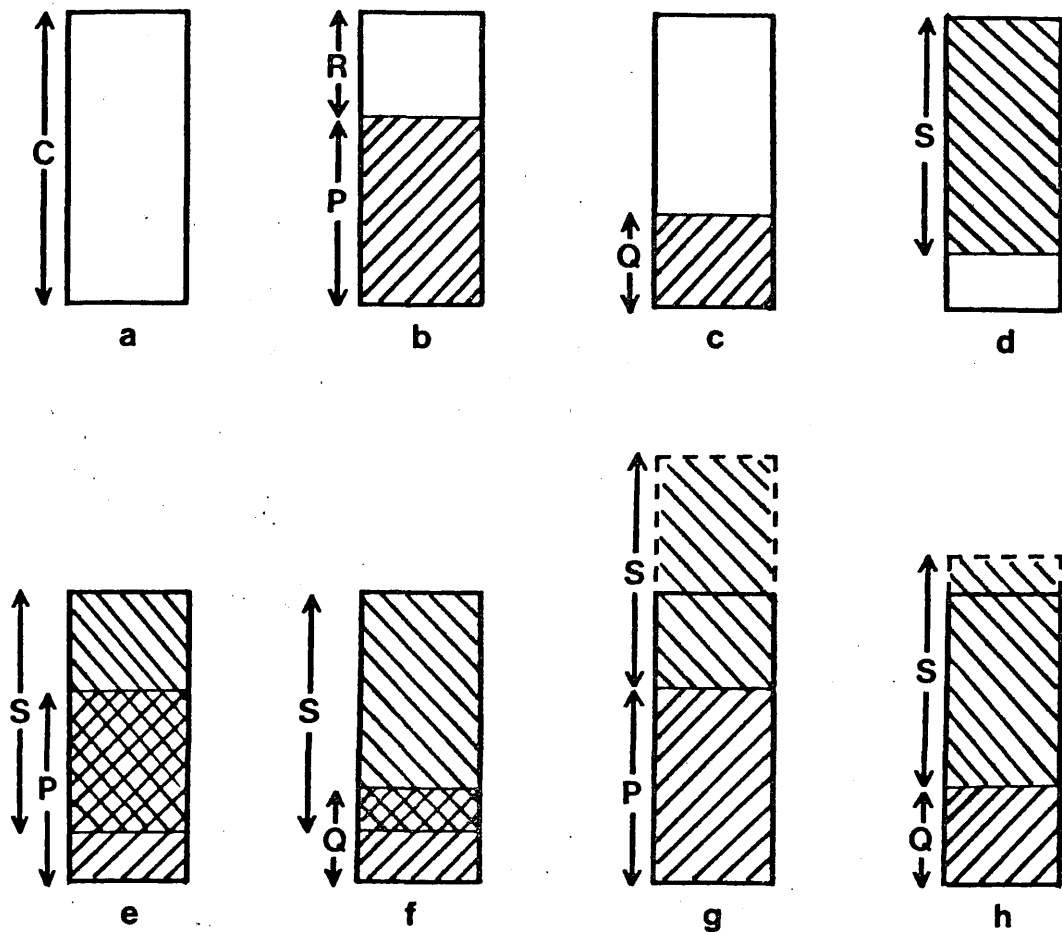


Figure 5.1. The capacity model and secondary task methods.

(a) Total capacity  $C$ . (b) Primary task requiring capacity  $P$ , leaving reserve capacity  $R$ . (c) Primary task requiring capacity  $Q$ . (d) Secondary task requiring capacity  $S$ . (e) (f) Secondary task used as a loading task. The amount of interference, shown by the overlap of the primary and secondary tasks, reflects the processing requirements of the primary tasks. (g) (h) Secondary task used as a subsidiary measuring task. The error rate on the secondary task, shown as that portion of the secondary task which "does not fit" within the limited capacity  $C$ , reflects the processing requirements of the primary tasks.

to be performed concurrently with the primary task. The secondary task should be chosen so that the combined loads  $P + S$  and  $Q + S$  are both greater than  $C$ . This secondary task may be combined with the primary task in two different ways.

We first consider the case in which the secondary task is to be used as a loading task. Its purpose will be to supply an additional processing load, effectively reducing the capacity available for the primary task. Interference between the tasks will show itself as a decrease in the standard of performance of the primary task (and possibly the secondary task as well); more interference, and hence a greater decrease, would be expected with the more demanding primary task (Figure 5.1e) than with the less demanding one (Figure 5.1f).

Alternatively, the secondary task may be used as a subsidiary measuring task. To use the secondary task in this way, the experiment is constructed so that the subject assigns priority to the primary task (usually he is instructed to do so), and he should, therefore, perform it as well as he can at all times. The secondary task is to be performed only when it can be done without detriment to the primary task. In this case the standard of performance of the secondary task gives a measure of the primary task processing load; the poorer the secondary task standard of performance, the higher the primary task processing load is deduced to be (Figure 5.1g, 5.1h).

It should be pointed out that in either mode of application it is not certain, nor is it assumed, that there is a precise quantitative relationship between the performance variable being measured and the primary task processing load. The techniques do have qualitative value, however, and would seem to permit ordering and comparison of



tasks in terms of the processing capacity they require.

### 5.5 Loading Task Experiments

Loading task experiments are constructed so that when the secondary task is applied, the effects of processing capacity being exceeded are directed into the primary task. The effects may, in addition, also be directed into the secondary task. Brown (1964) calls those tasks in which deterioration in standard of performance is measured in both tasks dual purpose second task experiments; I find it adequate for present purposes to class them as simple loading task experiments. In most loading task experiments, in any case, the performance standard of the secondary task drops unavoidably when performed concurrently with the primary task.

Brown (1964) emphasises that loading tasks can be used for "pin-pointing" faults in design and effects of stress on specific systems" (p.49). In this context experiments can be conducted to provide guidelines for redesign of man-machine systems to increase the effectiveness of the human operator. Garvey and Taylor (1959) report an example of this: using two types of loading task, they showed that the deterioration in the performance of a simple one-dimensional tracking task under load depended on the nature of the control system. Similar instances of loading tasks used in the evaluation of equipment are cited in a review by Rolfe (1971), which, together with reviews by Knowles (1963) and Brown (1964), should be consulted for more detailed discussions of this kind of application. In the remainder of this section, I shall be concerned with the use of loading tasks to investigate internal mental processing mechanisms.

## Memory

The study of various aspects of short-term memory has been approached using loading task techniques. Mitsuda (1968) showed that recall of pairs of single digits presented auditorily was adversely affected when pairs of two-digit numbers, presented alternately with the original pairs, had to be either written down or classified according to a simple scheme. Both the writing and classification tasks interfered with memory rehearsal to a similar degree; the tasks may be deduced to have been competing with rehearsal for the use of processing capacity. McNicol (1971) found that recall of sequences of binary (0 or 1) digits, in which the sequences could be subdivided into several groups of one repeated digit (for example, 00011011 could be divided into four groups), was decreased when subjects were required to repeat each digit as it was presented (shadowing). This was attributed to the shadowing process interfering with the recoding of the repeated items into one item, which is considered to require some central processing.

Kelly and Martin (1974), using both a simple arithmetical secondary task and a reaction time secondary task, demonstrated that retaining visual information in memory appeared to require processing capacity. Subjects were shown a sequence of two-dimensional shapes, and were required to identify certain shapes which had been shown to them previously. It was found that although the complexity of the shapes to be memorized had no significant effect on identification, shapes to which verbal labels could be easily attached significantly improved performance, suggesting that less processing capacity is required to rehearse the identity of such shapes.

Doost and Turvey (1971) presented subjects with iconic (visual

short-term sensory storage) memory tasks concurrently with secondary tasks in which either alphanumeric characters had to be repeated aloud and committed to memory, or letters had to be classified as vowels or consonants. Neither primary nor secondary tasks suffered significantly, from which it was concluded that iconic memory was relatively independent of the central processing mechanism. However, Chow and Murdock (1975), using a similar iconic memory task in combination with two types of secondary tasks, did detect some significant detriment to iconic memory, and hence drew the opposite conclusion. As in most types of experimental work, the sensitivity of the experiment is a crucial factor in the interpretation of results.

In an experiment in which subjects had to identify, after delays of up to 30 seconds, the locus of tactile stimulation on the upper side of the arm, Sullivan and Turvey (1972) found some grounds for suspecting that tactile short-term memory requires processing capacity. They suggest, however, that their results may be due not only to a reduction in available capacity because of the loading task, but also to the subjects responding in what they unconsciously felt to be a manner acceptable to the experimenters. Such involuntary alteration of behaviour to "please" experimenters is, by its nature, difficult to isolate and identify (Silverman, 1977).

Using a secondary task in which subjects had to identify the one repeated letter in a set of ten, presented over 50 seconds, Broadbent and Heron (1962) showed that the performance of tasks involving even a slight memory load is susceptible to distraction, as opposed to the performance of similar tasks with no memory load. The main purpose of this experiment was to investigate the effects of age on memory; it was shown that, in general, the performance of older subjects was very

much worse than that of younger subjects on at least one of the primary and secondary tasks, and the older subjects displayed far more variability in their standard of performance than the younger subjects.

The general feature of the memory experiments that have been described is that their main concern was to determine the degree of interference the secondary task causes in memory. When interference was detected, this suggested that the secondary task was competing with memory for processing capacity. Such interference had been found both in short-term memory and short-term sensory storage experiments.

The mechanism of short-term memory is regarded as rehearsal: to retain information a subject must constantly process it and return it to short-term memory store. It seems reasonable that this process should require central processing capacity. Short-term sensory storage, however, is traditionally believed by psychologists to take place peripherally to any central mechanism: consequently, to find that it apparently requires processing capacity is surprising. It is possible that in many memory experiments what is being measured is secondary task interference not only with the memory component of the primary task, but also with encoding and other mental operations associated with the primary task, although experimenters take elaborate precautions to avoid it.

The processes involved with the various types of memory (short-term sensory storage, short-term memory, and long-term memory) are discussed by Keele (1973).

## Learning

Loading tasks have been used to investigate human learning processes. Baker et al (1951) observed the development in the learning of a complex motor skill when a secondary task was assigned to different groups of subjects at varying stages of learning. For all groups there was an initial phase of heavy interference with the primary task when the secondary task was introduced. Following this phase, subjects' performance on the primary task continued to improve at much the same rate regardless of the stage at which the introduction of the secondary task had taken place. The secondary task lowered the absolute standard of performance eventually attainable on the primary task, probably due to structural interference, but there was little observable effect on the rate of learning the primary task. However, this is not always the case: Noble et al (1967) found that a loading task which involved anticipation of the next number in a sequence with varying degrees of randomness inhibited the learning of a tracking task as well as its performance.

Bahrick et al (1954) have observed that there is a "widely held notion that with continued practice tasks become less susceptible to interference" (p. 298). By running a secondary task at three stages during training in a primary task involving motor responses to a series of stimuli, Bahrick and Shelley (1958) showed that as training continued interference effects tended to vary inversely with the degree of redundancy (the extent to which one stimulus can be used to predict the next) in the stimulus series. For highly redundant stimuli a high degree of automatisation was possible, control of primary responses as training progressed changing from extroceptive to proprioceptive. The more automatised a task becomes, the less

processing capacity is required, and interference effects will accordingly lessen. Thus, the degree of automatisisation that is possible determines how effectively practice of a task can minimise interference with it. This effect applied as much to the secondary task as the primary task: Broadbent (1956) has demonstrated that practising a variety of loading tasks involving simple responses to an auditory signal will lessen their interfering effect on a primary perception task. In cases where learning is not the phenomenon being studied, precautions should therefore be taken to minimise the effects of progressive automatisisation of secondary task responses.

Stager and Muter (1971) investigated the effectiveness of different kinds of training for a primary decision task representing the interpretation of air traffic control radar displays. Subjects had to detect potential collisions in a series of strategic visual displays; as a loading task, they had to detect odd-even-odd sequences in random digits presented auditorily. As previous work had suggested (Stager, 1970), subjects who had been trained with instruction techniques involving pictorial displays processed information more effectively than those trained with apparently simpler verbal instructions. This was attributed to the differences in the processing procedures they acquired in training. Stager and Zufelt (1972) went on to confirm that method of instruction can substantially affect the required processing load; differences in processing procedures lead to differences in the load associated with a given information processing task.

#### Interference Effects

Schouten et al (1962) conducted an experiment to detect the

mutual impairment of two tasks due to limitations of perceptual capacity, running a variety of different tasks concurrently with a simple two-choice stimulus-response secondary task. For high standards of performance on the loading task, performance on the other tasks became very poor indeed: maze-tracking relapsed from adult standard to that of an eight-year-old child, and handwriting on a freely chosen subject deteriorated markedly both in technique and content, indicating interference with both motor control and reasoning. Kalsbeek (1971) comments that "it is amazing that the complete disintegration of both can be brought about experimentally simply by increasing the rate of binary choices in the distraction task" (p.109). In an experiment in which both the primary and the secondary task involved dual-choice reaction to stimuli, Kalsbeek and Sykes (1967) observed that performances on the tasks were inversely related. If a reduction was recorded in the degree of interference in one task, a corresponding increase would be recorded in the other, indicating different allocations of limited processing capacity to the two tasks.

Trumbo and Noble (1972) demonstrated that the greater the response uncertainty in the loading task (they used a response selection loading task with the number of possible stimuli ranging from one to five) the greater the interference in the primary task. The experiment indicated that the mental operation of response selection required processing capacity in proportion to the number of available alternatives. White et al (1975) showed that inserting a simple secondary task involving counting backwards by threes in the intervals between successive trials of a concept identification task interfered with the task in cases when theory suggested that a large number of hypotheses were present for evaluation by the subject. They did not attempt to examine

which aspects of hypothesis evaluation were susceptible to the secondary task interference.

## 5.6 Subsidiary Measuring Task Experiments

Subsidiary measuring task experiments are those in which a secondary task is employed as a measure of the processing capacity which is not required by the concurrent primary task. Errors due to processing overload should occur only in the secondary task; the primary task should be performed without detriment at all times. Brown (1964) claimed that, in comparison with the loading task technique, the subsidiary task technique appeared "to provide the more powerful tool in the evaluation of perceptual load and reserve capacity" (p.49), but he also noted that it was the "more controversial method" of the two (p.45).

The subsidiary task technique appears to offer a means of measuring the processing load associated with a task while keeping the disturbance due to the measurement at a minimum. The important point is that the primary task should be performed normally and should not be affected by the presence of a secondary task. Since measurement is occurring during times when the primary task does not present a processing load so great that breakdown of the task begins to take place, use of a subsidiary task should allow a better estimate of the processing load associated with a task under "normal" conditions.

### Vehicle Driving

One area where subsidiary tasks have been used effectively is in the measurement of processing load associated with vehicle driving.



In such experiments the methodological constraint that the primary task be performed without detriment at all times is imposed also by safety considerations, and clearly it would be unwise to attempt to carry out a loading task experiment on driving.

Brown and Poulton (1961) showed that error rates in both simple digit checking and mental addition subsidiary tasks increased when car drivers were required to drive in "shopping" as opposed to "residential" city areas, and they concluded that this was indicative of the different mental loads imposed by the driving conditions. Various subsidiary tasks were used by Quenault (1968) in an attempt to detect differences in performance between a group of drivers convicted of careless driving and a group of drivers chosen at random. No significant difference between the two groups was found, although effects due to the driving environment, similar to those detected by Brown and Poulton, were observed.

In studying fatigue effects on driving, Brown (1962a, 1962b) used both attention and memory subsidiary tasks to measure the reserve capacity of two groups of patrol policemen, one finishing an eight-hour spell of duty in the afternoon, and the other starting a spell of duty at the same time. Surprisingly, performance was consistently better for those finishing duty than for those starting. Attempting to account for this, Brown noted that those subjects who were starting duty had been awake for much the same length of time as those finishing, and had filled their day with various activities which might have had more tiring effects than driving. Another factor might be that policemen finishing duty might be more willing to co-operate fully with the experimenters than those starting.

Using a subsidiary task in which subjects were required to call out at a regular rate, the degree of regularity being a measure of the spare capacity available, Brown et al (1967) showed that continuous city driving for long periods does not necessarily affect driving performance adversely. In the same experiment, performance on a subsidiary vigilance task, involving detection of light signals through rear-view mirrors, was observed to improve over the long periods. This was attributed to the automatising of driving, leaving more capacity for perceptual skills. It seems possible that automatising of this type could also at least partially account for Brown's (1962a, 1962b) results.

Brown (1966) showed, using a digit checking subsidiary task, that in a class of trainee bus drivers, those who were ultimately successful in their five-week course had significantly greater reserve capacity while driving at the seventh day of training than their unsuccessful colleagues. At that early stage these two groups of trainees could not be distinguished by direct measures of driving performance. However, at the end of the course there was no longer a significant difference between the reserve capacities of the two groups. Similar follow-up work (Brown, 1968) confirmed that a subsidiary task method was a sensitive means of identifying drivers likely to be suitable for continued training.

### Mental Processes

Possibly the most ambitious use of subsidiary task techniques has been in the investigation of mental processes and operations. Using a subsidiary task involving sorting nuts and bolts by touch only, Dimond and Beaumont (1971) carried out an investigation of the

perceptual analysing systems of the individual cerebral hemispheres of the brain. Subjects were required to report on random pairs of digits perceived visually by individual hemispheres, a system of screens being used to direct information to the hemispheres independently. The left hemisphere, which controls speech, performed faster; to report digits perceived by the right hemisphere, it was apparently necessary to transmit the information to the left, causing delay. However, no significant differences in performance on the sorting task were observed for these two conditions, suggesting that similar capacity is required by the primary task in both cases.

Several experiments have attempted to subdivide mental processes into component operations and analyse the capacity required for each operation by the use of subsidiary tasks. Johnson et al (1970) attempted to split a verbal task into encoding, retention and recall components and they suggested tentatively that encoding requires more processing than retention, but less than recall. Martin (1970), performing a similar experiment to study organisational processes in memory, also found that retrieval required more processing than encoding, possibly because the organisation of memory was occurring at the retrieval stage. In an investigation into how the type of information affects the load involved with processing it, Martin et al (1973) showed that memory rehearsal required less capacity for lists of words which could be easily organised into categories compared with lists of words which could not.

The results of such experiments as those above, together with the loading task experiments that have been carried out on memory, are undoubtedly important. Mental processes are extremely difficult to investigate, and to bring "hidden" aspects of them into the open, as

these experiments appear to have done, is a valuable contribution to their understanding. All such experiments must be regarded carefully and critically, however. The elaborate precautions taken to ensure that mental tasks are correctly subdivided into their components, and the possibilities of structural interference effects, all permit more than one interpretation of results.

### Other Examples

Subsidiary tasks have also been used in a variety of other experimental contexts. Zeitlin and Finkelman (1975) presented subjects with a tracking task in which two different types of tracking control were used. They showed that a random digit generation subsidiary task was not sufficiently sensitive to pick up any load differences between the two conditions, but a delayed digit recall task indicated that a velocity control required more capacity than a position control. Another common type of experimental task, the vigilance task, is exemplified by Poulton (1958), who required two groups of subjects to monitor two and six dials respectively. Although performance between the two groups was comparable, a subsidiary memory task indicated clearly that the six-dial task was more difficult, requiring more capacity than the two-dial task.

The effects of noise on spare capacity have been investigated using subsidiary measuring tasks. Boggs and Simon (1968) showed that noise (recorded half-second bursts of a bandsaw cutting aluminium) had no significant effect on a reaction time primary task, but caused a subsidiary digit-checking task to be performed less accurately. Differences in subsidiary performance due to alterations in the complexity of the reaction time task were exacerbated by the noise,

which, Boggs and Simon suggested, "used up" spare capacity. Finkelman and Glass (1970), using both predictable and unpredictable bursts of white noise, found broadly similar results, and concluded that the "effort involved in noise habituation contributes to the loading of the organism's channel capacity" (p.213).

Brown (1964) emphasises that subsidiary tasks can measure the efficiency of information transmission in a task. The more reserve capacity available for a given task, the more efficient are the mental processing "programs" dealing with the task. Learning and automatising of tasks may be regarded as processes of developing and streamlining the efficiency of these programs. Bahrick et al (1954), in an experiment concerning the learning of a motor task, found that the scores on repetitive and random versions of the task were comparable. However, a subsidiary mental arithmetic task showed that the repetitive version, after a period of learning, required less capacity and was therefore being dealt with more efficiently than the random version, for which no learning effect could be detected. Knowles and Rose (1963), in a two-man lunar landing simulation task, showed using a subsidiary task that the degree of mental efficiency with which the operation could be dealt was dependent on the initial conditions of the landing, and that, most importantly, both crew members were dangerously close to overload during the final few seconds of the landing, an effect which practice could not remove.

It is clear that it is important for the success of a subsidiary task experiment that subjects do not permit the subsidiary task to affect their performance on the primary task. Any such effects should be below the level of significance if the experimental results are to

be acceptable. Bahrick et al (1952), using monetary incentives, concluded that subjects tend to give priority to the task they perceive as most important, and try to minimise interference due to other tasks. Thus it is necessary for the experimenter to ensure that the subjects consider the primary task to be the most important. In an experiment in which the primary task involved answering questions about a visual display, Broadbent (1956) found that his subjects consistently regarded the secondary task, involving various simple reactions to a buzzer, as the more important, and the experiment was therefore effectively a loading task experiment. Rolfe (1971) has argued that secondary task interference seriously weakens the applicability of subsidiary measuring tasks, pointing out that "early studies in series of experiments have often claimed to be successful in this respect, but as the measuring techniques have been improved so an impairment in primary task performance has become apparent" (p.145).

## 5.7 Discussion

Secondary task techniques have been employed with some success to investigate the degree of processing capacity required in the performance of various tasks. Following the categorisation of the bulk of such work by Brown (1964) into loading task experiments and subsidiary measuring task experiments, it can be seen that these two techniques each emphasise a different aspect of human information processing. Loading task experiments tend to show the effectiveness with which a task may be performed under load, highlighting points at which performance may break down. Subsidiary task experiments, by measuring reserve capacity, indicate, for a given level of effectiveness (usually optimal), the efficiency with which the appropriate processing operations are carried out.

The theoretical background to many loading task experiments is intuitively reasonable. The loading caused by the secondary task can often be readily regarded as effectively equivalent to real-life loading by a variety of attention-demanding stimuli, and its effects seem likely to be the same operationally. Certainly, loading task investigations of learning and man-machine systems would seem to lead to straightforward, clear results (Rolfe, 1971). Regarding the loading task experiments involving various types of memory, however, theory is less straightforward and the conclusions are consequently weakened to some extent. The rationale behind such work is that if the loading task can be shown to interfere with the primary memory task, then this is an indication that the memory task requires central processing capacity for which the loading task is competing. Extremely careful experimental design is required before the experimenter may be reasonably sure that the primary task is solely a memory task. In work involving iconic memory (Chow & Murdock, 1975), for example, the experiment must be constructed to minimise the encoding of visually presented information, since encoding is a different process.

The problem is just as serious with subsidiary measuring task experiments involving various components of mental processing. Methods of subdividing processing into non-overlapping components are often elaborate, and the results, although valuable, need to be treated critically. Kerr (1973) identifies five categories of mental operation (encoding, multiple input, rehearsal, transformation and responding). In reviewing secondary task studies of these individual operations, she reports consistent trends in results which indicate that such techniques are illuminating otherwise inaccessible structures of mental processing. However, the problems of the interpretation of results are considerable. The apparent lack of processing demand during

encoding, for example, may be due to insufficiently sensitive experiments rather than the nature of encoding itself.

Subsidiary measuring task experiments tend to be less well backed up by theory than those involving loading tasks. In essence, the interpretation of subsidiary task experiments requires a more sophisticated model of mental processing capacity. Although, when simply stated, the idea of directly measuring reserve capacity by means of a secondary task is elegant and looks promising, there is no general agreement among psychologists concerning the exact nature of mental processing and its capacity limits, and this has consequences for the interpretation of at least the more ambitious subsidiary measuring task experiments. A fundamental problem lies in deciding whether processing is totally serial in nature, or if parallel processing is possible and, indeed, the norm. Kahneman (1973), relating degree of attention to processing effort, considers that attention is certainly limited, but that the limit is variable, and the allocation of attention is divisible: that is, parallel processing indeed occurs. Attention may be regarded as being allocated in units, and concurrent tasks will interfere when there are not sufficient units to be allocated to both tasks. Then, depending on allocation policy, which may or may not involve a conscious decision, one or other or both tasks will suffer.

A thorough investigation of the microstructure of the interaction of two concurrent tasks by Fisher (1975a, 1975b, 1977) has tended to support the concept of a limited capacity central processing mechanism, but her results indicate that what in fact takes place is processing in sequence with rapid switching between the two tasks rather than true parallel processing. The nature of the switching process, however, is



not clear. Fisher emphasises that her results do not invalidate Kahneman's model of parallel processing, and show merely that "in some situations involving the combination of quite difficult tasks, a sequential strategy would be the best operating procedure" (Fisher, 1975a, p.288).

The difference between sequential and parallel processing may have important consequences for the interpretation of the more sophisticated secondary task experiments. In the parallel model, reserve capacity would be a measure of how much processing space is not actually utilised during the continuous performance of a task. In a sequential model, reserve capacity would quite simply be a measure of the processing time not allocated to the primary task. It is not clear that in all cases, particularly those involving the study of the components of mental operations, these would be operationally equivalent.

Kahneman (1973) draws attention to structural interference, which causes the processing demand of two tasks performed jointly to be greater than the sum of the individual demands of the tasks performed separately. This will occur when selected performance modes are incompatible or antagonistic or when both tasks require the same perceptual or response mechanism. Experiments should be constructed to avoid such interference as much as possible. Stimuli for different tasks should be presented via different modes (for example, visual and auditory) and responses should be similarly separated. Even so, it is doubtful that structural interference can ever be entirely eliminated in an experiment.

Rolfe (1971) has criticised the secondary task technique and its

exponents on three main points. Firstly, he suggests that often "experiments have been degraded by poor experimental procedures" (p.144). In particular he notes that detailed analysis of secondary task performance and its effect on primary task performance is often lacking. This is undoubtedly true of earlier, less sophisticated, work, but, in general, recent experimenters have considered these aspects in more detail. Secondly, Rolfe notes that there have been few attempts to "validate the technique and its findings outside the laboratory". I would agree that in cases where the secondary task "acts as a substitute for the additional load encountered or anticipated in the actual situation" (p.145), there is a need for "follow-up trials to check that the laboratory results are correct". However, this criticism seems inappropriate to the many recent experiments investigating mental operations.

Rolfe's third point is that in most subsidiary measuring task experiments the subsidiary task does seriously interfere with the primary task. Interference may be structural, or caused simply by the stress of introducing a second task. A similar point is raised by Brown (1978), following experimental work by Neisser and Becklen (1975). Brown states that "work-load may be largely a function of the characteristics of a task, rather than of the informational load imposed by its component parts" (p.??). The inherent structure of tasks "does not permit a reliable sharing of attention between inputs". This can be regarded as a form of structural interference. In general, then, most critics agree that some non-trivial form of interference in a subsidiary task experiment is unavoidable. Does this invalidate the subsidiary task method?

I would argue that it does not, although often it must weaken experimental conclusions. For some purposes subsidiary tasks would

still seem to be appropriate. For example, if the loads of two similar tasks are to be compared, careful construction of experiments permits subsidiary measuring tasks to be used with a reasonable degree of confidence in their validity.

Some work in recent years has suggested that the capacity model may require some further sophistication. Stager and Laabs (1977), in work concerned with performance on three simultaneous tasks, obtained results which led them to doubt the validity of reserve capacity as a direct indicator of processing load. There is not yet sufficient evidence, however, to recommend discarding all secondary task techniques on these grounds. It would, on the other hand, be extremely unwise to regard the techniques as proven beyond all reasonable doubt, or to approach each application with anything other than a critical frame of mind.

### 5.3 Extension of Secondary Task Application

In looking at the overall picture of secondary task experiments, it becomes apparent that all the primary tasks that have been studied can be categorised as tasks in which, to a large extent, response models appropriate to the task decisions are already available to the subject. In other words, given a set of input stimuli, the subject already possesses the essentials of a suitable response because he already possesses a mental model of the task which can generate such a response. To identify the response he need only consult the model. Little, if any, model construction or development is involved in the task process.

This survey has considered some experiments in which, clearly, some development of mental models as time progressed must have occurred.

Any experiment which displays a learning effect is an instance of this. Automatisation such as that reported by Bahrck and Shelley (1958) indicates that the mental models of their subjects were becoming more efficient as the redundancy in the task stimuli were incorporated in them. In general, too, it seems likely that the operation of a model becomes more efficient with continued use: its structure becomes more fixed and its manipulation more well-learned. As the efficiency increases, less processing capacity is required - for reasonably simple tasks, therefore, performance can be seen to become almost entirely automatic.

Yet although gradual model development over repeated trials of a primary task can often be detected, no primary tasks have been investigated in which the construction of a model is inherent in each individual trial. Such tasks would be problem-solving tasks, where, due to the complexity of the problems, straightforward mental models which provided appropriate immediate responses to inputs could not be available to the subject. An example of such a problem-solving task which might be investigated experimentally would be the planning of a move in a complex game such as chess or, indeed, the Organisational Control Game described in earlier chapters. The player would presumably be familiar with the rules of the game, but a solution would not involve simple reference to a prepared model. Rather, the situation would have to be assimilated from "first principles", and a model appropriate to that particular situation constructed. The investigation of such a problem-solving task by a secondary task technique, in particular a subsidiary measuring task technique, might give valuable insight into the processing load presented to a subject by the task.

Many decision-making or problem-solving processes can be regarded

as aggregates of simple mental processes, just as simple operations in a computer program can be combined to solve extremely complex problems, but this kind of reductionist approach to assessing processing load does not seem fruitful. If we were interested in assessing the processing load the problems might present to a non-intelligent machine this would undoubtedly be the way to proceed. In complex problems, though, it is doubtful that a person behaves as a machine: data and information may be missed, ignored, or treated preferentially; short cuts may be identified and taken; innovation and risk may be employed to a greater or lesser extent. An analysis of component operations would be unlikely to provide more than a rough approximation of the processing load imposed by the problem. A subsidiary task experiment, by measuring the reserve capacity of the individual throughout the problem solving process, has far greater potential for assessing the processing load.

What type of subsidiary measuring task would be most suitable for such a purpose? A problem-solving primary task would be carried out over a period of minutes rather than seconds, and it seems that a secondary task of a similar, continuous, nature might be most useful. In particular, it might be desirable to avoid the sort of task which requires nothing other than an almost immediate response to a stimulus. For such a task, a suitable strategy could be to switch from primary to secondary when the stimulus occurs, and switch back to the primary after the response has been made. The effect of this serial strategy on the performance of the primary task would probably be detrimental. The secondary task would be dependent on two decisions by the subject: a decision about whether to switch from the primary to the secondary task and a decision about the time to be spent on the secondary task before returning to the primary.

For this reason I would prefer a subsidiary task which is in some sense continuous, so that parallel processing is necessary if both tasks are to be attempted. It is likely that, when secondary task responses are required, the same type of decisions as for the serial case will be involved, but the quality of responses should be related to the amount of reserve capacity available for parallel processing of the secondary task over the duration of the primary task interval.

It seems possible that a task which involves a high degree of memory rehearsal (carrying information from a previous stimulus to a subsequent one) is likely to encourage parallel processing, and thus tasks which are predominantly memory tasks involving minimal motor or transformation processing (which would encourage a serial strategy) are seen as being suited to investigation of lengthy problem-solving tasks. Such a task has been used by Foulton (1958): messages involving numbers are presented auditorily every five seconds, and after ten messages, subjects were asked which number had been mentioned most frequently. This task requires continuous storage of information, and we might expect the quality of storage to deteriorate with a reduction of reserve capacity. The low response rate (once a minute) decreases task switching, but also decreases the amount of data available for statistical analysis. Zeitlin and Finkelman (1975), with their delayed digit recall task, in which subjects had to recall the digit before last in a sequence of random digits presented every two seconds, have avoided this problem at the expense of comparatively short (four second) recall intervals.

Other approaches to the problem of finding a continuous subsidiary task are possible. The interval production test successfully used by

Michon (1966) as a subsidiary measuring task seems worth consideration, if as seems indicated there is a relationship between the results of the test and processing load (discussed by Vroon & Vroon, 1973). Sorting unseen nuts and bolts, used by Dimond and Beaumont (1971), also can be regarded as continuous under certain circumstances. Zeitlin and Finkelman (1975) have noted that the "integrative nature" of random digit generation which tends to be insensitive to "moments of difficulty, averaging them over the entire duration of the digit sequences" (p.220) may be of particular use "in evaluating the effect of general loading conditions or steady-state environmental stressors". The integrative nature of the task may make it useful as a means of assessing long-term problem-solving load.

It seems that with careful choice of tasks, the use of secondary task techniques to measure the load associated with complex problems is a distinct possibility. In the next chapter, I describe in detail an experiment, conducted at Royal Holloway College, in which a subsidiary measuring task was used successfully to compare the load imposed by a number of problem-solving primary tasks.

## CHAPTER 6

### A SECONDARY TASK EXPERIMENT TO ASSESS PROBLEM COMPLEXITY

#### 6.1 Introduction

In the previous chapter I concluded that a secondary task experiment might well be a suitable means of measuring the mental processing load associated with a complex problem-solving task. An experimental technique of this type could be of value in the practice of Operational Research. Situations are often found in Operational Research in which it would be useful to have some objective measure of the complexity of a problem facing a decision-maker.

It seems reasonable to define a decision-making load to be the mental processing load associated with a complex decision problem. Different problems might be expected to present different decision-making loads to a decision-maker, and it would be of interest to know how decision-making load varies according to the precise features of the problem. In particular, it might be useful to be able to identify those problems which present decision-makers with particularly high decision-making loads. In the case of large-scale decision-making systems, this would allow the decision points at which satisfactory decision-making is likely to break down to be located. These points are not necessarily easily identifiable, either by inspection or by theoretical analysis of the systems.

An objective method of assessing problem complexity might be extremely valuable in a variety of areas. In problems of medical diagnosis, for example, where diagnoses based on a limited quantity of data are often required in a short time, an objective technique



would permit investigation of the way in which ease of diagnosis varies according to the quantity and quality of the data, its nature, and the way in which it is presented. Similar considerations apply to military decision-making in battle, where once again time is likely to be a severe constraint. What types of decision tend to overload military decision-makers, causing their decision-making to deteriorate? Identification of the critical features of decisions might indicate guidelines for the design of suitable decision rules and decision-making systems which would avoid, to some extent, the worst effects of unexpected overload.

A technique for assessing problem complexity would also have potential application in operational and experimental gaming such as that described elsewhere in this thesis. Indeed, as noted in the previous chapter, our interest in the measurement of problem complexity arose early in the course of our gaming work, when we began to seek a means of measuring and comparing the decision-making loads that would be imposed by different scenarios in the Organisational Control Game. Subsequently, as will be explained in Chapters 7 and 8, the direction of the gaming work changed and this particular aspect was no longer of immediate interest to us. However, our general interest in the field of assessing problem complexity remained, and the work described in this chapter is the result.

This chapter describes and discusses a secondary task experiment which we conducted at Royal Holloway College. The aim of the experiment was to evaluate the potential of a secondary task technique as a means of assessing problem complexity. As will be seen, the results are encouraging. They have already been reported in the O.R. literature (Klein & Cooper, 1981).

In the following section, I shall briefly recount relevant features of secondary task methods encountered in the previous chapter. This is followed by a consideration of the nature of complex decision problems, explaining the thinking behind our choice of decision problem. I then describe in detail the experiment and its results, before finally discussing the work and our conclusions.

## 6.2 The Secondary Task Technique

As we have seen in Chapter 5, the secondary task is based on a model of mental processing as taking place within a system of limited capacity. According to the model, any mental load lower than the capacity of the system will undergo satisfactory processing and errors due to overload will not occur. However, if capacity is exceeded, then overload occurs, and mental processing consequently deteriorates; additional errors in performance will occur in this case.

### The Subsidiary Measuring Task Technique

For the present purposes, the subsidiary measuring task technique is considered to be more suitable than the loading task technique. Essentially this is because the effect that is actually measured in a loading task experiment is the degree of deterioration in performance of the primary task itself. Most complex problem-solving tasks do not lend themselves obviously to an ordinal scale of performance measurement. It is often impossible to state whether a given solution is "right" or "wrong", let alone whether one solution is better than another. In Brown's terms (Brown, 1964) we are not interested so much in the effectiveness of problem resolution - effectiveness may be impossible to define appropriately - but rather with the efficiency with which the process of problem resolution is carried out.

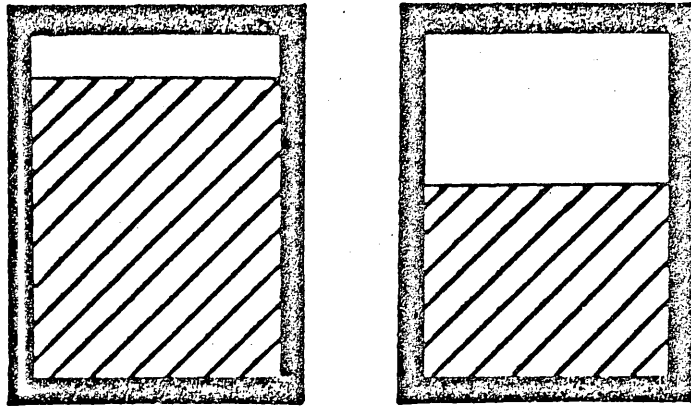
We shall briefly recapitulate the principles of the subsidiary measuring task technique. Consider two tasks, the loads of which are to be compared (Figure 6.1a). If neither task individually overloads the system, they will both be carried out properly, and no direct measure of load will be possible. We now add a secondary task of constant difficulty, which the subject is asked to perform at the same time as the main task (Figure 6.1b). This task is such that the combined loads of main and secondary tasks together exceed the capacity of the system. Now overload does occur, and processing errors are observed.

Subjects are asked always to give priority to the main, primary, task, so that this task is not affected by overload. The deterioration of performances is directed to the secondary task. The greater the deterioration of the secondary task, the greater must be the mental processing load presented by the main task.

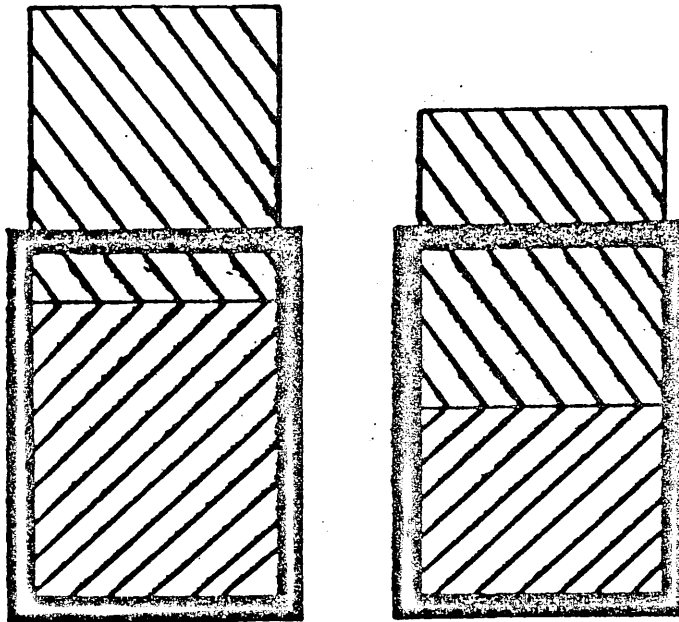
In the last chapter it was established that the most suitable types of secondary task for the present application would have a strong "continuous" element - that is, they would necessitate a significant degree of processing in parallel to the primary task, rather than intermittent switching from primary to secondary and back again. At the same time, a high number of secondary task responses is desirable, in order to maximise the quantity of data available for statistical analysis.

### 6.3 The Nature of Complex Decision Problems

The nature of the complex decision problems in which Operational Research scientists are likely to be interested differ in important



**a**



**b**

Figure 6.1. (a) The loads due to two primary tasks, neither of which overload the processing system. (b) The loads of the primary tasks, combined with a secondary task of constant load.

ways from the types of task that have been investigated by psychologists using secondary task techniques. In virtually all the studies reported in the psychological literature, the primary tasks can be categorised as short-term tasks in which complete or nearly complete solution models for decisions are already available to the subject (Figure 6.2a). In other words, for a given set of input stimuli, the subject already possesses the essentials of a suitable response. The response is implicit in his model. After perceiving the input stimuli, the subject consults his mental model of the situation, which he already has mentally available, and he generates the appropriate response. This process rarely takes more than a matter of seconds.

The types of mental process of interest here are different. Being those necessary to solve complicated problems, they are generally far more complex, and possibly less easily formulated. Straightforward solution methods will not be immediately available. A different process takes place to solve the problem (Figure 6.2b). Following the perception of the input stimuli, the subject has to construct a model of the situation. Only then can he use this model to generate a response. The important point is that the subject must construct his model as an integral part of the complex decision process. By constructing the model he is teaching himself how to solve the problem.

In this type of problem, there is unlikely to be a unique "best" solution. Indeed, it is often impossible to compare the quality of different solutions. Correctness may depend to a greater or lesser extent on subjective factors as well as on objective factors. In general, as already pointed out, solutions to complex decision problems can rarely be considered as "right" or "wrong", and the quality of the decision cannot be used as a guide to the load imposed by the problem, or its difficulty.

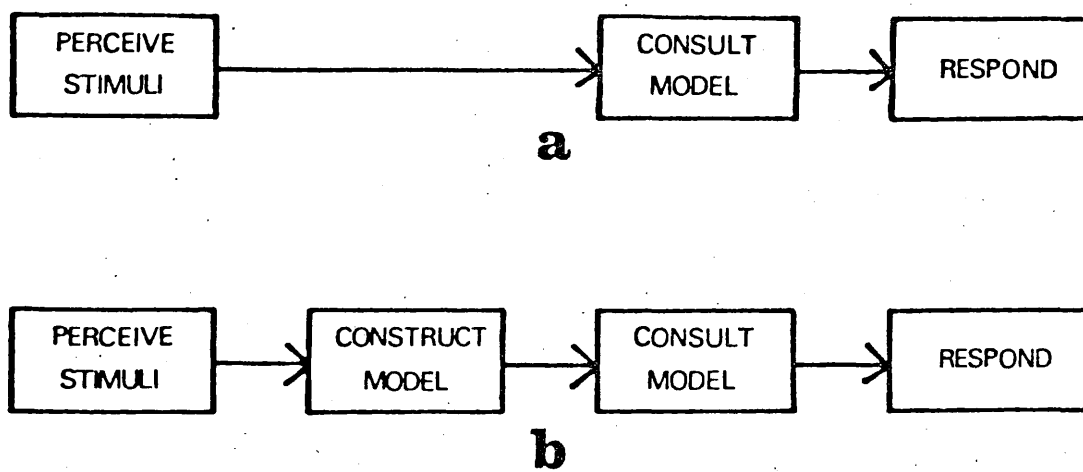


Figure 6.2. Solution processes for (a) simple and (b) complex problems.

It may be appropriate at this stage to mention a "resources" model of the decision-maker (Sharp & Dando, 1979). This model is, quite explicitly, a model of a decision-maker as a "limited capacity information processor" (p. 287) and resources are defined ("metaphorically") as the integral of capacity over time. Sharp and Dando suggest that resource use for a problem is proportional both to the complexity of approach to the problem and to the creativity required by the approach. This resource use is, of course, the mental processing load imposed by the problem over time, and, similarly, it can be seen that a subsidiary measuring task experiment will be using "spare resources" in order to measure this load.

#### 6.4 A Secondary Task Experiment

In this section I describe a subsidiary measuring task experiment to evaluate the application of the technique to the measurement of the mental processing load associated with complex problems. This description is somewhat more detailed than that which has already appeared in the literature (Klein & Cooper, 1981).

##### The Primary Task

A set of chess problems were used as the primary task. In comparison to many complex decision problems, a chess problem will clearly have a far more well-defined and formal structure. Nevertheless, provided the problems are of an appropriate level of difficulty, the solution process would seem to be of the type illustrated in Figure 6.2b. The chess-playing subject has a thorough grasp of the fundamentals of the game - that is, the rules, and certain other relatively simple concepts. However, these, in themselves, are insufficient to solve

the problems. For each problem, the situation on the chess board must be combined with the fundamentals to produce a model which is sufficiently sophisticated to generate a viable solution. In the case of chess, or in any complex decision environment, the construction and manipulation of an appropriate model is the main component of the problem-solving task, and is decidedly non-trivial. The outcome which the model generates is bound to be dependent on the way in which the model is put together.

We had originally intended to use chess problems of the classic "Meredith" form: "white to play, and mate in two moves", the subjects being asked to identify just the next move that white should make to proceed towards that mate. A collection of problems was selected from a published compilation (Barnes, 1976). It rapidly became clear, however, that these problems were very much beyond the ability of the subjects we were able to find (and also beyond the ability of the experimenters). Effectively the problems were all equally impossible: subjects would not even know how to go about constructing appropriate solution models. It was pointed out to us that as well as being difficult problems, this particular type of problem required skills not necessarily identical to those that might be developed by a player interested only in normal chess playing.

Accordingly, we devised our own chess problems. Two of these are reproduced here (Figures 6.3 and 6.4) in a similar form to the way in which they were presented to the subjects. Subjects were asked to select the move they would make next in order to win as quickly as possible, or, where appropriate, they were told that a mate in one move or two moves was possible, and asked to find it. To an expert chess player, most of the problems would be trivial. In many cases, they involved nothing more than identification of an immediate



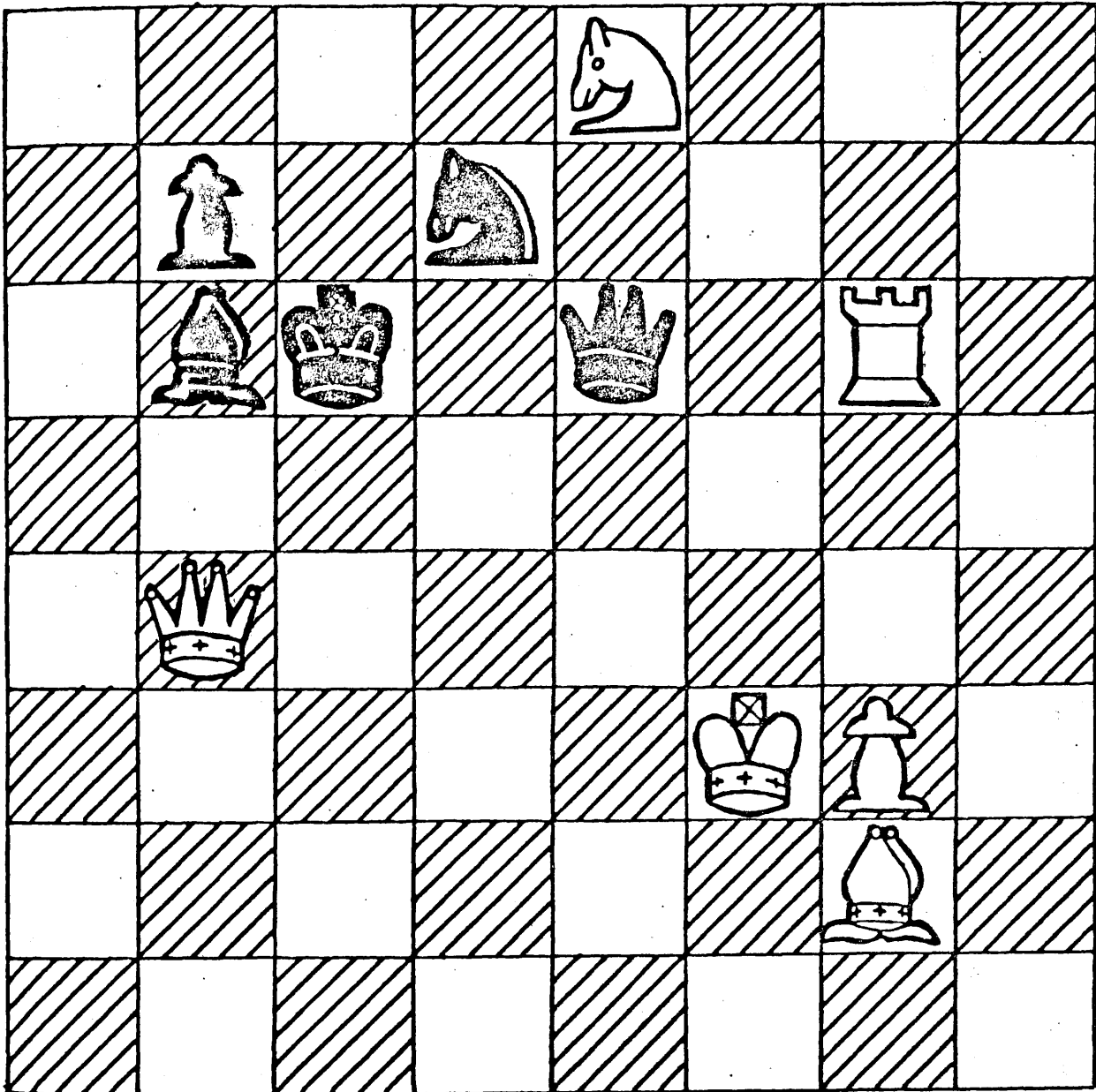


Figure 6.3. Chess problem 6 in the experimental set. White to move, playing up the board. K to E4 gives a mate.

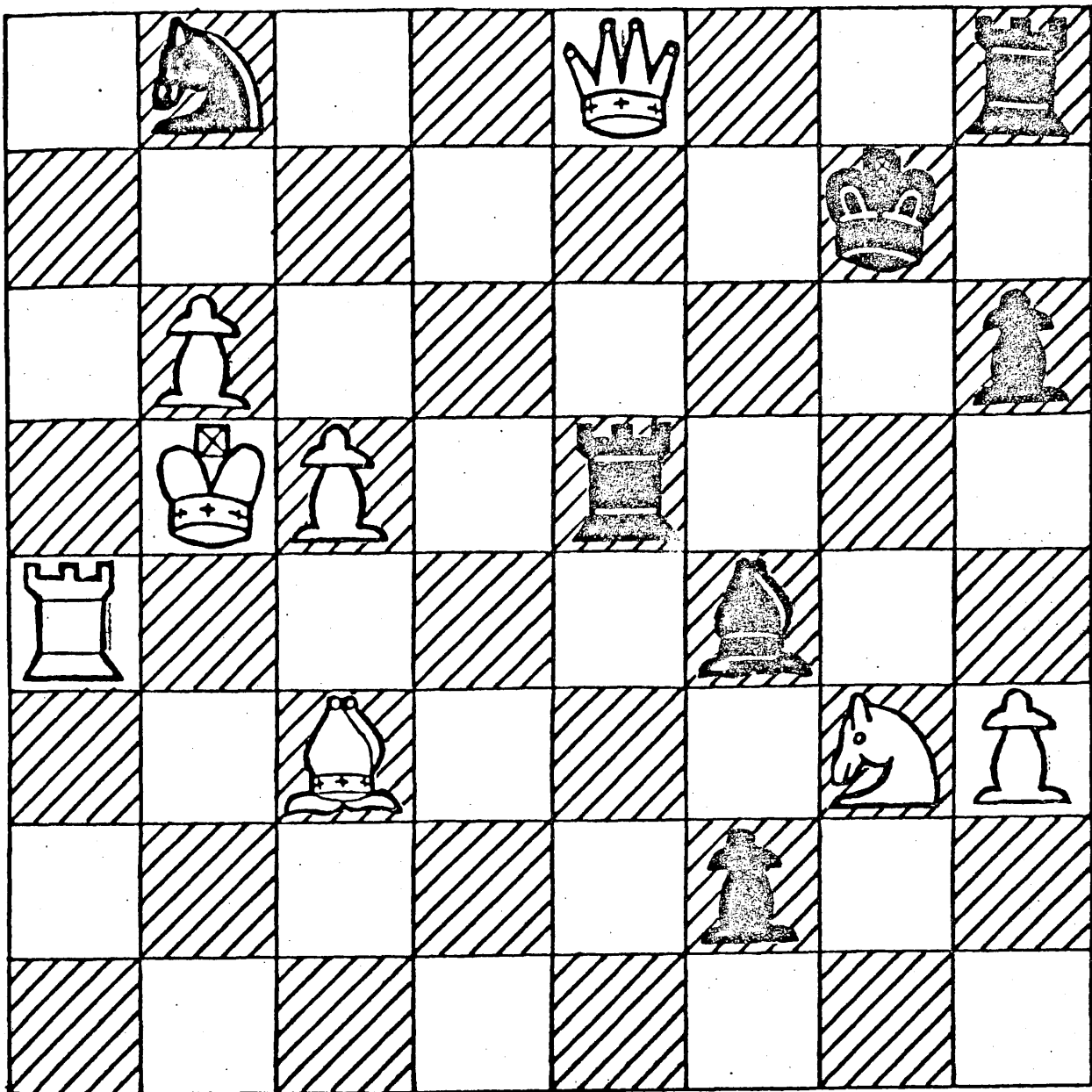


Figure 6.4. Chess problem 12 in the experimental set. White to move, playing up the board. N to R5 is the first move in a mate-in-two.

check-mate move. An expert player would be expected to have already available, stored as "fundamentals", models immediately applicable to these problems. Our subjects did not, as the results will show.

The subjects were six undergraduate students at Royal Holloway College. They were presented with a series of fourteen chess problems (Klein & Cooper, 1979) on separate sheets of paper in a predetermined order. Subjects were allowed two minutes to attempt each problem. At the end of each two minutes they were asked to write down their solution before proceeding to the next problem. The first five problems were "learning" problems, to familiarise subjects with the experimental format; only the results from the final nine problems were analysed.

#### The Secondary Task

The secondary task was a mental addition task involving short-term memory. A tape-recording of a female voice reciting a list of pseudo-random digits, from two to nine, had been prepared for playback during an experimental run. The digits were read out at five-second intervals. Subjects were asked to call out the sum of each pair of digits. For example, if the first two digits were "two" and "seven", the correct response would have been "nine", and if the next digit was "three", the next response ought to have been "ten".

The task requires that the digit prior to the last response be retained. The most common error subjects make is to add the latest digit to their own previous response, rather than the digit heard prior to it. Thus the task requires some conscious use of memory to be performed correctly. This, I believe, gives it a "continuous"

characteristic which I have already identified as being desirable. At the same time, a response rate of once every five seconds gives a high amount of data for analysis.

In addition to the digits, also recorded on the tape was a male voice giving all instructions necessary for the running of the experiment, such as instructions to write down an answer to a problem, or to start the next problem. Thus an experimental run was controlled entirely by the tape-recording, although an experimenter was always present to supervise the procedure.

An experimental run proceeded in the following manner. To begin with, the first chess problem was presented to the subjects without any digits. Then a two minute session of the arithmetic task alone was presented. For the second problem, the digits were available, but no arithmetic was required. For the rest of the problems, subjects were asked to perform the mental arithmetic while considering the problems. At the end of the experiment subjects were once again asked to perform two minutes of the arithmetic task alone. The number of mistakes and omissions in the arithmetic task was noted, and a total error score for each two minute sequence of digits was obtained.

### Results

The problems were ranked for each of the six subjects according to the number of errors occurring in the concurrent mental arithmetic. Assuming that the fewer errors that are made in the secondary task, then the less complex is the associated main problem, this gives a rank ordering of problem complexity for each subject. Agreement

between the rankings for individual subjects was found to be significant at the 2% level (Kendall concordance test,  $\chi^2 = 19.7$ ,  $\nu = 8$ ,  $p < 0.02$ , Kendall, 1975). This indicates that the secondary task-technique is measuring a property of the problems which is consistent from subject to subject.

The individual error rates were combined to produce a total ranking of the problems in order of increasing complexity. This ranking is shown in Table 6.1a. To produce this ranking, a subject's error rate for a given problem was expressed as the proportion of errors during that problem to his total errors over all the problems of interest. These proportions were then summed, for each problem, over all subjects. This procedure ensures that the proportional contribution from each subject to the overall result is the same.

As noted above, the individual rank orderings, showing a high degree of concordance, clearly indicate that the orderings correspond to a meaningful property of the problems. It has been assumed that this property is the complexity, or difficulty, of the problems. This seems a reasonable assumption, but undoubtedly an independent test of the meaning of the orderings would be desirable. In the case of these chess problems, such a test is possible. Most of the chess problems have one solution which is very clearly better than all other possible solutions (note that this property is not typical of complex problems in general, though I do not believe that this weakens the arguments presented here to any significant extent). In all, sixteen subjects have attempted the chess problems (as well as the six involved in this experiment, ten other subjects attempted the problems in the course of an experiment which will be briefly mentioned later). It seems reasonable to suggest that those problems which are correctly solved

Rank	1	2	3	4	5	6	7	8	9
(a) by secondary task	14	9	10	8	(6 13)	11	7	12	
(b) by solution rate	14	13	(8 10)	9	6	11	12	7	

Table 6.1. Problems ranked in order of increasing difficulty. Problem numbers in parentheses are of equal rank.

most often are the simplest, and present the least processing load, while those which are seldom solved are the most difficult, and present the greatest load.

The problems are ranked according to the number of times the correct solution was given in Table 6.1b. This ranking is significantly similar at the 5% level to the secondary task ordering (Kendall's  $\tau = 0.63$ ,  $p < 0.05$ , Kendall, 1975). Problem 14 was solved more often than any other problem, and the arithmetic task indicates that the load it presents is less than that presented by any other problem. At the other end of the scale, problems 7, 12 and 11 were solved least often, and they seemed to be the problems which presented the greatest load to the subjects. The five problems in the centre of the scales, however, display little apparent agreement in ranking.

Figure 6.5 gives some suggestion of why this might be so. Each point on the graph represents a problem. The scale on the horizontal axis represents the number of times (out of a possible maximum of sixteen) that each problem was incorrectly answered. The scale on the vertical axis is the arithmetic task error proportion for each problem, summed for all subjects. In other words, the scales are the interval equivalents of the ordinal rankings of Table 6.1.

At the bottom left hand corner of the graph is problem 14, often solved, and with a low processing load. At the top right are problems 7, 11 and 12, seldom solved, and with a high load. The central group consists of the other five problems. These problems cluster together on both scales. For these problems, the solution rates are all relatively similar, and so are the arithmetic error scores. This indicates that the two rankings of these problems show

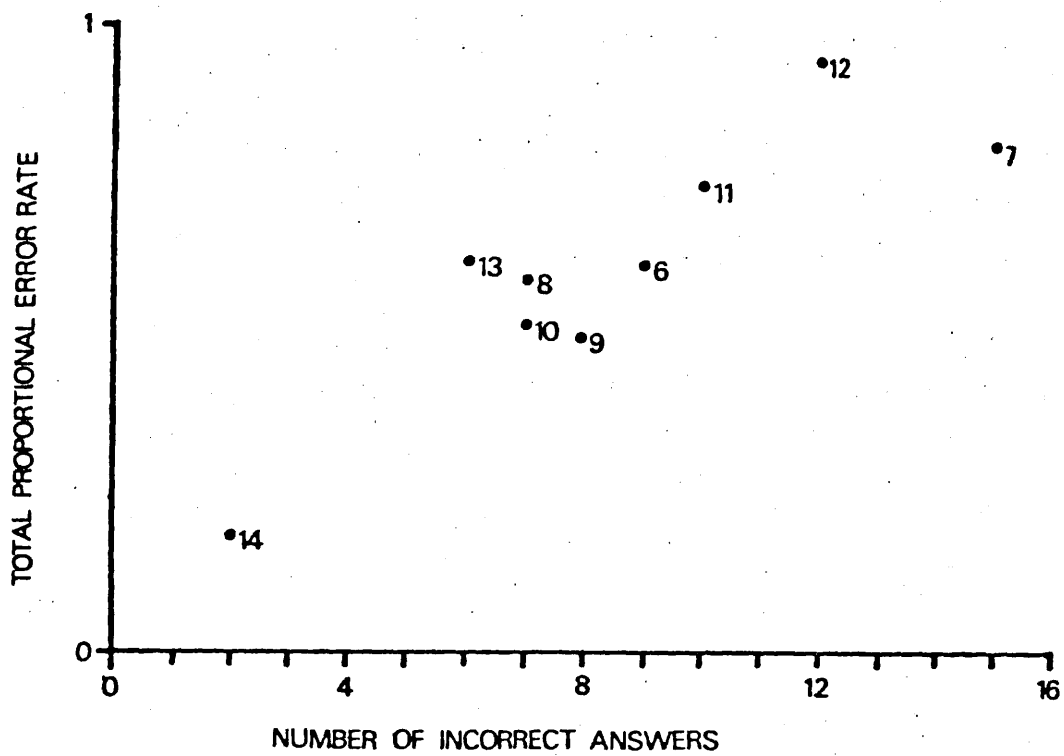


Figure 6.5. Total proportional secondary task error rate plotted against the number of incorrect answers, for problems 6 to 14.



agreement because the problems are of relatively similar complexity. Neither ranking method can differentiate between these problems. This effect may also account for the difference between the two rankings of the "high load, seldom solved" problems (problems 7, 11 and 1?).

#### 6.5 Conclusions and Discussion

The significant agreement in ranking of the problems by secondary task error rate and by solution rate clearly indicates that this type of mental arithmetic secondary task is suitable for measuring problem complexity as intended. This conclusion is pictorially reinforced by the graph in Figure 6.5, in which the points lie within a well-defined band running from bottom left to top right, as would be expected if the secondary task is measuring the load associated with the problem. I therefore believe that secondary task techniques of the type presented here can be applied in the general area of assessment of problem complexity.

I think it is worth pointing out explicitly that at present all measurements of the complexity or difficulty that a problem or task presents to a subject must be of an ordinal type. That is, it is quite meaningful to order tasks according to difficulty, but one would be treading on dangerous ground if one were to try to make comparisons of task difficulties according to some kind of interval scale. At present there is no interval scale of difficulty of mental processing load. Work such as that described in this and the previous chapter may ultimately lead to the setting up of one, but little of the literature reviewed in the previous chapter is concerned with this issue, and none considers it to be critical for immediate purposes. Ordinal comparisons suffice for the time being. It is plain to see that to

attempt to set up an interval scale of problem difficulty immediately raises questions as to what, precisely, is meant by problem difficulty. Does it, indeed, have any more than an ordinal meaning? If it does, I suspect that it would best be interpreted in terms of processing load in a limited capacity model. But at the moment this is of little practical assistance since, as far as secondary task techniques are concerned, it is neither certain nor assumed that there is a precise quantitative relationship between the performance variable being measured and the primary task processing loads. The secondary task measuring techniques must thus remain ordinal in nature.

For this reason, I have confined statistical measurements in this chapter to ranking tests. Similarly, the graph in Figure 6.5 should not be interpreted as anything other than a comparison of the data on which two rankings were based. It was invoked to illustrate the clustering of certain points on both scales, indicating that a ranking of these points alone by either method is unlikely to be meaningful. In this sense, the corresponding problems may be said to be "relatively similar" in complexity. It would be unwise to draw any further conclusions concerning magnitudes of differences between the complexities of the problems on the basis of the graph. It should be regarded only as an illustration of the significant concordance of rankings.

#### Improvements in Sensitivity

In future work of this type, steps might be taken to make the secondary arithmetic task more sensitive. It is felt that some such modifications are possible. In the experiment I have described, a maximum of twenty-three responses to the secondary task could be made

by a subject in the course of a two-minute problem. A subject rarely makes more than ten errors, however, and the average is considerably less. This leads to ties in the ranking of a subject's arithmetic task score, and indicates that the method may lack sensitivity. I suggest that by increasing the digit presentation rate from one every five seconds to one every four (or perhaps even three) seconds the sensitivity would be increased because the task would be more difficult, and because the maximum number of possible responses would be greater. These factors would lead to increased total error rates and a greater spread of errors. Similar effects might also be produced by use of a different, more difficult, type of arithmetic task.

With such improvements, the sensitivity of the secondary task experiment I have described might be improved to a degree at which the resolution of the technique would be great enough to separate significantly the complexity of problems such as those that appear in the central group of Figure 6.5. However, it may be that an arithmetic (or similar) task will ultimately prove not to be the most suitable type of secondary task.

The most worrying aspect of the arithmetic secondary task technique is the lack of control over the time-sharing strategy of the subjects. Although subjects are instructed to give priority to the primary task, it is impossible to ensure that they do this, even with their full cooperation. A subject presented with a particularly difficult problem, for example, may simply give up and turn to the arithmetic task instead, consequently obtaining a misleadingly high score on the arithmetic. For reasons of this nature, we looked briefly at the use of a secondary task which had a far smaller overtly cognitive aspect than the arithmetic task.

### A Foot-Tapping Secondary Task

Subjects were asked to tap a foot pedal at a constant rate, while concentrating on the primary task. Such a secondary task is an example of an interval production test. It has been found (for example, Michon, 1966 and Vroon & Vroon, 1973) that both the regularity of interval production and mean interval production rate have corresponded significantly to the independently assessed complexity of primary tasks in such interval production tests. Such a secondary task would eliminate to a large extent the problems of time-sharing decisions: the way in which foot tapping is regulated seems relatively independent of other thought processes.

Seven subjects attempted the same sequence of chess problems as used for the experiment already described but using foot tapping as the secondary task. A further three subjects were presented with the series with problems 7 to 14 in reverse order. As before, only problems 6 to 14 were considered in analysis. To detect possible fatigue effects, the subjects were required to carry out the tapping task without any chess problem at the beginning and end of an experimental run. It was suggested that subjects tap one of their feet (whichever they preferred) at a rate of about one tap every 1.5 seconds. The actual rates varied considerably from subject to subject. As in the arithmetic task experiment, two minutes were allowed for each problem.

The tapping was recorded using a foot pedal connected to a microprocessor-based data collection system (described fully by Storr *et al*, 1979), which stored data on cassette tape for future analysis. For each subject, the tapping associated with each problem was scored according to several parameters: mean tapping rate; deviation from the

mean; and various running mean measurements to show up any gradual consistent changes. Thus both tapping rate and regularity were considered in the analysis. Rankings of problems for each subject were produced for all measures.

None of the rankings showed significant concordance over all subjects. The experiment did not indicate that a foot tapping interval production test was likely to be useful in assessment of task complexity. It is quite possible that an interval production test is not a suitable type of secondary task to assess the load due to complex problems. The problem may be that the primary and secondary tasks do not in this case compete sufficiently for processing capacity, or it may be that the technique is simply not sensitive enough. The problem is unlikely to lie with the measuring equipment, which was extremely accurate.

I would be reluctant, however, on the basis of this one experiment, to discard all interval production test techniques as secondary tasks for assessing problem complexity. The interaction of a primary task with an interval production test is in general less well understood than with other secondary tasks, and it seems possible that an experiment constructed after more lengthy consideration of the theoretical aspects might show a positive result. Certainly, the advantages that an interval production test appears to offer make some further effort in this area worthwhile.

#### Concluding Remark

The arithmetic secondary task experiment that has been presented in this chapter indicates that secondary task techniques can be used

to measure the complexity of "difficult" problems for which no obvious solution is available. The chess problems with which this work has been concerned are similar in important ways to the complex decision problems with which decision-makers are likely to be faced, and which arise frequently in the practice of Operational Research. I firmly believe that the method has potential for extension to "real" decision problems, the precise nature of which may be difficult to define, and where unique "best" solutions do not exist. The secondary task method can provide a new technique for Operational Research scientists to use in the investigation of a variety of complex decision-making tasks.

## CHAPTER 7

### EXPERIENCE WITH THE ORGANISATIONAL CONTROL GAME

#### 7.1 Introduction

This chapter, and that following, is concerned with describing and discussing the research carried out in our programme using the Organisational Control Game. It will be recalled from Chapter 4 that in total three series of games were played. These were:

- I. The "Student" games, in which the players were students at Royal Holloway College.
- II. The "Emergency Planning Officer" games, in which the players were Emergency Planning Officers from various counties.
- III. The "Major" games, in which the players were serving army officers, all with the rank of major at the time of play.

The primary aim of the first series of games, played with students, was to test the Superior Commander system and the Organisational Control Game. This testing has already been covered in Chapters 3 and 4, and I shall not refer to the "Student" series in this chapter except to compare it with later games. This chapter will consider the research output of the second and third series of games. I begin by describing and discussing the "Emergency Planning Officer" games, which were intended partially as development games, but were also envisaged as providing some research output.

## 7.2 Games Played with Emergency Planning Officers

The players in the "Emergency Planning Officer" series of games were eight retired military officers, all of whom held senior positions in United Kingdom local government at the time of play. All these senior positions were to do with local government emergency planning: indeed, the majority of players were County Emergency Planning Officers. Thus, the players all had had previous military experience, and were, at the time of play, professional decision-makers. We could expect, therefore, to learn more from their comments than we could from the students who had played previously. As a result, we would gain more confidence about the realism of the military scenarios and of the decision-making environment.

### Experimental Procedure

Two players were invited each week to Royal Holloway College to play the Organisational Control Game. One player played Scenario 1 and the other Scenario 2. They were briefed and debriefed together, and played as if they were two separate Divisional Commanders within the same army. During play there was no contact between them. In total, four players played Scenario 1, and a different four players played Scenario 2. A research phase was run at three stages in each game (in rounds 1, 3 and 5) in which each player was asked to list the decision options he saw open to himself and to the enemy.

On no occasion did a player express disbelief about the reports he received from his Task Force Commanders, or suggest in any way that it seemed automatic and not a response to his suggestion or directives. We were able to conclude, therefore, that the Superior Commander deception can be made to work not only with students but also with experienced professionals.



### A Counter Replacement Test

For some of the players an additional experiment was devised. At the end of rounds 2, 4 and 6, four of the players (two playing each scenario) had the counters representing all the units removed from the map board and they were asked to replace them as accurately as they could remember. Notice was given in the briefing of this subsidiary experiment though neither the exact timing nor its purpose was specified. The first replacement was designed as a control, to allow for learning, and the other two as measures of performance. The purpose was to attempt to measure player involvement in the game by investigating whether errors in replacement or the change in error with time varied from one part of the map to another. In particular, counters that were a more prominent feature of the game might be replaced more accurately than others not so heavily involved.

For each group of counters (representing a Task Force or combat unit of similar size) the distance in hexes between the true position and the player's estimate of it was measured for each counter. Individual counters within a group were not distinguished; errors were recorded so as to minimise the error sum of squares in the group.

In Scenario 2 the two players tested taken together made a significant improvement in their overall accuracy between rounds 2 and 4 and between rounds 2 and 6 (significant at the 5% level, Wilcoxon Test, in Siegel, 1956). The players replaced the counters representing one of the friendly units significantly less well than those of the other friendly units, using the same test and criteria as above. This unit was moving independently to engage with the enemy

later, while the others were heavily involved in battle. Similarly, with the enemy forces, a unit which was moving to a supporting position was less accurately replaced than enemy units which were heavily engaged in battle.

No such relationships could be found for Scenario 1, however. There could be a number of explanations for this, the most plausible perhaps being that in this case even the units engaged in battle move a substantial distance each round. The red units not in contact are quite close to those that are, which could be confusing to the players. Further, there are no specific named geographical features associated with each part of the battle. However, in Scenario 2, two units get locked together in battle in the small area of Harris Hill and a further two fight it out in Olivier. These two features obviously aid accurate replacement.

The results of this test indicate that specific scenario effects were more important than intrinsic player aptitude in the counter replacement task, and the test serves to highlight these scenario differences. We did not continue with these tests as they were clearly very likely to interfere with the smooth running of the game, and our main aim was to extract data concerning the perception of decision options. Further details of the counter replacement test may be found in our second progress report (McDowell et al, 1979a).

### Perception of Options

Each player was asked to list the decision options he saw open to himself and to the enemy in a research phase at the ends of rounds 1, 3 and 5. The number of listed options are shown in Table 7.1 for

Player	Scenario	Round					
		1		3		5	
		B	R	B	R	B	R
1	1	2	3	4	2	3	2
2	1	2	3	1	2	1	2
3	1	5	4	2	3	3	3
4	1	2	3	1	1	1	1
5	2	2	1	2	2	4	3
6	2	6	4	5	2	5	2
7	2	2	3	3	2	6	4
8	2	3	3	5	2	4	3

Table 7.1. Player's perceptions of number of options open to Blue (B) and Red (R) forces.

friendly Blue forces (B) and enemy Red forces (R). A form was used which had numbered spaces provided and the figures quoted are the number of separate spaces filled in. Some players tended to list alternatives as a single option and others listed as separate options what were in fact minor variations of a single one. However, these effects can be taken into account by the method of analysis described below, assuming players were consistent throughout a game.

At first it is difficult to see any trend in the results at all. Consider, however, whether a player sees more options open to himself than the enemy or vice versa. Taking each round in turn, the ratio of Blue to Red options may be assigned a positive grade (+, where  $B > R$ ), a neutral grade (0, where  $B = R$ ), and a negative grade (-, where  $B < R$ ). These grades are shown in Table 7.2. Using values of +1, 0, and -1 for the three types of grade, the grades are summed. Then the players can be ranked from 1 to 8, assigning the value 1 for the most positive set of ratios. In Table 7.2 the ranks for each player are shown, mid-ranks being used for equal scores. Using the Wilcoxon Test (Siegel, 1956) the ranks show a significant difference between the two scenarios at the 5% level, indicating that in Scenario 1, the Defence, players see less options open to themselves compared to the enemy than in Scenario 2, the Advance-to-Contact. This confirms what might be expected: in the Defence, players would tend to perceive the enemy as having the initiative, whereas in the Advance-to-Contact, players would tend to perceive the friendly forces as having quite a high degree of initiative.

Considering changes in perceptions of options perceived during the game, the ratio of options seen for Blue forces for Period 3/period 1, for Periods 5/Period 3, and for Period 5/Period 1, were graded

Player	Scenario	Period			Rank
		1	3	5	
1	1	-	+	+	4.5
2	1	-	-	-	3
3	1	+	-	0	6
4	1	-	0	0	7
5	2	+	0	+	2.5
6	2	+	+	+	1
7	2	-	+	+	4.5
8	2	0	+	+	2.5

Table 7.2. Ratio of Blue to Red options (+, where  $B > R$ ; 0, where  $B = R$ ; -, where  $B < R$ ), and ranks of players by these ratios.

as above. This was also done for the Red options, and the results are shown, together with their rankings, in Tables 7.3 and 7.4. For the friendly options differences between scenarios would only be significant by the Wilcoxon Test if we took 15% as the criterion, and for the enemy options 7%, instead of the usual 5%. I do not feel that this demonstrates any effect conclusively.

This type of analysis, based on the counting of decision options, was what we had originally envisaged as being the primary method of analysis of the game output. It may now be clear, however, that it was not showing itself to be the fruitful approach we had originally imagined it to be. This practical disappointment with our analysis techniques coincided with some theoretical misgivings about our approach which we developed during the play of this second series of games. I shall describe these misgivings, and the consequent reorientation of our approach, shortly.

First of all, however, it will be of value to compare the general decision-making behaviour of our emergency planning officers, whom we may designate "experienced" commanders, with the student players in our previous series of games, who are clearly "inexperienced" Commanders. Such an analysis is qualitative and subjective, but nevertheless important, since it indicates a direction in which such decision-making research as ours might usefully proceed. We have already commented on this analysis in the O.R. literature (Cooper et al, 1980).

#### Experienced and Inexperienced Commanders

The difference in the behaviour of the two groups of players can

Player	Scenario	Friendly Options			Rank
		<u>Round 3</u> Round 1	<u>Round 5</u> Round 3	<u>Round 5</u> Round 1	
1	1	+	-	+	3.5
2	1	-	0	-	7
3	1	-	+	-	5
4	1	-	0	-	7
5	2	0	+	+	2
6	2	-	0	-	7
7	2	+	+	+	1
8	2	+	-	+	3.5

Table 7.3. Ratios of Blue options in different rounds.

Player	Scenario	Enemy Options			Rank
		<u>Round 3</u> Round 1	<u>Round 5</u> Round 3	<u>Round 5</u> Round 1	
1	1	-	0	-	6
2	1	-	0	-	6
3	1	-	0	-	6
4	1	-	0	-	6
5	2	+	+	+	1
6	2	-	0	-	6
7	2	-	+	+	2
8	2	-	+	0	3

Table 7.4. Ratios of Red options in different rounds.



be grouped into three categories, although these categories are inter-related and should not be considered in isolation.

1. The officers who played the game used their previous experience of command. They considered factors which they had found valuable in practice, even though these factors were not mentioned explicitly in the game rules or briefing. They injected their own ideas, assumptions and preconceptions into the game; to an extent they "built their own scenarios" using the one we had provided as a basis. The students, on the other hand, tended to accept the scenario as presented, probably due to their experience with commercial wargames (in which game rules are more rigid and there is less flexibility).
2. The officers had a more developed knowledge than the students, performing equally well on both the Defence scenario and the Advance-to-Contact. They were more conscious of the losses that were possible in each case, and they had a better grasp of the principles of war (although these were violated on occasion). The students preferred attack to defence, and showed little appreciation of possible losses. While they managed advance-to-contact quite well they had almost no idea about the way to conduct a defence.

3. The experienced players viewed the problems posed by the scenarios as complex and ill-defined, with no "best" solution. In reports to superior and subordinate commanders in the game they assessed situations more fully and formed better strategic views of the scenarios. In contrast, the students made rather cursory assessments, looking for unique "best" solutions. They tended to see situations in uncomplicated terms, viewing all data as known with certainty, with no "fog of war". They also tended to regard all data as equally important, whereas the experienced players were more selective in their approach.

Both groups of players exhibited a wide range of playing styles, and both groups appeared very involved in the game. As commented on earlier, the deception inherent in the Superior Commander System, which was known to work for student players, was also achieved with the experienced players: they were convinced that the game was interactive. As noted above, the experienced players made assumptions which "expanded" the original scenario, perhaps to make it more lifelike; it is suspected that this process, together with the preconception that no attempt would be made to deceive senior managers who were doing the experimenters a favour by playing the game, is important to the success of the deception.

It seems clear that, although student players are very useful in the initial process of designing the general structure of a game, planning experiments and testing for gross breaches of validity,

experienced players must be involved in the final detailed design (see also related comments by Hartley et al, 1979). After each of the second series of games a debriefing session was held, when the players themselves made valuable contributions and suggestions for improving the realism and level of detail of the scenarios, the mechanics of representing units and transmitting information, and methods for extracting research data. The conceptual abilities and verbal skills of the experienced players were decisive in this process. Thus scenario creation should be a dynamic, interactive procedure, involving both the game designers and players with suitable background and experience, and this procedure should be a continuous one. Each series of games then becomes a learning process for the game designers, who benefit by having their scenarios probed each time the game is played (see also Brewer, 1978).

A simple but nevertheless important point demonstrated by the games so far described is that students do not behave in the same way in a game environment as experienced decision-makers. Therefore, it would be unwise to generalise game results obtained with student players to professional decision-makers. Differences in behaviour are both quantitative and qualitative. We expected to find quantitative differences in the alternative options considered and in the levels of skill displayed. However, we also observed important qualitative differences. The students approached the game differently from the experienced players, and exhibited a different range of decision-making styles. The quantitative and qualitative differences are apparent in all three of the categories identified at the beginning of this section.

Many behavioural studies in the Operational Research and

management literature have used student subjects to derive conclusions about professional decision-making. It is clear that extreme care is required in making such derivations. This point had been demonstrated before, notably in accounting and financial contexts (see Abdel-Khalik, 1974; Alpert, 1967; Copeland et al, 1973; Fleming, 1969). Birnberg and Nath (1968) identify two ways in which student and professional decision-makers are likely to differ: in "common skills and experience" and "basic personality traits" (p.40). These factors are likely to result in the quantitative and qualitative differences noted above. Students, having considerably less experience and training in the decision-making environment, display less quantitative skill when required to work in that environment. Their different personality traits, including cognitive styles, values and beliefs are bound to have a qualitative effect on their decision making.

In general, qualitative differences in decision-making styles are likely to become apparent in comparisons between any distinct groups of individuals. This is seldom explicitly recognised in Operational Research, although instances where it is implied can readily be found. Bowen (1978), for example, notes the limitations of having a Royal Navy Officer representing an enemy submarine commander (p.43). Thus, the recognition and study of different decision-making styles is essential to the understanding of decision-making.

### 7.3 A Re-examination of Research Strategy

Consideration of our first and second series of games have led to the explicit realisation that different types of player approach a decision-making environment in different ways. This may seem to be a rather self-evident finding. However, the difference between

different types of player is not the whole point. Rather, as I tried to emphasise in the last section, different types of decision-maker appear to differ from each other in systematic ways which are related to their backgrounds and skills. The problem is: how can we describe the differences usefully, and determine measures of the differences? The comments of the previous section should be regarded not so much as results, but as statements of the characteristics which we should be studying and attempting to quantify, in some sense, or at least demonstrate explicitly.

It should be realised, too, that despite our declaration that different types of decision-makers - in this case, experienced and inexperienced - differ in their behaviour in a decision-making environment, we are definitely not trying to say that all decision-makers of one type behave in precisely the same fashion. To state this would be a parody of our results. It is very clear that decision-makers are individuals, who come to their decision-making environments from their own idiosyncratic backgrounds and provide their own skills, peculiar to themselves. Although similarities between members of a group of decision-makers of one type may exist, these should not blind us to the fact that each individual decision-maker approaches his decision-making environment in an entirely unique way.

It seems that any useful study of decision-making must attempt to approach the process of decision-making in terms appropriate to the individual decision-maker. Such an individualistic approach would enable us to develop a deeper and more valid understanding of decision-making than it would seem that we are likely to achieve by attempting to fit decision-making processes into a rigid model that concentrates on rather arbitrary indicators as a guide to the processes.

It will be recalled from Chapter 1 that our initial research strategy was to select a decision-making hypothesis, and test it using the game. This would lead to the formulation of a decision-making rule, and its possible subsequent testing using the game. The primary measure of the decision-making hypotheses in which we were interested were the players' perceptions of the number of decision-making options open to both the players and the enemy. The way in which the hypotheses were formulated (see Milburn, 1972) led naturally to a measure which would be intimately related to enumeration of options.

However, it now becomes clear that what one player considers to be describable as a valid and distinct option may be entirely meaningless to another player. Players may differ in the detail they feel should be ascribed to a distinct option, the degree of forward planning incorporated in its formulation, the constraints imposed on its formulation by the decision-making environment, and the features which characterise it is an individual entity to be considered separately from other courses of action. In short, the processes of thought by which a set of possible courses of action are formulated, and then considered, are likely to differ between even two apparently similar decision-makers.

Differences in this decision-making style are likely to be related to personality characteristics (or, indeed, to define them) as well as criteria such as professional experience. For example, Hudson (1972) describes a dimension of personality which ranges from "convergent" to "divergent". He explains (p.67): "divergers are more open to the irrational elements in their own mental functioning"; convergers "construct robust 'ego-boundaries', and include within

them only what is rational", whereas divergers "form relatively weak ego-boundaries and allow their own irrational impulses to suffuse their perception of who they are". It seems unlikely that decision-makers who differ in their convergent/divergent profile will approach a decision problem in the same fashion (indeed, divergers are noted for their "open-ended" approach to problem solving). There are numerous aspects of personality which are likely to have a bearing on decision-making, some well-defined, such as introversion/extraversion (see, for example, Eysenck, 1973), others less so.

If, as I have argued, it would be difficult to get two decision-makers to agree over precisely what constitutes an option, it follows that as researchers we will have problems in using the idea of options as units of decision-making. We have a choice. We can define the decision-making environment in our own terms, and measure how decision-makers behave in our terms. If we do this, we are measuring in part the degree to which the decision-maker's model of his environment corresponds with our own. This measure and the measure of decision-making performance we are supposed to be studying will be inextricably and indefinably intertwined - such a situation is clearly unsatisfactory for research purposes.

Alternatively, we can carry out our research entirely in the terms of the decision-maker. This means, essentially, relinquishing our own model of the decision-making environment, and discovering and adopting that of the decision-maker himself (a different model for each decision-maker, of course). The work so far described in this thesis has led me to believe that this is the right approach. But it has far-reaching consequences. As things stand, we have no knowledge of the decision-maker's model of his decision-making environment. We

cannot define what is meant, in the decision-maker's terms, by an option. When, further, we try to describe processes such as "rigidity of thought", as met in Chapter 1, in the decision-maker's terms, we are in even deeper water. Before we can begin to discuss such ideas, we need a way of modelling the decision-making environment as the decision-maker sees it.

This means, then, that the research strategy as stated in Chapter 1, of hypothesis selection and testing, is somewhat premature. Before we could even begin to consider such a programme, we would need a technique or language for describing explicitly the subjective environment of the decision-maker. It is to this end that our third series of games was directed, and I can now present our revised research aim, which is, quite simply, to develop a means of describing the environment of the decision-maker in his own terms, in a way that provides a model of that subjective environment rich enough to be discussed and studied.

In the next chapter, I describe in detail a technique, cognitive mapping, which we adopted to achieve this aim. In the next section of this chapter I describe our third, "Major", series of games which provided the raw material of the cognitive map analysis.

I conclude this section by noting that this subjective approach to decision-making appears to be gaining popularity in Operational Research at the moment. Two relatively recently developed techniques of decision analysis, the hypergame technique (Bennett, 1980) and analysis of options (Radford, 1975) emphasise the importance of the different perceptions of various actors in the decision situation, and their value as decision aids stems in large part from the way in which



a consideration of the perceptions of other actors is demanded by the methodologies. The work of Eden et al (1979), which will be met in the next chapter, is very much concerned with making the perceptions of the decision-maker explicit to himself, the idea being that the decision-maker is aided by a fuller understanding of the full ramifications of his own formulation and feelings concerning the decision problem.

Such a mood indicates a trend away from rigorous problem definition, in objective fashion, to idiosyncratic and subject-dependent definition. In these terms a problem can only be said to be resolved when the decision-maker acknowledges it to be so. It goes without saying that the criteria by which the solution is assessed are defined by the decision-maker himself. Decision-making aids may be seen as means whereby a decision-making environment is defined as fully and naturally as possible to allow the decision-maker's unique capabilities, such as invention and intuition, to be utilised. Decision-making aids may be catalysts to problem resolution, rather than themselves the means of resolution. Such aids may have something in common with the ancient Chinese divination technique the "I Ching", of which Watts (1957) explains (p.35): "one does not consult the oracle without proper preparation, without going quietly and meticulously through the prescribed rituals in order to bring the mind into that calm state where the 'intuition' is felt to act more effectively". A similar concern with catalysis of intuition and invention can be seen in modern, sometime outrageously off-beat, decision aids (for example, a set of "oblique strategy" cards by Ino & Schmidt, 1975).

#### 7.4 Games Played with Army Majors

The third, "Major" series of games were to have been played under different conditions of stress, generated by noise. In the event, we were unable to obtain permission to subject army officers to such stress, although we did get as far as preparing and testing equipment. Because of the arguments outlined in the previous section, however, this setback did not greatly disturb us. We believed that we had identified a more fundamental problem than that of decision-making behaviour under stress - namely, that of developing means of discussing decision-making behaviour in a useful way. The "Major" series would be played with this end in mind.

We did, however, include two different level-of-data conditions for communications coming up to the experimental players from the Task Forces. In the "aggregated" condition players received communications from Task Force Commanders, as in the previous games. In the "disaggregated" condition, communications came from the four commanders of individual battalions within each Task Force. In both cases, however, the information in the reports amounted to the same. In analysis of the games we would be looking for differences in decision-making behaviour which could be related to these experimental conditions.

Seven players, labelled A, B, ... G, each played through both of the two game scenarios. At the time of play, all players were serving army officers holding the rank of major. Table 7.5 shows the experimental design (originally we had hoped for eight players, but one had to drop out unexpectedly, and could not be replaced).

Week	Scenario 1	Scenario 2
1	C	G
2	A	B
3	E*	D
4	-	F*
5	G*	C*
6	F	
7	B	A*
8	D*	E

\* indicates a disaggregated game; other games were played in the aggregate condition.

Players A, C and B played Scenario 1 before Scenario 2.  
 Players E, D, F and G played Scenario 2 before Scenario 1.

Player B played two aggregate condition games.  
 Players A, C, D and G played aggregate condition before disaggregate condition.  
 Players E and F played disaggregate condition before aggregate condition.

Each Scenario was played three times in disaggregate condition.

Table 7.5. Experimental design for the "Major" series of games.

I note immediately that in the analysis of the games, no effect could be detected due to the aggregation/disaggregation condition. It seems that, at least at the levels that we used, players were easily able to assimilate the data into usable form irrespective of the format in which it was presented. Further consideration of this aspect of decision-making is interesting but beyond the scope of this thesis. It may be of interest to compare our results with those of Daniel (1980); we have commented briefly in this vein in our Final Report (McDowell et al, 1979b). The aggregation/disaggregation variable will not be mentioned further in the analysis of the games to be presented here.

#### Game Communications

The rawmaterial for the analysis of a player's behaviour in the game was to be his communications in the game. As before, in each of the six rounds of the game the player, in his role as Divisional Commander, issued a directive statement to his subordinate Task Force Commanders, and a report statement to his Superior Corps Commander. In addition, during a research phase in the first, third and fifth rounds, the player provided a commentary on his recent decision-making to the game directing staff.

All three types of material - directives, reports, and decision-making commentaries - were recorded directly onto tape by the players. After the games, the recordings were transcribed (by a skilled secretary) into written form. These transcriptions are virtually word-for-word renderings of the players' speech - only very rarely was it necessary to guess at the content of an inaudible phrase or remark. A sample transcript, for player C in Round 1 of Scenario 1,

is shown in Table 7.6; this communication set is one of the shortest we encountered.

Transcription of verbal communications was an efficient and effective way of recording decision-making data in the game. Occasional ambiguities of meaning did arise in the transcripts - these corresponded to ambiguities in the recordings made by the players. There is no way of resolving these ambiguities. Had the Organisational Control Game been being played "for real", without the deception inherent in the Superior Commander System, a superior or subordinate player might well have queried these ambiguities, had he detected them. It should be emphasised that ambiguities were found extremely rarely. Where they occurred, an arbitrary decision was taken as to the intended meaning. At no time was there any danger of this arbitrary procedure severely altering the meaning of a substantial section of a transcript.

#### Initial Impressions of Communication Transcripts

Some impression of the output of the games was gained by myself and the others associated with the project during the preparation of the transcripts. When the transcripts were complete, they were reread as a set, bearing in mind that we were looking for anything that could be regarded as an indicator or measure of decision-making, and that could be compared from game to game.

It is worth making one point about the movements of reserve forces in the late stages of the game. It becomes immediately clear on studying these movements that it is impossible to attribute meaning to them on the basis of the geographical movements alone. In other

Scenario 1, Player C, Round 1

Task Force Communication Phase

Orders for 3 Task Force: 9,10,11 and 12 battalions. No change, hold your positions.

Orders for 1 Task Force: 1,2,3,4 battalions of 1 Task Force. No change, hold your positions.

Corps Communication Phase

Situation enemy North. Our units in contact are 7,8 and 9 Regiments of Z Division, are in contact and in battle at the moment. Further South 3,1 and 2 Regiments of F Division are moving into contact all in the area West of Pontet-Canet. In the South in contact with our forward elements is 52,51 and 53 Regiments of 7 Division; South-West of Batailley. In the middle of the area we have 1, 3 and 2 Regiments of X Division, whose present location is uncertain, but they are North of Croizet-Bages, are moving East and South-East. Z Division and 7 Division are the two divisions in contact at the moment.

Estimate of strength, material means, morale and knowledge of our situation is: no change.

Conclusions covering courses of action open to the enemy: I estimate the Z Division will be moving forward up the axis Pontet-Canet to Cantemerle and continuing to, if possible, overwhelm our forces in that area. To the South-West the 7 Division will be moving up the road Batailley to Langoa towards Cantemerle.

Our situation: activities of our own forces. No change from previous situation report. Those units in contact are still in contact. Activities of forces not attached: no change.

Conclusions covering courses of action open to our own forces: we are under orders to defend our present locations where in contact at the moment, and will continue to do so.

Research Phase

My appreciation in the decision-making process here was that in both the North and South of my area my directive was to hold my position as long as I could and I had reports from all the Task Force Commanders that although in contact in both of the Task Force areas, they were able to hold their positions, and they were in a reasonable position in which to defend the two thrust axes, so I decided at the moment to leave them where they were.

As far as resupply is concerned, appreciation there was to move up particularly quickly to the Southern Task Force (number 1) more resupply units, in order that they can continue fighting that intensive battle to the South.

Table 7.6. Transcript of communications for Player C in the first round of Scenario 1.

words, a particular movement could have been made for a variety of quite distinct reasons. To identify the reasons, the transcripts of the games must be consulted. That this is so might have been predicted from the arguments already put forward in this chapter. It is, however, interesting to compare this with a similar observation recently reported in the O.R. literature (Huxham et al, 1981) for the case of simple matrix games, where the unsuspecting experimenter might (wrongly) have supposed motivation for particular plays to be fairly obvious, and thus deducible from observation of the plays.

The immediate impression gained from the transcripts of the games was the large variety in the ways in which the players described their interpretation of and behaviour in their decision-making environment. For a given scenario, in which all seven players were led through exactly the same sequence of objective situations and events, individual players would describe the game environment completely differently. We were left in no doubt that their perceptions of the same objective situations differed widely. This is, of course, consistent with our earlier exhortation to regard decision-makers as unique individuals.

Despite what we have already stated in this chapter, in some respects the observation was still somewhat surprising; in other respects it now seemed quite natural. It was surprising, because our players, we know, were of broadly similar backgrounds - they were all army men, all held the rank of Major, all were articulate and apparently competent, and all were presumably of similar cultural and doctrinal background. Thus, that the most obvious feature of their perceptions in the game were the differences was, in one sense, unexpected. Yet, taking a less rigid view of the game, remembering

that the players were not merely subjects with measurable behavioural responses, but were, presumably, as complex as the experimenters themselves, we realised that to expect a high degree of consistency between players might be a ridiculous idea. After all, we had partly chosen the method of tape-recording of free communication in order to permit players to give free reign to their own views rather than constrain them to terse orders or reports (as written communication tended to produce) or to thinking along lines suggested by the experimenters (as a questionnaire methodology would tend to do).

Given the written transcripts of the players game communications, there now remains the problem of translating them into a format more amenable to study and comparison. We carried this out using the technique of cognitive mapping. Since I believe the cognitive map analysis of these games to be the most important part of our decision-making research (as opposed to our work on game design and secondary tasks), I have chosen to devote the entire next chapter to this analysis.



## CHAPTER 8

### A COGNITIVE MAP ANALYSIS OF THE ORGANISATIONAL CONTROL GAME

#### 3.1 Introduction

In this chapter I shall describe the analysis of our third, "Major", series of games by the technique of cognitive mapping. This is a technique from psychology which offers an approach to the analysis of decision-making that begins to tackle the problems that were described in the previous chapter. Its use can be seen as part of a recent trend in some areas of Operational Research to explore methods and techniques which emphasise the individuality of the decision-maker and the uniqueness of his own conception and understanding of his decision-making environment. Some of the work described in this chapter has been published in the Operational Research literature (Klein & Cooper, 1982).

The aim of Operational Research is to aid decision-makers. In the past, this has been achieved largely by the provision of techniques, together with guidelines for their application, directed towards the resolution of reasonably well-defined problems. In this area Operational Research has been unquestionably successful.

Recently, however, attention has been increasingly directed towards a different type of problem. It has now been widely recognised that many of the problems faced by decision-makers exist within complex, ill-defined and interconnected "messes" (Ackoff, 1979), located in fundamentally turbulent environments (Radford, 1978). Accordingly, much attention is now directed towards the problems of strategic decision-making in conflict and crisis. It is recognised

that purely analytical approaches are inadequate to deal with such problems, and the need for the development of a more general "science" of decision-making has been identified (Dando et al, 1977).

An important part of the development of a decision-making science is the provision of more detailed models of the individual decision-maker than exist at the moment. Psychological research provides a wealth of information about the behaviour of individuals, while political science and international relations have long been concerned with high-level decision-making, but it is only recently that combined approaches have been adopted. The interdisciplinary stance of O.R. makes it ideally suited to take an active role in this more unified approach to the study of decision-making. O.R. can both contribute to this area, and learn from it.

The cognitive mapping technique to be described in this chapter is a means of examining the behaviour of individual decision-makers. A cognitive map is a representation of the perceptions and beliefs of an individual about his own subjective world. As will be described in the next section, the technique has been employed in the fields of international relations and administrative science, and recently there have been applications in O.R. It seems certain that cognitive mapping is a fruitful method of analysing decisions in a number of O.R. areas.

In the next section of this chapter, I shall briefly describe the cognitive mapping technique, and concisely review the literature concerning applications in related areas. This will be followed by an example of the preparation of a cognitive map from the output of a round of the Organisational Control Game. I then proceed to present

and discuss the cognitive map analysis of the series of games. The results deal mainly with the initial perceptions of the players in the games, but include some consideration of dynamic aspects and the way in which perceptions alter as the game progresses. Finally, I draw some conclusions about the usefulness and applicability of cognitive mapping, both for the type of decision-making experiments described in this thesis, and also in a wider O.R. context.

### 3.2 Cognitive Mapping

Traditionally, scientists (and therefore by implication most of those involved in O.R.) have been concerned with the real, objective world, where phenomena can be observed and measured. However, human decision processes never take place in this objective world. All that can be observed there are some of their effects. Human decision processes always take place within the subjective world of the individual decision-maker. Cognitive mapping offers a window on this subjective world. It provides a means of representing the way in which a decision-maker models his decision-making environment, in terms of the concepts he himself uses. The technique has been used to examine the causal belief systems of decision-makers in a variety of contexts, including O.R.

A cognitive map is derived from a text, a record of the verbal or written statements of an individual. From the text are identified the concepts used by the individual, and the causal relationships between them. Only two types of relationship are considered in the application discussed here: positive and negative. A positive relationship occurs when a change in a predecessor concept causes a

similar change in the successor: an increase (decrease) in the first causes an increase (decrease) in the second. With a negative relationship, an increase (decrease) in the predecessor causes a decrease (increase) in the successor.

An example will serve to clarify these definitions. Discussing a particular friendly unit, one player stated that "... morale is high, as they are advancing with little opposition ...". Two concepts may be identified in this text: "high morale", and "unopposed advance". These concepts are linked by a positive relationship, since the player believes that an "unopposed advance" causes or promotes "high morale". The two concepts, and the relationship between them, are shown graphically in Figure 3.1.

The coding procedure is generally more complicated than in this example. Relationships between concepts are often obscure and sometimes ambiguous, but with care large cognitive maps containing many concepts and relationships may be coded from a text and pictured graphically. Coding techniques are described in detail by Wrightson (1976). An example of our own coding procedures is given in the next section.

When a text has been fully coded, what will have been constructed is a signed directed graph. Graphs of this type will of course vary in size and complexity. We shall see some examples of those constructed in our own work later in the chapter. For the time being I shall illustrate the "finished article" with a purely fictional example, shown in Figure 3.2, in which the letters representing concepts and the relationships between concepts have been designated with no other purpose in mind than to demonstrate various structural features of a cognitive map.

Text

...morale is high, as they are  
advancing with little opposition...

Concepts

A: High morale.

B: Unopposed advance.

Map

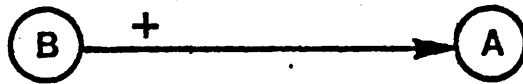


Figure 3.1. Two concepts linked by a positive relationship.

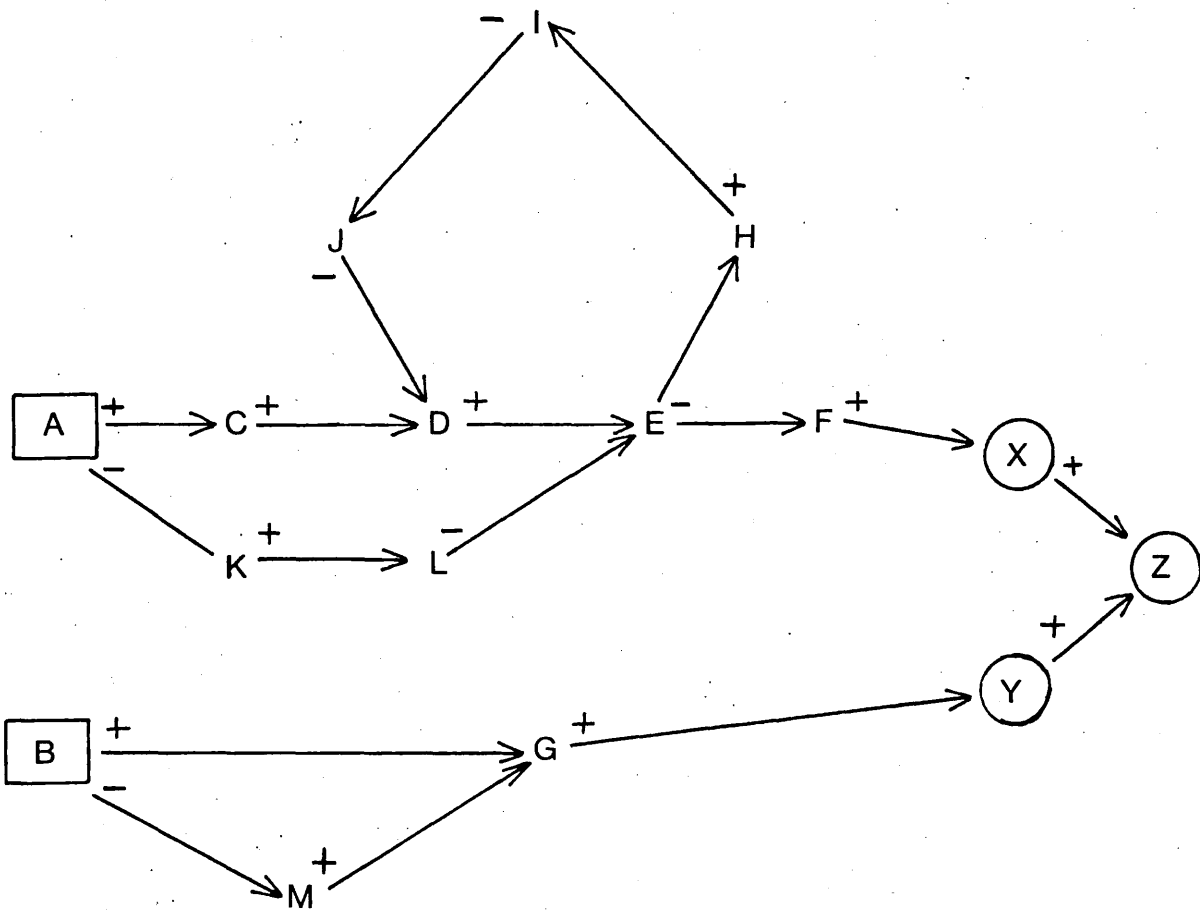


Figure 3.2. A "fictional" cognitive map, illustrating: goals (in circles); actions (in rectangles); a negative feedback loop (D, E, H, I, J, D); a pair of balanced paths (A, C, D, E) and (A, K, L, E); and a pair of unbalanced paths (B, G) and (B, M, G).

To be a valid representation of an individual's cognitive processes, the concepts which appear in the cognitive map must be those which the individual himself uses, and the relationships must be ones which he believes exist. To attribute another set of concepts and beliefs to the individual would defeat the whole object of the exercise, which is essentially to construct a model of the belief system of an individual.

The cognitive mapping technique permits complex and subtle diagrams of individual decision-making to be constructed, and these diagrams can elucidate and give insight into the processes by which a decision-maker organises his thoughts so as to make his decisions. A first step in looking at a cognitive map in more detail is to identify certain types of concept within its structure. The first of these types are goals, which can be broadly defined as concepts which the decision-maker wishes to maximise, at least in part, for their own sake, and not merely because they facilitate some other concept. In a military situation, an obvious example of a goal would be to defeat the enemy. That goal, however, might be regarded as possibly too long-term and non-operational, and might not be the concept with which the decision-maker would naturally work. He might prefer goal concepts which are more immediately realisable, like prevention of forward movement by enemy, or the breaking of enemy supply routes. In a cognitive map it is probably most useful to regard the goals as those ends towards which the decision-maker is immediately working. In our notation, we enclose goal concepts in circles in the cognitive map, as shown in Figure 8.2.

The other type of concept in a cognitive map which it can be useful to identify are actions (enclosed in rectangles in our notation, as in Figure 8.2). These represent real actions that the decision-

maker, and possibly other actors in the environment, might take. An instance of such an action concept might be to attempt to outflank the enemy. The decision-maker can choose to do this, or not, and the decision will have consequences.

By this identification of goals and actions, cognitive maps are given a dynamic aspect. We can now trace paths of positive or negative relationships leading from actions to goals, and we can begin to see how specific actions affect specific goals in the decision-making environment.

As we trace paths and examine structures in a cognitive map, certain features can be identified. The distance between actions and goals is one such feature - sometimes, it is found that several concepts intervene between an action and the goal which it affects, leading to long, linear paths of consequences appearing in the map. In other cases, actions may be very closely linked to goals. Compare the paths (A,C,D,E,F,X) and (B,G,Y) in Figure 3.2.

There is of course, no reason why feedback loops should not occur in the maps. A positive feedback loop, which must have an even number of negative links in its structure, would suggest that any alteration in a concept within a loop would tend to simplify that alteration. On the other hand, a negative feedback loop, with an odd number of negative links, would tend to cancel out such alterations. The structure (D,E,H,I,J,D) in Figure 3.2 is a negative feedback loop.

Other structures that will often appear in maps are multiple paths between two concepts. It seems clear that the situation is relatively well-defined when all the paths have, in total, the same sign. For



example, if all the paths are positive (such as (A,C,D,E) and (A,K,L,E) in Figure 3.2), then obviously an increase in the initial variable causes an increase in the final variable. But what happens if we have, say, one positive path, and one negative path (such as (B,G) and (B,M,G) in Figure 3.2)? The precise relationship between initial and final variable is no longer clear. In fact, there is not enough information in the map to define it. Some measure of path magnitude would be needed to make the overall relationship precise. It would be possible to attempt to assign rough magnitudes to paths to deal with this type of situation; however, it can be argued that the ambiguity in the simple picture may well reflect quite accurately the reality of the decision-maker's cognitive process, that the ambiguity is uncomfortable, and that the decision-maker may, in fact, delete relationships until he is left with a situation which exhibits path balance. That is, in some sense he deliberately fails to consider possibilities which lead to unbalanced path structures, because this strategy will lead to easier decision-making. This type of decision-making behaviour is hypothesised by Axelrod (1976a).

#### Cognitive Map Literature

Cognitive mapping was developed by Robert Axelrod as a means of analysing decision-making in political science. His approach is described in detail in the book "Structure of Decision" (Axelrod, 1976a). As well as a thorough examination of the principles of the political science work of Axelrod and his colleagues, the book includes a complete practical guide to their methods of coding texts - that is, how concepts and relationships are identified in a piece of material to be coded (Wrightson, 1976). The book also contains five case studies.

The case studies include a comparison of the cognitive maps, derived from transcripts, of the participants of the "British Eastern Committee" of 1913, the purpose of which was to determine British policy in what was then known as Persia (Axelrod, 1976b). The author finds the number of causal assertions made by the participants surprisingly high, and records his amazement that a decision-maker could usefully handle so many assertions. In another of the case studies, Hart (1976), investigating international cooperation in the exploitation and conservation of ocean resources, presents cognitive maps of large organisations (rather than individuals) in which the beliefs mapped were derived with the aid of a panel of judges who were experts on the subject. Hart was particularly interested in the structure of the goal concepts of the organisations.

Hart has also used cognitive mapping to compare the belief systems of three Latin American policy makers (Hart, 1977) and also three presidents of the United States (Hart, 1973). Other applications of cognitive mapping in the field of political science include an application to Norwegian oil policy (Bonham et al, 1978), an analysis of the perceptions of American foreign policy officials towards the Arab-Israeli conflict in the early 1970s (Bonham et al, 1979), and an analysis of Japan's decision to go to war in 1941 (Levi & Tetlock, 1980). More general discussions of cognitive mapping and allied techniques in the context of political science include those by Shapiro and Bonham (1973), Bonham et al (1976), and Heradstveit and Narvesen (1978).

Roos and Hall (1980) provide a recent example of the use of cognitive mapping in the field of administrative science. Explaining that essentially they are using influence diagrams whose structure

is determined by the perceptions of the analyst, they use the diagrams they are able to derive to aid analysis of power within an organisation.

One of the most exciting uses of cognitive mapping recently has been in the field of Operational Research, carried out by Colin Eden and his colleagues, and described in detail in their book "Thinking in Organizations" (Eden et al, 1979). Their approach is rather different to that so far described: rather than being primarily concerned with research, they are engaged in actively assisting decision-making by their clients in the real world. They see cognitive mapping as a tool to aid that process.

The approach, or philosophy, of Eden and his colleagues to their clients is that it is the decision-maker with the problem who actually owns the problem, who has the potential to understand it best, and who ultimately has to make the decision. Thus the role of the consultant is not to take over the problem and proceed to solve it, but rather to assist the decision-maker in learning about his problem, understanding it fully, and ultimately leading him to a position in which he feels able to make a decision. This often takes the form of making explicit knowledge that is usually implicit, unformulated, and which may even be regarded as "illegitimate" (that it ought not to be relevant to the problem, even if it is).

One of the primary ways in which Eden and his colleagues assist their clients is by constructing cognitive maps (with computer assistance) in conjunction with the clients. Thus the clients see a model of their own thinking processes being built up. As the process goes on they find themselves questioning the validity of each component of the

map. As the map becomes more accurate, awareness of the full ramifications of the decision-problem increases. The cognitive map is a kind of mirror of the thinking processes of the clients.

The cognitive mapping techniques used by Eden and his colleagues are not surprisingly somewhat different in style from those used in the context of political science. One particular difference is that Eden has blended cognitive mapping with psychological construct theory (Eganister & Fransella, 1971). Two case studies employing this and related methodologies have recently appeared in the literature: one is a description of aid given to the management of an ailing journal (Eden & Jones, 1980); the other describes assistance given to a community group assisting unemployed youth in the city of Bath (Jones & Eden, 1981).

To this account of the cognitive mapping literature I would add our own contribution (Klein & Cooper, 1982), which essentially reports on work covered in this chapter. It may well be the only reported use of cognitive mapping to date as a means of analysis of a highly controlled decision-making experiment. This may not be surprising, however, in view of the fact, made clear in previous chapters, that complex decision-making experiments are rarely subject to such a high degree of control.

In concluding this brief review of the literature it is probably worth making two points, to avoid possible confusion. The first is that the term "cognitive mapping" has been used to describe the construction of mental representations of geographical areas (see, for example, Neisser, 1976): a related area, but not quite the same.

The term "cognitive map" has also been used in an educational context (Skemp, 1979); once again, the concept is related, but quite distinct from the usage here.

The second point is that some of the principles of cognitive mapping can be employed in the field of artificial intelligence (see, for example, Holland & Reitman, 1978). It is clear that this type of explicit modelling of human thought might be particularly valuable in a field which could be considered to be concerned mainly with the computer simulation of mental processes. This application, although fascinating, is beyond the scope of this present brief survey.

### 2.3 Coding Cognitive Maps

To code the Organisational Control Game, the texts used were the transcripts of the players' communications in the game, already described in the previous chapter. Maps were coded from these texts in two phases. The initial texts of each player - his directives to subordinates and his first reports to his superior and to the game controllers - were used to construct maps of each player's initial perceptions of the game environment. Subsequent texts were then used to revise each player's map as his perceptions altered with the progress of time and the development of the game.

The details of the coding process were specifically tailored to the output of the game, and not surprisingly differ in various ways from other coding processes reported in the literature. It may be of value, therefore, to illustrate our coding process by showing how the sample of game output reproduced in the previous chapter (Table 7.6)

is coded. Note that this was one of the shortest texts, and produces one of the smallest maps.

Table 3.1 reproduces the text. In this text concepts have been underlined and labelled in the right hand margin (labels are as used in the scheme employed for all the maps). Table 3.2 lists these concepts concisely, and tabulates the causal links between them. Finally the concepts are put together to form the cognitive map shown in Figure 3.3.

A number of points should be made. First of all, regarding identification of concepts, we were anxious to avoid coding straight descriptions of what was actually happening. Thus, we were only looking for: (a) stated goals; (b) speculation or estimates of relevant factors, including enemy behaviour; (c) friendly actions. Thus the first section of the Corps Communication Phase report was not coded.

Secondly, in listing the concepts we have abbreviated them, making every effort not to lose the sense of the concept as originally stated by the player. It will be seen that the same concept can apparently occur more than once in a text; in such cases the coder must decide whether the concepts are truly the same or not, and code them appropriately (see concepts B1C and B3C).

Thirdly, casual links must often be inferred. For example, no explicit mention is made of the fact that the enemy actions R71 and R7Z1 act against the player's goal G2a, yet this seems clearly to be so, and the coding has been carried out accordingly. Many of the

Scenario 1, Player C, Round 1

<u>Task Force Communication Phase</u>	
Orders for 3 Task Force: 9,10,11 and 12 battalions. No change, hold your positions.	
	B3C
Orders for 1 Task Force: 1,2,3,4 battalions of 1 Task Force. No change, hold your positions.	B1C
<u>Corps Communication Phase</u>	
<p>Situation enemy North. Our units in contact are 7,8 and 9 Regiments of Z Division, are in contact and in battle at the moment. Further South 3,1 and 2 Regiments of 7 Division are moving into contact all in the area West of Pontet-Canet. In the South in contact with our forward elements is 52,51 and 53 Regiments of 7 Division, South-West of Batailley. In the middle of the area we have 1, 3 and 2 Regiments of X Division, whose present location is uncertain, but they are North of Croizet-2ages, are moving East and South-East. Z Division and 7 Division are the two divisions in contact at the moment.</p>	
<p>Estimate of strength, material means, morale and knowledge of our situation is: no change.</p>	
	RFZ1 P8
<p>Conclusions covering courses of action open to the enemy: I estimate the Z Division will be moving forward up the axis Pontet-Canet to Cantemerle and continuing to, if possible, overwhelm our forces in that area. To the South-West the 7 Division will be moving up the road Batailley to Langoa towards Cantemerle.</p>	
	R71
<p>Our situation: activities of our own forces. No change from previous situation report. Those units in contact are still in contact. Activities of forces not attached: no change.</p>	
<p>Conclusions covering courses of action open to our own forces: we are under orders to defend our present locations where in contact at the moment, and will continue to do so.</p>	
<u>Research Phase</u>	G2a B1C, B3C G2a P9
<p>My appreciation in the decision-making process here was that in both the North and South of my area my directive was to hold my position as long as I could and I had reports from all the Task Force Commanders that although in contact in both of the Task Force areas, they were able to hold their positions, and they were in a reasonable position in which to defend the two thrust axes, so I decided at the moment to leave them where they were.</p>	
	B1C, B3C B1CS P10
<p>As far as resupply is concerned, appreciation there was to move up particularly quickly to the Southern Task Force (number 1) more resupply units, in order that they can continue fighting that intensive battle to the South.</p>	
	S8

Table 8.1. Text for Player C, in the first round of Scenario 1,

with concepts underlined and labelled.

Concepts

Goal:	$G2a$	Defend present Task Force locations.
Blue actions:	B1C	Hold position of 1 Task Force.
	B1CS	Move supply to 1 Task Force particularly quickly.
	B3C	Hold position of 3 Task Force.
Red actions:	RFZ1	Z Division to move towards Cantemerle.
	R71	7 Division moves towards Langoa.
Other Variables:	S8	Ability of 1 Task Force to continue fighting intensive battle.
	P8	Change in strength of enemy.
	P9	Assessment of ability to hold position.
	P10	Suitability of Task Force positions to defence.

Links

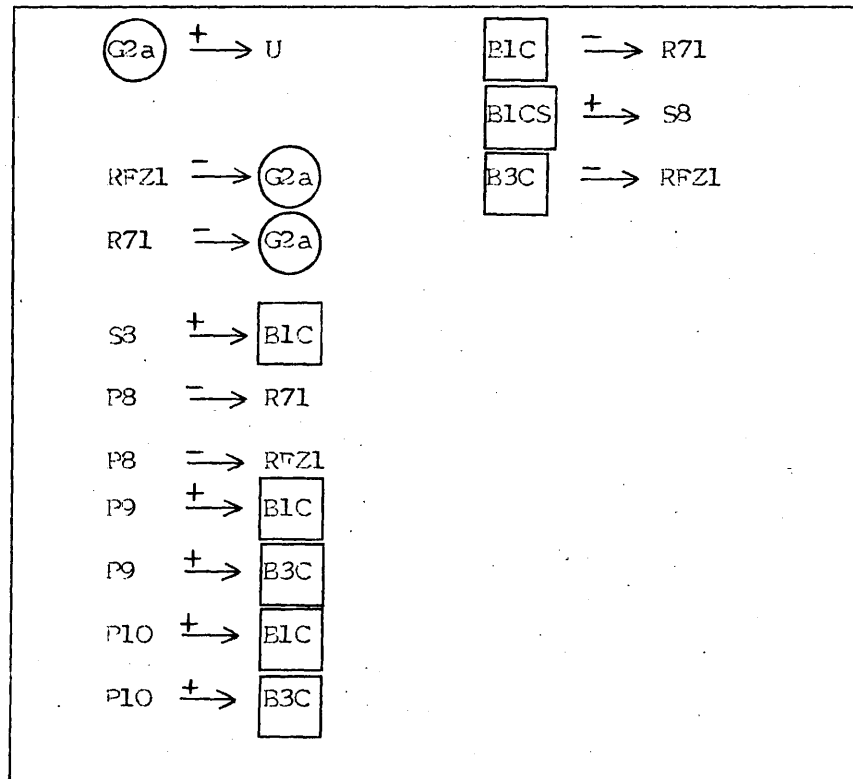


Table 8.2. Concepts and links between concepts for Player C, Scenario 1.



C:1

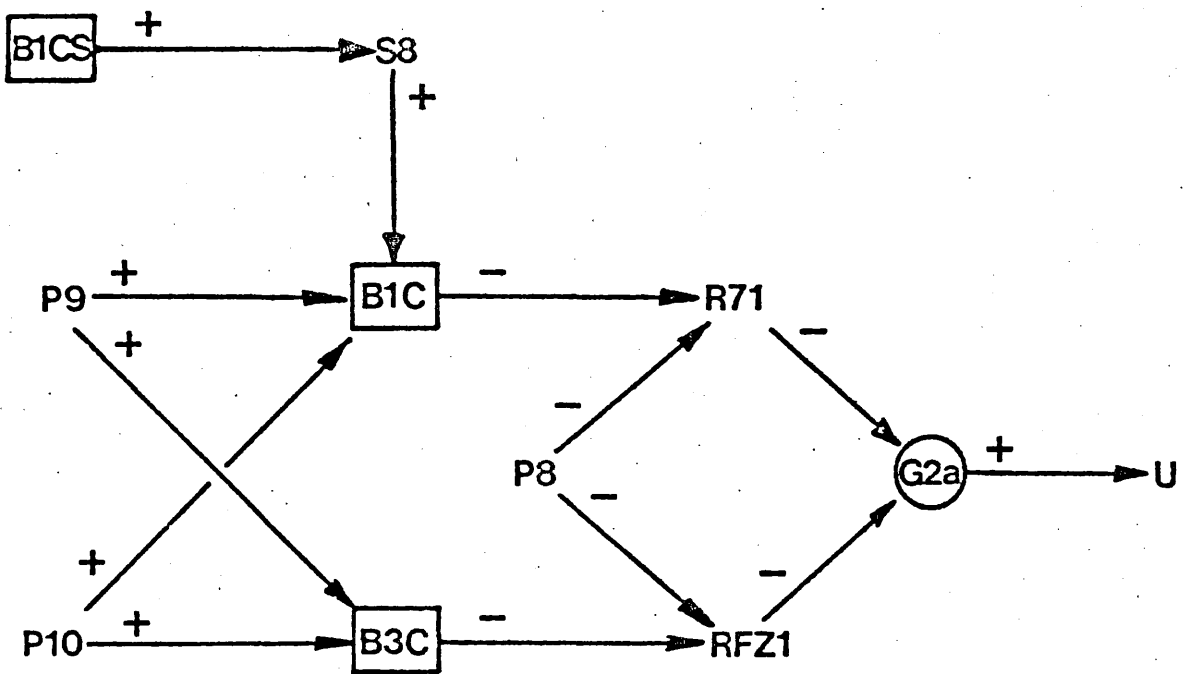


Figure 8.3. Cognitive map of Player C in Scenario 1.

links between concepts must be deduced from the overall sense of the text in this way.

Finally, notice that in the final map, an additional utility variable, U, has been added, purely for completeness. All goal variables will have positive causal links to U, and we may say that the player's formal aim in the game is to maximise U. This may be a useful device for analysis when several, possibly conflicting, goals appear in the cognitive map.

#### 8.4 Results

Coding the initial texts yielded 14 cognitive maps (from the seven players, in each of the two scenarios). Broadly, analysis of the maps indicates marked differences between the players, yet, for any player, marked similarities across the scenarios. We have already seen Player C's map in Scenario 1 (Figure 8.3 and Table 8.2). This map may be compared with the maps for Players A and C in Scenario 2, which, with respect to number of concepts, are the extreme cases for this scenario (Table 8.3 and Figure 8.4; Table 8.4 and Figure 8.5).

#### Concepts and Relationships

The most obvious difference between players is the number of concepts identified in their maps. This ranges from 10 (Player C) to 26 (Player F) in Scenario 1, and from 8 (Player C) to 23 (Player E) in Scenario 2 (see Table 8.5). However, for four of the seven players, Players A, B, C and G, the numbers of concepts in the two scenarios is quite similar, despite the fact that they involve very different military situations. (Recall that in Scenario 1, the

Goals:	<p>(G1)</p> <p>(G2)</p> <p>(G3)</p>	<p>Recapture Olivier.</p> <p>Secure Railway.</p> <p>Proceed to join rest of Corps eventually.</p>
Blue actions:	<p>B2A</p> <p>B3A</p> <p>B3AF</p> <p>B3AH</p> <p>B3E</p> <p>B1A</p>	<p>2 Task Force to attack Olivier.</p> <p>3 Task Force to move on as fast as possible.</p> <p>3 Task Force fall back to Eouscaut if necessary.</p> <p>3 Task Force take up position on Harris Hill.</p> <p>3 Task Force hold Eouscaut.</p> <p>Deploy 1 Task Force to counter B Regiment.</p>
Red actions:	<p>RA?</p> <p>RA3</p> <p>RE1</p> <p>RE?</p>	<p>A Regiment defend Harris Hill Area.</p> <p>A Regiment to move towards gap in hills.</p> <p>B Regiment to go West.</p> <p>F Regiment to go South.</p>
Other Variables:	<p>N1</p> <p>N2</p> <p>N3</p> <p>N4</p> <p>P1</p> <p>P2</p> <p>P3</p> <p>P4</p>	<p>Ability of A Regiment to hold Harris Hill.</p> <p>Present opposition to 3 Task Force.</p> <p>Destruction of A Regiment.</p> <p>Danger of rear attack by A Regiment.</p> <p>Speed of Blue Forces.</p> <p>Knowledge of enemy intentions.</p> <p>Enemy strength and morale.</p> <p>Lack of scope to change anything.</p>

Table 8.3. Concepts for Player A, Scenario 2.

A:2

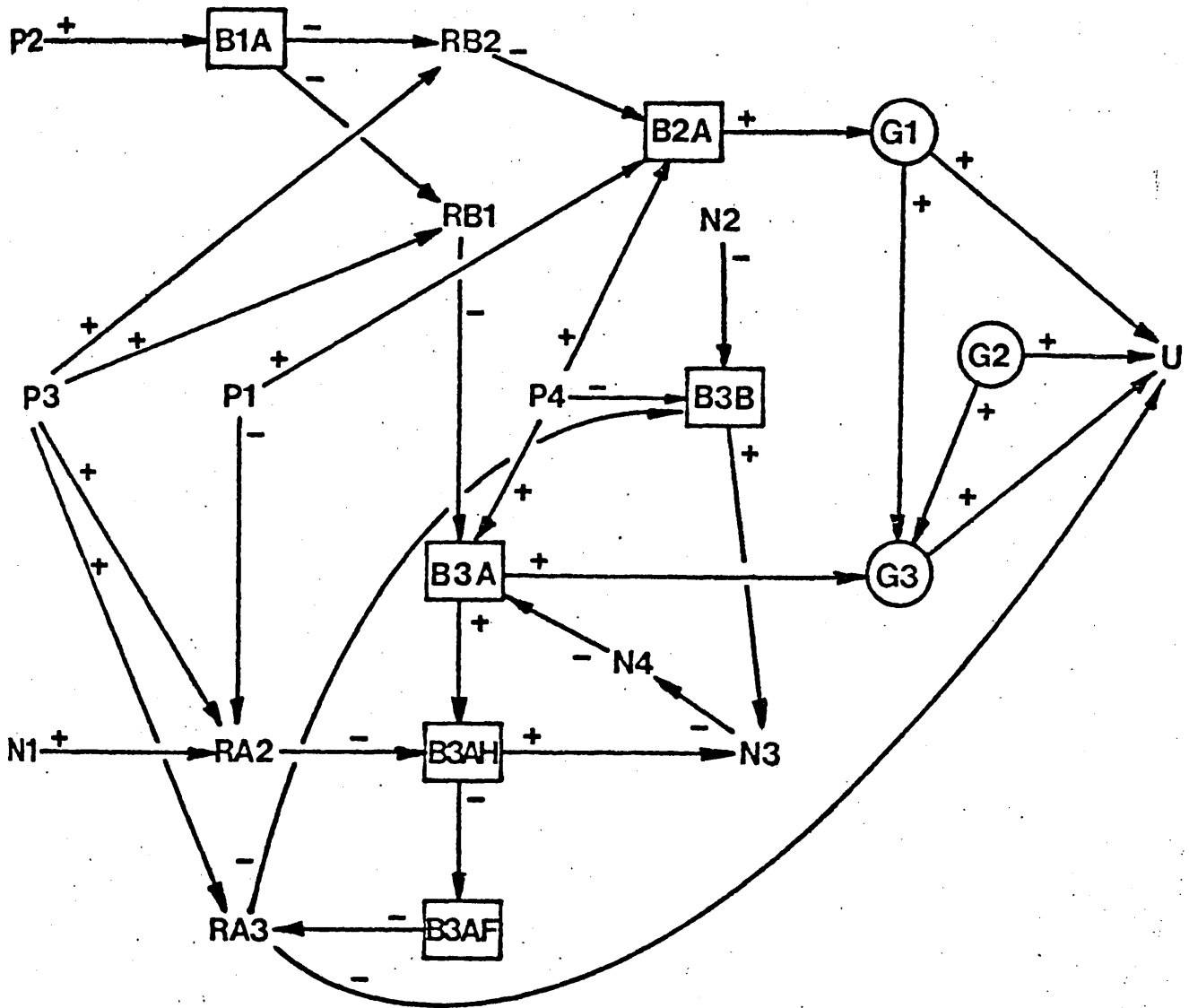


Figure 3.4. Cognitive map of Player A in Scenario 2.

Goals:	<p>(G1)</p> <p>(G2)</p> <p>(G3)</p>	<p>Recapture Olivier.</p> <p>Secure railway.</p> <p>Proceed to join rest of Corps eventually.</p>
Blue actions:	<p>B2C</p> <p>B3D</p> <p>B5</p>	<p>2 Task Force to clear Olivier quickly.</p> <p>3 Task Force to clear advancing enemy.</p> <p>Task Forces to link up eventually.</p>
Red actions:	<p>RA1</p> <p>RCL</p>	<p>A Regiment to move on Eouscaut.</p> <p>C Regiment to hold Olivier.</p>

Table B.4. Concepts for Player C, Scenario 2.

C:2

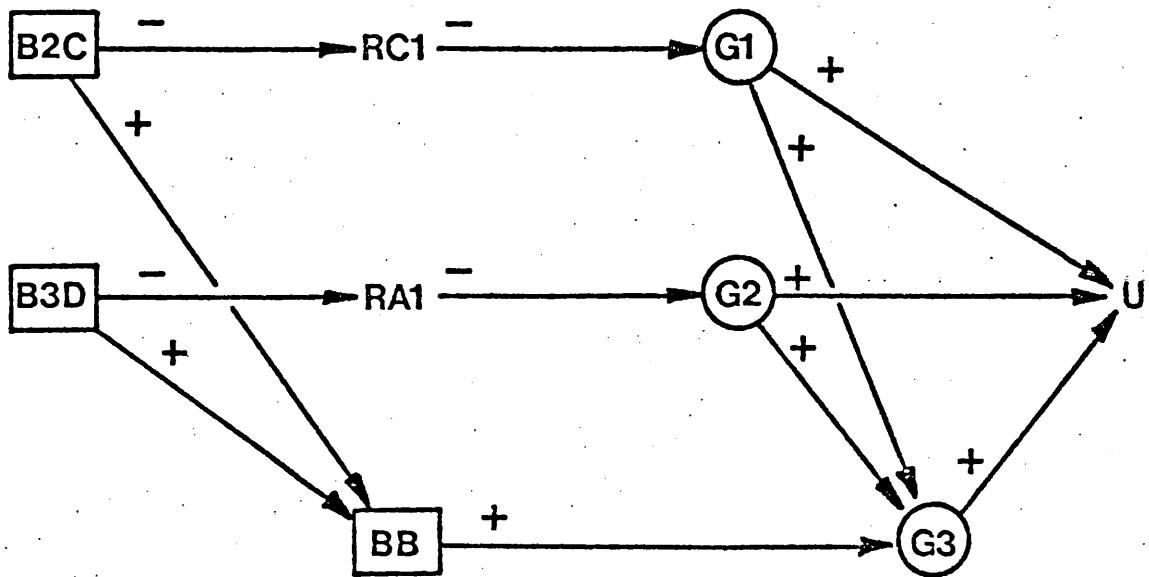


Figure 3.5. Cognitive map of Player C in Scenario 2.

SCENARIO	NUMBER OF CONCEPTS, N		NUMBER OF LINKS, L		DENSITY, D	
	1	2	1	2	1	2
A	22	21	38	33	•082	•079
B	22	22	26	29	•056	•063
C	10	8	13	12	•144	•214
PLAYER D	23	13	32	16	•063	•103
E	16	23	21	39	•038	•077
F	26	18	39	24	•060	•078
G	21	19	27	24	•064	•070

Table 3.5. Concepts, links and densities of the cognitive maps.

Defence, the player is opposed by a superior enemy force, and he has little freedom of action; on the other hand, in Scenario 2, the Advance-to-Contact, the player has the initiative and military superiority.) Some of the decision-makers, therefore, appear to be modelling their decision-making environment using a relatively constant number of concepts. This number apparently has little to do with the objective situation, and may represent some limit to the quantity of different concepts the decision-maker feels he can usefully cope with at any one time. This limit may vary quite widely between decision-makers. There is no suggestion, however, that the quality of decision-making exhibited by those players with fewer concepts is in any way inferior, at least in these games.

The number of causal links in the maps follow a similar pattern to that shown by the numbers of concepts. It is common to calculate the map density,  $D$ , defined as the number of links observed,  $L$ , divided by the maximum number of links possible between  $N$  concepts:

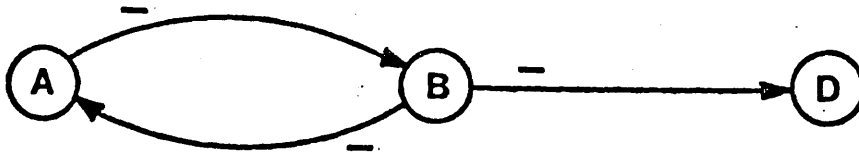
$$D = L/N(N-1).$$

Numbers of links and densities are also shown in Table 8.5. With only three exceptions, the densities of all the maps lie between 0.05 and 0.1. The three maps with the largest densities are the three smallest maps - in fact, they are exceptionally small. The concepts in these small maps tend to be of central importance to the situation, and the decision-makers acknowledge many casual relationships between them. The larger maps all contain similar central sections, but there are also many more peripheral concepts, of limited influence and affecting only a small fraction of the others. The more concepts a decision-maker uses, the less each concept is likely to have a direct



- A: Friendly holding action.
- B: Enemy advance.
- C: Enemy manoeuvre.
- D: Friendly goals.

**Case 1**



**Case 2**



Figure 3.6. Alternative views of the same objective situation.

causal effect on any other concept. As might be expected, the central sections of different players maps are frequently similar for a given scenario, with many concepts and links appearing in the same or slightly altered form in several players' maps, although this is not always the case.

#### Differences in Detailed Interpretation

Often, concepts that one player regards as fundamental do not appear on another player's map, and from time to time the same concepts can be found to be related in entirely different ways by different players. Figure 3.6 shows a specific example from Scenario 2. Case 1 illustrates the reasoning of most players: a friendly holding or delaying action will limit an enemy advance; if allowed to proceed, the advance would be detrimental to friendly goals; but a determined advance by the enemy will make the holding action more difficult. Case 2 shows the contrasting reasoning of Player G: a friendly holding action will cause the enemy to manoeuvre; this forced manoeuvre will ultimately put the enemy at a disadvantage, and so will have a positive effect on the attainment of friendly goals. The positive as opposed to negative links in Case 2 reflect the fact that only this player felt that he had the initiative in this situation, a point he makes clearly in the text. This is a particularly fine example of the way in which one decision-maker can see the same objective situation in a completely different way from another.

#### Confidence of the Players

It became clear on reading the texts that one of the main differences in the way players approached a given scenario was in

the degree of confidence they displayed in their ability to achieve their goals. A subjective rating was made by the experimenters of the confidence of each player in each scenario (Table 3.6). This rating was compared with two quantitative measures, based on the cognitive maps and the texts.

In each cognitive map, paths of relationships leading from actions, events or circumstances to either an increased goal attainment or a decreased goal attainment were identified. Each goal was considered separately. Paths leading to a particular goal either begin with an action the decision-maker can take, or with some external factor over which the decision-maker has no control, but to which he usually assigns a value. In the first case, where a path begins with an action, it is assumed that the decision-maker will always act in order to increase his goal attainment - thus such a path is always goal-increasing. In the second case, paths may be goal-increasing or goal-decreasing, depending on the value assigned to the initial factor and on the number of negative relationships on the path to the goal. The number of goal-increasing paths for all goals, divided by the total number of goal-increasing plus goal-decreasing paths, defines a goal-increasing score (Table 3.6).

This goal-increasing score, which we devised specifically for analysis of these maps, may be illustrated using the example of Player C in Scenario 1 (Figure 3.3 and Table 3.2). There are three "external" concepts which begin causal paths, and, using the text (Table 3.1), it seems reasonable to assign the values "high" to P9 and P10, and "low" to P8. We now have to identify all possible paths leading to the goal G2a. (P9, B1C, R71, G2a) is a path that leads to a goal increase;

## SCENARIO 1

PLAYER	ORDER OF PLAY	SUBJECTIVE RANK	GI SCORE	GI RANK	DS SCORE	DS RANK
A	1,2	5	.493	4	.602	4
B	2,1	4	.294	7	.456	7
C	1,2	2	.714	3	.821	1
D	2,1	6	.417	5	.560	5
E	1,2	3	.875	1	.647	3
F	2,1	7	.311	6	.549	6
G	2,1	1	.800	2	.661	2

## SCENARIO 2

PLAYER	ORDER OF PLAY	SUBJECTIVE RANK	GI SCORE	GI RANK	DS SCORE	DS RANK
A	1,2	6	.300	6	.622	4
B	2,1	3	.231	7	.660	3
C	1,2	2	1.000	1	.799	1
D	2,1	4	.778	2	.587	5
E	1,2	7	.538	5	.557	6
F	2,1	5	.769	3	.537	7
G	2,1	1	.722	4	.762	2

Table 8.6. Subjective ranks, goal-increasing (G.I.) scores and ranks, and description-speculation (D.S.) scores and ranks for the two scenarios.

on the other hand, (P8, R71, G2a) leads to a goal decrease. In fact, there are seven possible paths in total in the map, of which five are goal-increasing, and two goal-decreasing. Thus the goal-increasing score is 0.714.

Intuitively, it seems reasonable to suggest that a high goal-increasing score might be associated with a high degree of confidence, since it indicates that the player perceives a relatively high number of factors leading to increased as opposed to decreased goal attainment. The goal-increasing score measures two aspects of the decision-maker's view of his decision environment. It reflects both his perceptions of external factors affecting his goals, and also his freedom to act without external factors imposing constraints on him.

A content analysis of each text was carried out to count the numbers of words of 'description', related to the current game situation, and 'speculation', concerned with possible future developments. A description-speculation score was defined as the number of words of description divided by the total number of words of description plus speculation (Table 3.6).

From Table 3.6 it can be seen that the players who were subjectively judged as confident are characterised by both high goal-increasing scores and high description-speculation scores. The rankings based on subjective judgement, the goal-increasing scores and the description-speculation scores show a high level of agreement: using a test based on Kendall's coefficient of concordance (Kendall, 1975), the agreement is significant for Scenario 1 at the 1 per cent probability level ( $S = 212$ ), and only just fails to achieve significance for Scenario 2 at the 5 per cent probability level ( $S = 152$ , critical

value of  $S$  at 0.05 level = 157.3). Players with low description-speculation scores are those players who are thinking ahead. They consider more external factors which might adversely affect the attainment of their goals, and so they tend to have lower goal-increasing scores and they appear less confident.

#### Sequence Effects

A marked sequence effect is detectable when scores on the first and second games played by a player are compared. Goal-increasing scores and description-speculation scores are generally lower on the second occasion an individual plays, irrespective of the scenario which he plays first. This suggests that on their second encounter with the game, the players may be taking a more serious and professional approach. By the time they played for the second time, the officers may have realised that the game was very similar in many ways to the Command Post exercises with which they were familiar, that it was not an "amateur" game, and that it posed some interesting and non-trivial military problems. Subjective classification of players' confidence in the scenarios also suggests that a more cautious view is being taken on the second play (Table 8.7).

#### Dynamic Aspects

As the games develop through time, the military situations change, the players' perceptions change, and in consequence the maps are amended. The rate at which the maps must be amended varies from player to player. Table 8.8 and Figure 8.7 show the amended cognitive map for Player C in Scenario 1 in the second round of the game, which can be compared with his original map (Table 8.2 and Figure 8.3).

	FIRST SCENARIO PLAYED	SECOND SCENARIO PLAYED
CONFIDENT PLAYERS	C, E, G	C, G
LESS CONFIDENT PLAYERS	A, B, D, F	B
LEAST CONFIDENT PLAYERS		A, D, E, F

Table 3.7. Sequence effects on subjective estimates of the players' confidence.

Goal:	G2a	Defend present Task Force locations.
Blue actions:	B1C	Hold position of 1 Task Force.
	B1CS	Move supply up to 1 Task Force particularly quickly.
	B3C	Hold position of 3 Task Force.
	2B3	Move 3 Task Force to higher ground.
Red actions:	RFZ1	2 Division to move towards Cantemerle.
	R71	7 Division moves towards Langoa.
Other variables:	S8	Ability of 1 Task Force to continue fighting intensive battle.
	P8	Change in strength of enemy.
	2P9S	Ability of 1 Task Force to hold position.
	2P9N	Ability of 3 Task Force to hold position.
	2P10S	Suitability of 1 Task Force position to defence.
	2P10N	Suitability of 3 Task Force position to defence.

Table 3.8. Concepts for Player C, Scenario 1, second round.



C:1'

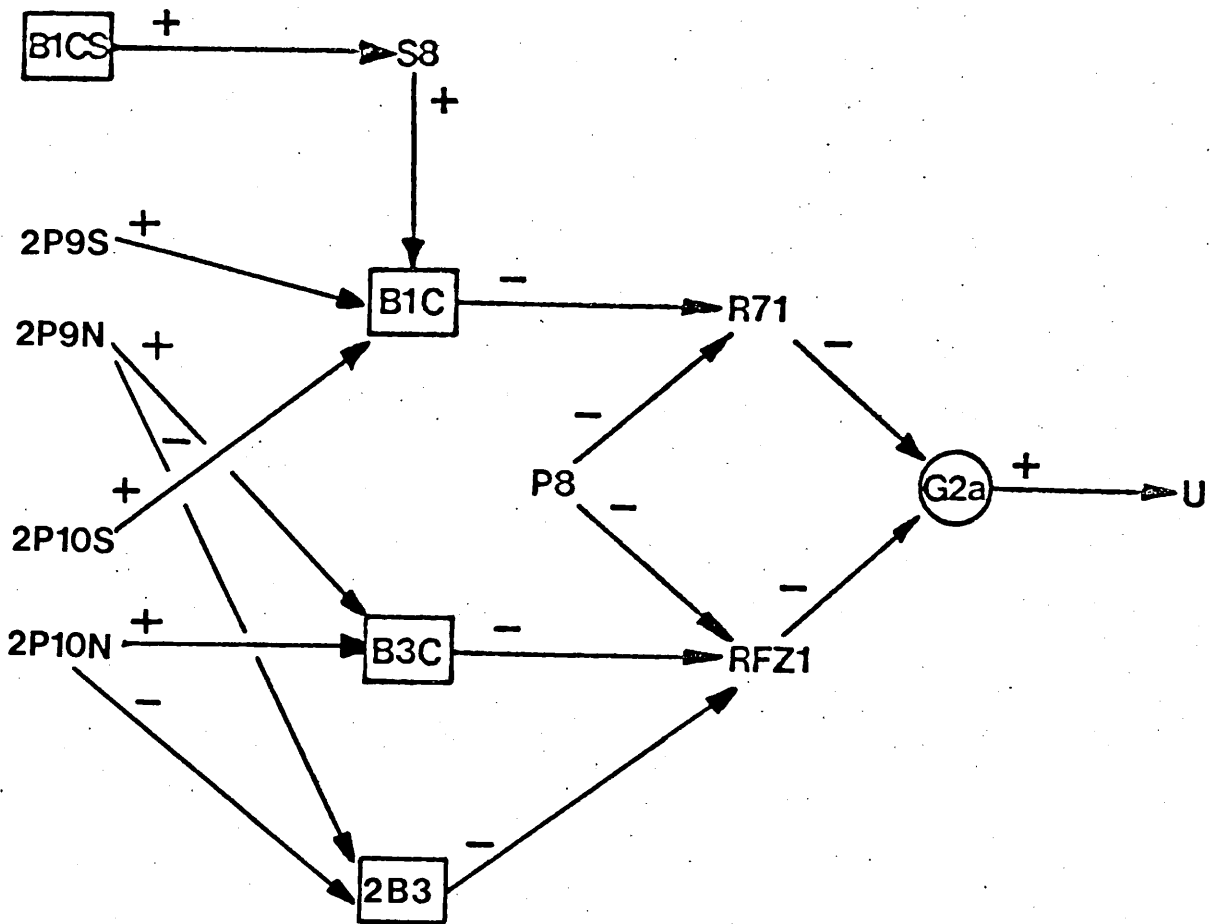


Figure 8.7. Cognitive map of Player C in Scenario 1, second round.

The way in which the cognitive maps of the players are amended indicates a tendency for some sections to become elaborated and considered in finer levels of detail, while other parts are lost from consideration. Which sections are expanded and which are lost vary for different players. Arguably, both scenarios develop towards crisis situations; the focussing by the players on subsections of the maps as time passes may be related to the general hypothesis that, in crisis, decision-makers tend to concentrate on certain aspects of their decision-making environment at the expense of others (see, for example, specific hypothesis cited by Milburn, 1972).

#### 8.5 Discussion

The results that I have presented in this chapter show dramatically that the perceptions individuals have of a common decision-making environment differ in noticeable and sometimes surprising ways. In consequence, their decision-making behaviour might be expected to differ too. The differences that have been observed fall into distinct classes, related to the size and complexity of the cognitive maps, their detailed interpretation, the players' confidence and anticipation of the future, and the way in which maps are altered as time progresses. These differences occur precisely because the decision-makers are individuals.

Do these differences, though, indicate real differences between the decision-makers? In other words, how valid is a cognitive map as a representation of real cognitive processes? This question can be approached at two levels: first, the validity of a player's assertions as a statement of his true beliefs; second, the validity of a cognitive map as a model of the mechanism of thinking.

Concerning the validity of the players' assertions, it will be recalled that an essential feature of the game methodology was the way in which most of the texts which were analysed were generated as a natural part of play. In fact, as far as the players were concerned, the game could not proceed without them. There was very little prompting by the experimenters, and thus the texts should reflect what the players themselves thought important. Additionally, the players were used to taking part in exercises of this kind, as part of their normal peacetime military role. There is no reason to believe that they engaged in deliberate deception about their true views and aims. There may, however, have been unintentional biases, in that players may tend not to articulate concepts and relationships which they have considered explicitly, but which they have assessed to be in some way irrelevant, inconsequential or "illegitimate". To detect this kind of effect, other independent evidence is needed.

Regarding the validity of cognitive mapping as a way of modelling the mechanism of decision-making, there is some doubt. Certainly, in some cases, more sophisticated treatment will be required to provide adequate models of the structure and complexity of thought. However, despite the simplicity of its approach, cognitive mapping can provide a useful analogy, enabling predictive inferences to be drawn about the outcomes of decision-making processes.

Irrespective of questions of validity, it is clear that the cognitive maps described here have elicited features of the decision-making behaviour of individuals which are both consistent and capable of interpretation. The maps of particular players show marked similarities across the two scenarios, indicating that meaningful components of an individual's decision-making process are being

detected, and these components are independent of the particular decision-making environment.

The experience of the authors, coupled with the reports of related work, indicate that cognitive mapping and allied techniques can play a number of useful roles in Operational Research. A cognitive map may be viewed as an external model of a decision process, a model which can present the complex ideas and interrelationships perceived by a decision maker in a concise, tangible and manageable form. In this framework, a cognitive map is a means of communication, not only for the transfer of perceptions and ideas between different decision-makers, but also for aiding an individual decision-maker to elucidate his own perceptions.

One particular form of this communication is the process of problem negotiation between client and analyst described by Eden et al (1979). The use of cognitive maps to facilitate communication can also be envisaged in the teaching of decision-making skills, and for assessing the ways in which decision-makers acquire problem-solving strategies as they learn about a new problem and its environment.

I also feel that cognitive mapping has a role to play in decision-making research, as presented in this chapter. The work that has been described here is essentially preliminary. It was an attempt to assess whether the combination of two comparatively new techniques - research gaming and cognitive mapping - was a worthwhile enterprise. The results have indicated that this joint approach can be a fruitful one.

It may be appropriate to end this chapter with a warning. The experiment that has been described here was limited in size and scope.

The players in the game were special, in that they were used to accepting a game world as one to be taken seriously. Nevertheless, although it cannot be claimed that the results obtained can be generalised, they do form a basis from which hypotheses can be generated for more extended experiment and testing.

## CHAPTER 9

### SUMMARY AND CONCLUSIONS

#### 9.1 Introduction

I begin this final chapter of the thesis by presenting a summary of the work described in the previous chapters. The summary will take the form of a recapitulation of the steps in the arguments and discussions which form the bulk of this account. Where appropriate, specific results and conclusions will be identified and numbered.

Following this summary, I shall discuss the work briefly, with a view to identifying where its value particularly lies, and in what directions future work might usefully proceed.

#### 9.2 A Summary

My work, as presented in this thesis, can be conveniently divided into three areas: that concerned with gaming methodology and game design (Chapters 1, 2, 3 and 4); that concerned with secondary task methods (Chapters 5 and 6); and that concerned more specifically with decision-making research (Chapters 7 and 8). This summary follows this method of division.

#### Gaming Methodology and Game Design

In Chapter 1 I identified the relative lack of techniques provided by Operational Research for coping with decision-making in complex and ill-defined problem situations. In particular, decision-making under conditions of crisis, where human-decision-making may be at its

most susceptible to failure, was identified as an area in which research might usefully be carried out. Thus, the initial formulation of the aim of the work was arrived at: to develop a suitable methodology and use it to test the validity of models of decision-making.

I proceeded to define what is meant by "crisis", and discussed hypotheses concerning decision-making behaviour in crisis - in particular, hypotheses formulated by Milburn (1972). It was, specifically, a number of Milburn's hypotheses that we initially conceived as being suitable for testing in our research. I went on to argue that a gaming approach was appropriate, research gaming being the experimental approach to decision-making.

Emphasis was made of the criticality of the processing load imposed by a game scenario, and of the nature of the relationship between stress and crisis. This led to a formulation of our initial research programme, which was based around a main programme of decision-making hypothesis testing using a research game which we would design and develop. Against this main programme, two subsidiary programmes would be developed. The first was the development of a secondary task method for measuring mental processing load. The second was the development of an understanding of the stress aspect of crisis (ultimately this second subsidiary programme was not pursued).

In Chapter 2 I began to collect and develop ideas about gaming that would ultimately lead to the development of our own research game. I considered Bowen's (1978) classification of games by purpose, the major value of which is to emphasise the importance of a clear idea of the intended aim of a game during its design. I then turned to

Bowen's (1978) classification of games by structure and showed that a desirable research game structure would combine the rich interaction property of the three-room game with the replicability property of the single-decision game. If deception is resorted to, such a game may be possible. Stated more formally:

A research game requires the single-decision structure  
 $B(P,N,C) \cdot R(A,N,C)$ , but should appear to the player to  
have the three-room structure  $B(P,M,C) \cdot R(P,M,C)$ . (1)

I described Sharp and Dando's (1977) Intelligence Man gaming system, which partially meets this requirement, before describing Cooper's (1978, 1979) Superior Commander system, which we were to use in our own work:

The Superior Commander gaming system permits the running  
of a game with structure  $B(P,N,C) \cdot R(A,N,C)$ , which appears  
to the player to have structure  $B(P,M,C) \cdot R(P,M,C)$ . (2)

Chapter 3 was concerned with the content rather than the structure of research games. Most of the discussion in this chapter is based on our early gaming experiments. We were able to identify six requirements for a research game, and three requirements for the game methodology employed:

A research game requires: (a) simple, easily learnt, rules.  
(b) relatively fast play.  
(c) easy recordability.



- (d) a strategic aspect.
- (e) to be interesting.
- (f) natural extraction of experimental  
data. (3)

A research gaming methodology requires:

- (a) replicability.
- (b) complexity.
- (c) emphasis on player's game  
participant role. (4)

It then presented the Organisational Control Game, which we designed with the above requirements in mind.

The Organisational Control Game, played according to the Superior Commander gaming system, meets in full the requirements for a research game and gaming methodology listed in (3) and (4). (5)

Chapter 4 consisted of a detailed description of the Organisational Control Game. Important aspects of the design of the game were emphasised: in particular, the two completely different game scenarios, one depicting a defence, the other an advance-to-contact. Points concerned with the practical running of the game were also mentioned.

The four chapters described above contain an account of considerable experience in the design, construction, testing and development of a research game, and may be of particular interest, therefore, to prospective game designers.

## Secondary Task Methods

Chapter 5 was a lengthy literature review devoted to consideration of secondary task methods for assessing mental processing load. In the course of the chapter the capacity model of human information processing was described. There are two distinct types of secondary task method - that in which the secondary task is used as a loading task, and that in which it is used as a subsidiary measuring task. Both methods were clearly explained.

The review covered applications to the study of such cognitive processes as memory and learning, as well as to the characteristics of tasks such as vehicle driving. However, it became clear that all the tasks that had been studied can be categorised as tasks in which, to a large extent, response models appropriate to the task decisions are already available to the subject:

Secondary task methods have not been applied to complex problem-solving tasks in which straight-forward mental models providing appropriate immediate responses are not available to the subject. (6)

Chapter 6 described an experiment we conducted to determine whether the gap could be filled. As the problem-solving task we used a series of chess problems, and the experiment attempted to rank these problems in order of complexity by using a mental arithmetic secondary task with a number of subjects. The experiment showed that this was a workable means of ranking the problems.

Secondary task methods can be used to measure the processing load associated with complex problem-solving tasks.

(7)

Thus, we have shown the viability of a technique for measuring the processing load imposed by a variety of tasks that might be of interest to Operational Research scientists. It might, for example, be applicable to the assessment of the complexity of game scenarios, as we had originally intended. Operational Research scientists may also find several of the applications described in the literature review chapter of interest.

#### Decision-Making Research

Chapter 7 began by discussing a series of games played with eight retired military officers, all of whom held senior positions in United Kingdom local government at the time of play. In particular, these players were compared with earlier student players, the two groups being classified as "experienced" and "inexperienced" players. Three differences between the groups were observed:

Experienced players made use of their previous experience, injecting their own ideas, assumptions and perceptions into the game; inexperienced players tended to accept the game precisely as presented.

(8)

Experienced players exhibited greater knowledge, performing equally well on both scenarios; inexperienced players managed advance-to-contact quite well, but had little idea about the way to conduct a defence.

(9)

Experienced players viewed problems as complex and ill-defined; inexperienced players made rather cursory assessments, and looked for unique "best" solutions. (10)

From this we drew the obvious but important conclusion:

Students do not behave in the same way in a game environment as experienced decision-makers. (11)

At the same time, we were developing some doubts about the immediate course of our research. We began to feel that before we could begin to consider a research programme such as that originally formulated, we would need a technique or language for describing explicitly the subjective environment of the decision-maker. This can be stated:

In decision-making research, it is essential to develop a means of describing the environment of the decision-maker in his own terms, in a way that provides a model of that subjective environment rich enough to be discussed and studied. (12)

In our final series of games, played with army majors, we attempted to follow up this belief. We hoped to use cognitive mapping as a way of describing subjective decision-making environments. The cognitive mapping analysis of the games was described in Chapter 3. A number of results were presented.

The cognitive maps of the players in the game show marked differences between the players, yet, for any player, marked similarities across the scenarios. (13)

The number of concepts and relationships between concepts in a cognitive map of a situation varies widely from player to player. (14)

There is a suggestion that some players may employ a relatively constant number of concepts to describe different situations. (15)

Cognitive maps show up differences in detailed interpretation of the decision-making environment by different players. (16)

Confident players are characterised by high goal-increasing scores in their cognitive maps, and high description-speculation scores in their game texts. (17)

Two general conclusions can be drawn as a result of the cognitive map analysis.

Cognitive mapping may be a useful way of revealing the individual decision-making characteristics of subjects, particularly regarding their perceptions of their decision-making environments. (18)

The combination of the two comparatively new techniques of research gaming and cognitive mapping can be a fruitful approach to the study of decision-making. (19)

Chapters 7 and 8, therefore, can be seen to develop ideas about decision-making research, and in particular were a plea to put the horse before the cart: before attempting to study decision-making, in prescriptive fashion, we should ensure that we are in a position to adequately discuss and describe that which we wish to study. Having made this plea, I then proceeded to do something to answer it, I believe with some degree of success.

It is worth making two points concerning the cognitive map analysis. The first point is that the results that have been presented derive from an experiment that was limited in both size and scope. Results (13) to (17) above were obtained in the context of a military decision-making game played by a relatively small number of military decision-makers. Therefore there is a clear need to examine whether the cognitive mapping technique gives similar results when applied to decision-making by further subjects in both similar contexts and different contexts. In this way we would be able to gain some idea of both the generality and area of applicability of these results. Since there is a clearly subjective element in the analysis procedure, it would be desirable that such work be carried out by a number of different workers. This would give some indication of the effects due to the particular perceptions of the analyst.

The second point concerns the degree to which the cognitive map analysis technique provides a valid representation of the decision-making process and the characteristics of individual decision-makers. Conclusions (18) and (19) have been deliberately formulated in rather tentative fashion because it is, at present, not certain how good this representation is. Ideally we would require a completely independent means of identifying and measuring the characteristics of decision-

making in order to confirm that these are genuinely reflected in a cognitive map representation. However, because decision-making takes place in subjective worlds, we can only realistically hope to work with representations of decision-making processes. We should therefore be looking for consistency between representations, and also consistency within the cognitive mapping representation. The cognitive map analysis I have presented, taken together with other cognitive mapping work (see Chapter 8), shows a fair degree of these types of consistency: results are both self-consistent and show some correspondence with external notions of decision-making. I believe cognitive mapping does provide a good and useful representation of decision-making - although I note that every representation must have its limits. Further work, demonstrating features of consistency, will enable us to strengthen the tentative conclusions I have presented here.

### 9.3 Final Discussion

In this section I identify what I see as the main value of the work described in this thesis. I also mention future directions in which work might usefully proceed. Once again, it is convenient to split the work into three sections.

Beginning with game methodology and game design, the work in this thesis can be seen as part of a general rethinking that has taken place in recent years concerning the value of gaming. Game designers have become more critical of their work, and are increasingly concerned with defining their objectives more accurately and ensuring that a research game measures what it is intended to measure. Bowen's (1978) work, an almost entirely theoretical monograph on the subject of games, can be seen as indicative of this trend.

On the negative side, the trend has resulted in an increasing number of workers becoming disillusioned with gaming as a research tool. Relatively little of this disillusionment has reached the

literature, but most of us in gaming have heard criticisms concerning the validity of any results we may offer. This stems from the wide realisation that a successful gaming experiment cannot simply be approached as one might approach an experiment on the behaviour of laboratory rats. Questions of the player's subjective role in the game, and the context in which the game is presented to him are critical and cannot be ignored. The failure to take these issues into account quite rightly calls into question the validity of much early gaming work.

However, I believe that the work in this thesis shows that these difficult issues can be faced and dealt with to a most encouraging extent. As I believe the thesis demonstrates, in the course of our research programme we spent a great deal of time considering what it was we wanted to look at, how we could look at it, and what we could and could not measure. Issues such as the control of a game and the subjective role of the player have been met squarely and dealt with. These considerations led us to develop the Organisational Control Game and Superior Commander system. This combination constitutes a worthwhile game which I believe can be used as an effective research tool - indeed, that this is so is demonstrated practically in the final chapters of this thesis. Furthermore, the Superior Commander system could clearly be used in a variety of contexts as the basis for an effective research game.

Thus the success of the Organisational Control Game and Superior Commander system is seen as a demonstration that even with a far more critical approach to gaming, it is possible to carry out a valid research gaming programme. I believe that this type of approach will continue to be of value in the field of decision-making research.



Turning now to the secondary task work, the experiment described in Chapter 6 very clearly demonstrates the applicability of secondary task techniques to the assessment of decision-making loads imposed by complex problem-solving and decision-making tasks. I have already noted that this technique could be used to assess the complexity of game scenarios. A variety of other applications are likely to be found in other areas: for example, the design of effective decision-making systems. At the same time, the technique as used for other types of task, as described in the literature review in Chapter 5, may be of value to the Operational Research scientist who happens to find himself faced with certain types of ergonomic problem.

As far as decision-making research is concerned, the value of the work lies in two areas. To begin with, I regard the identification of the need for a means of describing the individual's subjective decision-making environment as an important feature of the work. Once again, this can be seen as part of a recent trend - in this case, a trend to question initial assumptions, particularly concerning the perceptions of individual actors in a situation. Clearly, if such matters are now seen to be of central importance, rather than an implicit framework on which research is based, researchers must be certain that they are able to adequately describe these matters. Yet researchers have tended to rush into decision-making research without fully establishing this essential basis of their research. The formulation of our research programme in Chapter 1 is to some extent an example of this: consequently, as I have shown in Chapter 7, it was quite radically revised.

The second area of value is the cognitive mapping work described in Chapter 3. I have already written at length in that chapter of the

value of the technique in Operational Research. Here I wish only to point out the value of the cognitive mapping technique as a means for analysing and investigating decision-making of various kinds. This point is demonstrated in this thesis in the context of research gaming, and others (Eden et al, 1979) have demonstrated it in the context of practical assistance to decision-makers.

It is in this area that I feel that much interesting future work lies. In conjunction with a research game, a cognitive map approach might be adapted to begin to study the type of hypothesis in which we were originally interested. Hypotheses can be formulated in cognitive mapping terms, and in this form can be tested. For example, in Chapter 1 we considered two of Milburn's (1972) hypotheses:

Hypothesis 13 (p.274): Crises increase a tendency towards rigidity of perception and thought.

Hypothesis 17 (p.275): In a crisis one's own alternatives appear to contract while the other side's options seem to grow.

These hypotheses suggest various new hypotheses that can be expressed in terms of cognitive maps. For example, Hypothesis 13 gives rise to:

In a crisis, the cognitive map of a decision-maker exhibits relatively little development and elaboration as time progresses.

In a crisis, new data tends to be ignored when its use would tend to require the revision of parts of an established cognitive map of the decision-maker.

Hypothesis 17 suggests:

In a crisis, the number of concepts representing actions available to the decision-maker in his cognitive map will decrease, whereas the number of concepts representing actions available to opponents will increase.

In a crisis, the goal-increasing score of the cognitive map of a decision-maker tends to decrease.

Of course, this type of formulation does not automatically grant a straightforward hypothesis-testing experiment. In particular, unless we choose to use some stress definition of crisis, what actually constitutes a state of crisis may remain poorly defined. Possibly a cognitive map definition might be formulated. We recall, from the definitions in Chapter 1, the importance of threat, time and surprise in the definition of crisis. Thus, we might describe a decision-maker as being in a state of crisis when a large amount of new data causes a sudden, possibly catastrophic, change in his cognitive map (surprise), and when the new cognitive map explicitly indicates lack of time and high threat as being perceived as central to the situation in which the decision-maker is placed.

Even if the specific suggestions presented here turn out to be of little consequence, the fundamental message of the cognitive mapping

work should not be lost sight of: decision-makers are individuals. Any science of decision-making must accept this as a fundamental feature, perhaps the most important feature, of the raw material which it is intended to account for and elucidate. Techniques which are intended to develop this fundamental point, by understanding decision-making within the context of the decision-maker's own subjective world, are essential. This thesis has, in its entirety, described a journey towards this realisation. This may be, ultimately, the most important point which is made in this thesis.

## REFERENCES

- A.R. ABDEL-KHALIK (1974) On the efficiency of subject surrogation in accounting research. Accounting Review, 49, 743-750.
- R.L. ACKOFF (1979) The future of Operational Research is past. Journal of the Operational Research Society, 30, 93-104.
- J.A. ADAMS & R.W. CHAMBERS (1962) Response to simultaneous stimulation of two sense modalities. Journal of Experimental Psychology, 63, 198-206.
- B. ALPERT (1967) Non-businessmen as surrogates for businessmen in behavioural experiments. Journal of Business, 5, 203-207.
- AVALON HILL (1970) Panzer Blitz (game). Avalon Hill Game Company, Baltimore, Maryland.
- AVALON HILL (1974) Panzer Leader (game). Avalon Hill Game Company, Baltimore, Maryland.
- AVALON HILL (1976) Diplomacy (game). Avalon Hill Game Company, Baltimore, Maryland.
- R. AXELROD, editor (1976a) Structure of Decision. Princeton University Press, Princeton, New Jersey.
- R. AXELROD (1976b) Deceision for neoimperialism: the deliberations of the British Eastern Committee in 1918. In R. AXELROD, editor (1976a) Structure of Decision. Princeton University Press, Princeton, New Jersey.

- H.P. BAHRICK, P.M. FITTS & R.E. RANKIN (1952) Effect of incentives upon reactions to peripheral stimuli. Journal of Experimental Psychology, 44, 400-406.
- H.P. BAHRICK, M. NOBLE & P.M. FITTS (1954) Extra-task performance as a measure of learning a primary task. Journal of Experimental Psychology, 48, 298-302.
- H.P. BAHRICK & C. SHELLY (1958) Time sharing as an index of automatization. Journal of Experimental Psychology, 56, 288-293.
- C.E. BAKER, R.C. WYLIE & R.M. GAGNÉ (1951) The effects of an interfering task on the learning of a complex motor skill. Journal of Experimental Psychology, 41, 1-9.
- D. BANNISTER & F. FRANSELLA (1971) Inquiring Man. Penguin Education, Harmondsworth.
- B.P. BARNES (1976) Pick of the Best Chess Problems. Elliot Right Right Way Books, Kingswood, Surrey.
- P.G. BENNETT (1980) Hypergames: developing a model of conflict. Futures, 12, 489-507.
- J.G. BIRNBERG & R. NATH (1968) Laboratory experimentation in accounting research. Accounting Review, 43, 38-45.

- D.H. BOGGS & J.R. SIMON (1968) Differential effect of noise on tasks of varying complexity. Journal of Applied Psychology, 52, 148-153.
- G.M. BONHAM, D. HERADSTVEIT, O. NARVESEN & M.J. SHAPIRO (1978) A cognitive model of decision-making: application to Norwegian oil policy. Cooperation and Conflict, 13, 93-108.
- G.M. BONHAM, M.J. SHAPIRO & G.Z. NOZICK (1976) A cognitive process model of foreign policy decision-making. Simulation and Games, 7, 123-152.
- G.M. BONHAM, M.J. SHAPIRO & T.L. TRUMBLE (1979) The October War. International Studies Quarterly, 23, 3-44.
- K.C. BOWEN (1978) Research Games. Taylor & Francis, London.
- G.D. BREWER (1978) Scientific gaming: the development and use of free-form scenarios. Simulation and Games, 9, 309-338.
- G.D. BREWER & M. SHUBIK (1979) The War Game. Harvard University Press, Cambridge, Massachusetts.
- D.E. BROADBENT (1957) Effects of noise on behavior. In C.M. HARRIS, editor (1957) Handbook of Noise Control. McGraw-Hill, New York.
- D.E. BROADBENT (1956) Listening between and during practised auditory distractions. British Journal of Psychology, 47, 51-60.

D.E. BROADBENT (1958) Perception and Communication. Pergamon Press, London.

D.E. BROADBENT (1971) Decision and Stress. Academic Press, London.

D.E. BROADBENT & A. HERON (1962) Effects of a subsidiary task on performance involving immediate memory by younger and older men. British Journal of Psychology, 53, 189-198.

I.D. BROWN (1962a) Measuring the 'spare mental capacity' of car drivers by a subsidiary auditory task. Ergonomics, 5, 247-250.

I.D. BROWN (1962b) Studies of component movements, consistency and spare capacity of car drivers. Annals of Occupational Hygiene, 5, 131-143.

I.D. BROWN (1964) The measurement of perceptual load and reserve capacity. Transactions of the Association of Industrial Medical Officers, 14, 44-49.

I.D. BROWN (1966) Subjective and objective comparisons of successful and unsuccessful trainee drivers. Ergonomics, 9, 49-56.

I.D. BROWN (1968) Some alternative methods of predicting performance among professional drivers in training. Ergonomics, 11, 13-21.

I.D. BROWN (1978) Dual task methods of assessing work-load. Ergonomics, 21, 221-224.



- I.D. BROWN & E.C. POULTON (1961) Measuring the spare 'mental capacity' of car drivers by a subsidiary task. Ergonomics, 4, 35-40.
- I.D. BROWN, D.C. SIMMONDS & A.H. TICKNER (1967) Measure of control skills, vigilance, and performance on a subsidiary task during 12 hours of car driving. Ergonomics, 10, 665-673.
- S.L. CHOW & B.B. MURDOCK (1975) The effect of a subsidiary task on iconic memory. Memory and Cognition, 3, 678-688.
- J.K.S. CLAYTON (1981) Viewpoint. Journal of the Operational Research Society, 32, 1150-1152.
- D.F. COOPER (1977) Decisions in crisis. Working paper, Department of Mathematics, Royal Holloway College, University of London.
- D.F. COOPER (1978) On the design and control of crisis games. Omega, 6, 460-461.
- D.F. COOPER (1979) The Superior Commander: a methodology for the control of crisis games. Journal of the Operational Research Society, 30, 529-537.
- D.F. COOPER & J.H. KLEIN (1980) Board wargames for decision-making research. European Journal of Operational Research, 5, 36-41.

- D.F. COOPER, J.H. KLEIN & M.R.C. McDOWELL (1978) Crisis gaming progress report. Department of Mathematics, Royal Holloway College, University of London.
- D.F. COOPER, J.H. KLEIN, M.R.C. McDOWELL & P.V. JOHNSON (1980) The development of a research game. Journal of the Operational Research Society, 31, 191-193.
- R.M. COPELAND, A.J. FRANCIA & R.H. STRAWSER (1973) Students as subjects in behavioral business research. Accounting Review, 48, 365-372.
- L.R. CREAMER (1963) Event uncertainty, psychological refractory period, and human data processing. Journal of Experimental Psychology, 66, 187-194.
- M.R. DANDO (1981) HOT SEAT - a critical review of gaming in a civil defence training exercise. Journal of the Operational Research Society, 32, 1146-1156.
- M.R. DANDO & K.L. BROWN (1981) The development of an industrial relations teaching game. Journal of the Operational Research Society, 32, 625-637.
- M.R. DANDO, A. DEFRENNE & R.G. SHARP (1977) Could OR be a science? Omega, 5, 89-92.
- D.W. DANIEL (1980) What influences a decision? Some results from a highly controlled defence game. Omega, 8, 409-419.

- R. DAVIS (1957) The human operator as a single channel information system. Quarterly Journal of Experimental Psychology, 9, 119-129.
- R. DAVIS (1962) Choice reaction times and the theory of intermittency in human performance. Quarterly Journal of Experimental Psychology, 14, 157-166.
- S. DIMOND & G. BEAUMONT (1971) Use of two cerebral hemispheres to increase brain capacity. Nature, 232, 270-271.
- R. DOOST & M.T. TURVEY (1971) Iconic memory and central processing capacity. Perception and Psychophysics, 9, 269-274.
- E. DUFFY (1962) Activation and Behaviour. Wiley, New York.
- C. EDEN & S. JONES (1980) Publish or perish? A case study. Journal of the Operational Research Society, 31, 131-139.
- C. EDEN, S. JONES & D. SIMS (1979) Thinking in Organizations. Macmillan, London.
- A. ELITHORN & C. LAWRENCE (1955) Central inhibition - some refractory observations. Quarterly Journal of Experimental Psychology, 7, 116-127.
- J.R. EMSHOFF (1971) Analysis of Behavioral Systems. Macmillan, New York.
- B. ENO & P. SCHMIDT (1979) Oblique Strategies (set of cards). Third edition.

- H.J. EYSENCK (1973) Eysenck on Extraversion. Crosby Lockwood Staples, London.
- J.M. FINKELMAN & D.C. GLASS (1970) Reappraisal of the relationship between noise and human performance by means of a subsidiary task measure. Journal of Applied Psychology, 54, 211-213.
- S. FISHER (1975a) The microstructure of dual task interaction. 1. The patterning of main-task responses within secondary task intervals. Perception, 4, 267-290.
- S. FISHER (1975b) The microstructure of dual task interaction. 2. The effect of task instructions on attentional allocation and a model of attention-switching. Perception, 4, 459-474.
- S. FISHER (1977) The microstructure of dual task interaction. 3. Incompatibility and attention switching. Perception, 6, 467-477.
- J.E. FLEMING (1969) Managers as subjects in business decision research. Academy of Management Journal, 12, 59-66.
- W.D. GARVEY & F.V. TAYLOR (1959) Interactions among operator variables, system dynamics, and task-induced stress. Journal of Applied Psychology, 43, 79-85.
- G.I. GIBBS, editor (1974) Handbook of Games and Simulation Exercises. Spon, London.

- J.I. HARRIS (1974) Some concepts of a conflict game. DOAE  
Memorandum M7420, Defence Operational Analysis  
Establishment, West Byfleet.
- J.A. HART (1976) Comparative cognition: politics of international  
control of the oceans. In R. AXELROD, editor (1976a)  
Structure of Decision. Princeton University  
Press, Princeton, New Jersey.
- J.A. HART (1977) Cognitive maps of three Latin American policy  
makers. World Politics, 30, 115-140.
- J.A. HART (1978) Cognitive maps of presidential documents. Paper  
presented at the 86th American Psychological Association  
Annual Convention, Toronto, Ontario, Canada,  
August 28 - September 1, 1978.
- D.A. HARTLEY, P.V. JOHNSON, A. FITZSIMONS, J. LOVELL, B. CHIPPENDALE  
& J.K.S. CLAYTON (1979) A case study on the development  
of the home defence training game HOT SEAT.  
Journal of the Operational Research  
Society, 30, 861-871.
- D. HERADSTVEIT & O. NARVESEN (1978) Psychological constraints on  
decision making. Cooperation and Conflict, 13, 77-92.
- C.F. HERMANN (1972a) Some issues in the study of international crisis.  
In C.F. HERMANN, editor (1972b) International Crises:  
Insights from Behavioral Research. Free Press,  
New York.

- C.F. HERMANN, editor (1972b) International Crises: Insights from Behavioral Research. Free Press, New York.
- J.H. HOLLAND & J.S. REITMAN (1978) Cognitive systems based on adaptive algorithms. In D.A. WATERMAN & F. HAYES-ROTH, editors (1978) Pattern-Directed Inference Systems. Academic Press, New York.
- L. HUDSON (1972) The Cult of the Fact. Jonathan Cape, London.
- C.S. HUXHAM, P.G. BENNETT, M.V. LOZOWSKI & M.R. DANDO (1981) The study of human decision-making: a cautionary tale from the world of experimental gaming. Journal of the Operational Research Society, 32, 173-185.
- P.V. JOHNSON & D. HARTLEY (1981) Viewpoint. Journal of the Operational Research Society, 32, 1152-1155.
- W.A. JOHNSON, S.N. GREENBERG, R.P. FISHER & D.W. MARTIN (1970) Divided attention: a vehicle for monitoring memory processes. Journal of Experimental Psychology, 83, 164-171
- A.J. JONES (1980) Game Theory: Mathematical Models of Conflict. Ellis Horwood, U.K.
- S. JONES & C. EDEN (1981) O.R. in the community. Journal of the Operational Research Society, 32, 335-345.

- D. KAHNEMAN (1973) Attention and Effort. Prentice-Hall, Englewood Cliffs, New Jersey.
- J.W.H. KALSBECK (1971) Sinus arrhythmia and the dual task method in measuring mental load. In W.T. SINGLETON, J.G. FOX & D. WHITFIELD, editors (1971) Measurement of Man at Work. Taylor and Francis, London.
- J.W.H. KALSBECK & R.N. SYKES (1967) Objective measurement of mental load. Acta Psychologica, 27, 253-261.
- S.W. KEELE (1973) Attention and human performance. Goodyear, Pacific Palisades, California.
- R.T. KELLY & D.W. MARTIN (1974) Memory for random shapes: a dual-task analysis. Journal of Experimental Psychology, 103, 224-229.
- Sir M. KENDALL (1975) Rank Correlation Methods. Fourth edition. Griffin, London.
- B. KERR (1973) Processing demands during mental operations. Memory and Cognition, 1, 401-412.
- J.H. KLEIN & D.F. COOPER (1979) Chess problems for secondary task assessment. Set C. Department of Mathematics, Royal Holloway College, University of London.
- J.H. KLEIN & D.F. COOPER (1981) Assessing problem complexity. European Journal of Operational Research, 6, 243-247.

- J.H. KLEIN & D.F. COOPER (1982) Cognitive maps of decision-makers in a complex game. Journal of the Operational Research Society, 33, 63-71.
- J.H. KLEIN & M.R.C. McDOWELL (1981) HOT SPOT: a civil defence training game for senior officers of metropolitan boroughs. Department of Mathematics, Royal Holloway College, University of London.
- W.B. KNOWLES (1963) Operator loading tasks. Human Factors, 5, 155-161.
- W.B. KNOWLES & D.J. ROSE (1963) Manned lunar landing simulation. Proceedings of the National Winter Convention on Military Electronics, Volume 2, I.E.E.E., California.
- M. LAVER (1979) Playing Politics. Penguin, Harmondsworth.
- A. LEVI & P.E. TETLOCK (1980) A cognitive analysis of Japan's 1941 decision for war. Journal of Conflict Resolution, 24, 195-211.
- D.C.F. LLOYD (1978) An introduction to business games. Industrial and Commercial Training, January 1978, 11-18.
- R. MANDEL (1977) Political gaming and foreign policy making during crises. World Politics, 29, 610-625.



- D.W. MARTIN (1970) Residual processing capacity during verbal organization in memory. Journal of Verbal Learning and Verbal Behavior, 9, 391-397.
- D.W. MARTIN, P.T. MARSTON & R.T. KELLY (1973) Measurement of organizational processes within memory stages. Journal of Experimental Psychology, 98, 387-395.
- J. MAYNARD SMITH (1974) The theory of games and the evolution of animal conflicts. Journal of Theoretical Biology, 47, 209-221.
- M.R.C. McDOWELL, D.F. COOPER, J.H. KLEIN & P.V. JOHNSON (1979a) Crisis gaming progress report II. Department of Mathematics, Royal Holloway College, University of London.
- M.R.C. McDOWELL, D.F. COOPER, J.H. KLEIN & P.V. JOHNSON (1979b) Crisis gaming final report. Department of Mathematics, Royal Holloway College, University of London.
- D. McNICOL (1971) Short-term recall of messages containing repeated items. Australian Journal of Psychology, 23, 241-252.
- J.A. MICHON (1966) Tapping regularity as a measure of perceptual motor load. Ergonomics, 9, 401-412.
- T.W. MILBURN (1972) The management of crisis. In C.F. HERMANN, editor (1972b) International Crises: Insights from Behavioral Research. Free Press, New York.

- M. MITSUDA (1968) Effects of a subsidiary task on backward recall. Journal of Verbal Learning and Verbal Behavior, 7, 722-725.
- U. NEISSER (1976) Cognition and Reality. Freeman, San Francisco.
- U. NEISSER & R. BECKLEN (1975) Selective looking: attending to visually specified events. Cognitive Psychology, 7, 480-494.
- M. NOBLE, D. TRUMBO & F. FOWLER (1967) Further evidence on secondary task interference in tracking. Journal of Experimental Psychology, 73, 146-149.
- N. PALMER (1977) The Comprehensive Guide to Board Wargaming. Arthur Barker, London.
- J. PARROT (1971) The measurement of stress and strain. In W.T. SINGLETON, J.G. FOX & D. WHITFIELD, editors (1971) Measurement of Man at Work. Taylor and Francis, London.
- E.C. POULTON (1958) Measuring the order of difficulty of visual-motor tasks. Ergonomics, 1, 234-239.
- E.C. POULTON (1971) Skilled performance and stress. In P.B. WARR, editor (1971) Psychology at Work. Penguin Education, Harmondsworth.

D.G. PRUITT & M.J. KIMMEL (1977) Twenty years of experimental gaming: critique, synthesis, and suggestions for the future.

Annual Review of Psychology, 28, 363-392.

S.W. QUENAUULT (1968) Task capability while driving. RRL Report LR-166, Road Research Laboratory, Crowthorne, U.K.

R.H.C. REPORTS (1980) RHC/MOR/GLC1 - RHC/MOR/GLC5. Department of Mathematics, Royal Holloway College, University of London.

R.M.C.S. REPORTS (1977) OR/C/1 - OR/C/7. Department of Management Sciences, Royal Military College of Science, Shrivenham.

K.J. RADFORD (1975) Applications of metagame theory in managerial decision making. Omega, 3, 303-312.

K.J. RADFORD (1978) Decision-making in a turbulent environment. Journal of the Operational Research Society, 29, 667-682.

J.A. ROBINSON (1972) Crisis: an appraisal of concepts and theories. In C.F. HERMANN, editor (1972b) International Crises: Insights from Behavioral Research. Free press, New York.

J.M. ROLFE (1971) The secondary task as a measure of mental load. In W.T. SINGLETON, J.G. FOX & D. WHITFIELD, editors (1971) Measurement of Man at Work. Taylor and Francis, London.

- L.L. ROOS & R.I. HALL (1980) Influence diagrams and organisational power. Administrative Science Quarterly, 25, 57-71.
- J.F. SCHOUTEN, J.W.H. KALSBEK & F.F. LEOPOLD (1962) On the evaluation of perceptual and mental load. Ergonomics, 5, 251-260.
- R.G. SHARP & M.R. DANDO (1977) The Intelligence Man system - an approach to assisting managers in complex conflicts. European Journal of Operational Research, 1, 257-288.
- R.G. SHARP & M.R. DANDO (1979) Decision resource management and practical intervention in conflicts - a model and its implications. European Journal of Operational Research, 3, 283-295.
- M.J. SHAPIRO & G.M. BONHAM (1973) Cognitive process and foreign policy decision-making. International Studies Quarterly, 17, 147-174.
- R.W. SHEPHARD (1963) War gaming as a technique in the study of Operational Research problems. Operational Research Quarterly, 14, 119-130.
- M. SHUBIK (1975a) Games for Society, Business and War. Elsevier, New York.
- M. SHUBIK (1975b) The Uses and Methods of Gaming. Elsevier, New York.

- M. SHUBIK & G.D. BREWER (1972a) Models, Simulations and Games - a Survey. R-1060-ARPA/RC, RAND Corporation, Santa Monica, California.
- M. SHUBIK & G.D. BREWER (1972b) Reviews of Selected Books of Gaming and Simulation. R-732-ARPA, RAND Corporation, Santa Monica, California.
- M. SHUBIK, G.D. BREWER & E. SAVAGE (1972) The Literature of Gaming, Simulation, and Model-Building: Index and Critical Abstracts. R-620-ARPA, RAND Corporation, Santa Monica, California.
- S. SIEGEL (1956) Nonparametric Statistics for the Behavioral Sciences. McGraw-Hill Kogakusha, Tokyo.
- I. SILVERMAN (1977) The Human Subject in the Psychological Laboratory. Pergamon Press, New York.
- R.R. SKEMP (1979) Intelligence, Learning and Action: A Foundation for Theory and Practice in Education. Wiley, Chichester.
- P. STAGER (1970) Effects of information load and instructions on performance in a complex task. Journal of Experimental Psychology, 84, 123-130.
- P. STAGER & G.J. LABBS (1977) The effect of divided attention on probe reaction time in multiple-task performance. Canadian Journal of Psychology, 31, 174-183.

- P. STAGER & P. MUTER (1971) Instructions and information processing in a complex task. Journal of Experimental Psychology, 87, 291-294.
- P. STAGER & K. ZUFELT (1972) Dual-task method in determining load differences. Journal of Experimental Psychology, 94, 113-115.
- P.A. STORR, J. WENNEL, D.F. COOPER & M.R.C. McDOWELL (1979) A microprocessor based system for traffic data collection. Traffic Engineering and Control, 20, 156-158.
- STRATEGY AND TACTICS (1977) Wargame Design. Strategy and Tactics Staff Study No. 2, Simulations Publications Incorporated, New York.
- E.V. SULLIVAN & M.T. TURVEY (1972) Short-term retention of tactile stimulation. Quarterly Journal of Experimental Psychology, 24, 253-261.
- TIME OUT (1981) Rumbles. No. 563, 30 January 1981, 11.
- D. TRUMBO & M. NOBLE (1972) Response uncertainty in dual-task performance. Organizational Behavior and Human Performance, 7, 203-215.
- P.A. VROON & A.G. VROON (1973) Tapping rate and expectancy in simple reaction time tasks. Journal of Experimental Psychology, 98, 85-90.
- A.W. WATTS (1957) The Way of Zen. Penguin, Harmondsworth.

A.T. WELFORD (1952) The 'psychological refractory period' and the timing of high-speed performance - a review and a theory.

British Journal of Psychology, 43, 2-19.

R.M. WHITE JR, J.R. WOODS & W. WOZNIAK (1975) Effects of interpolating a secondary task during the intertrial interval of a concept

identification problem. Memory and Cognition, 3,

648-652.

A. WILSON (1970) War Gaming. Penguin, Harmondsworth.

M.T. WRIGHTSON (1976) The documentary coding method. In R. AXELROD, editor (1976a) Structure of Decision. Princeton

University Press, Princeton, New Jersey.

L.R. ZEITLIN & J.M. FINKELMAN (1975) Research note: subsidiary task techniques of digit generation and digit recall as

indirect measures of operator loading.

Human Factors, 17, 218-220.

APPENDIX : PUBLISHED MATERIAL

This appendix contains reproductions of four published papers written jointly by myself and other authors, and related to work described in this thesis.

1. D.F. COOPER & J.H. KLEIN (1980) Board wargames for decision making research. European Journal of Operational Research, 5, 36-41.
2. D.F. COOPER, J.H. KLEIN, M.R.C. McDOWELL & P.V. JOHNSON (1980) The development of a research game. Journal of the Operational Research Society, 31, 191-193.
3. J.H. KLEIN & D.F. COOPER (1981) Assessing problem complexity. European Journal of Operational Research, 6, 243-247.
4. J.H. KLEIN & D.F. COOPER (1982) Cognitive maps of decision-makers in a complex game. Journal of the Operational Research Society, 33, 63-71.



# Board wargames for decision making research

Dale F. COOPER

*Department of Accounting and Management Economics,  
University of Southampton, Southampton SO9 5NH,  
United Kingdom*

and

Jonathan KLEIN

*Department of Mathematics, Royal Holloway College  
(University of London), Egham, Surrey TW20 0EX,  
United Kingdom*

Received November 1978

Revised June 1979

This paper discusses some aspects of a game for research into decision making. From our experience with commercial board wargames, we identify six requirements for a research game, and three requirements for its control. The Organisational Control Game, a wargame which we have developed, is shown to satisfy the requirements. The Superior Commander system is used to control the game, and the importance of the dummy task is discussed. We conclude that commercial board wargames can be adapted to examine decision making scientifically, we offer some guidelines for game development, and we consider extensions to other contexts, particularly those related to business games.

## 1. Introduction

There have been many studies of the behaviour of the Decision Maker in crisis, ranging from detailed observations of the actors in specific international crises to broader generalisations concerning the type of effects that might be caused by crisis (see, for example, [9]). However, comparatively little has been done to investigate models of decision making in

The advice and encouragement of Mr. K.C. Bowen, Head of Research, Defence Operational Analysis Establishment, and of Professor M.R.C. McDowell, Royal Holloway College, is gratefully acknowledged. This research is supported in part by the Ministry of Defence under Contract Number ER3/9/4/2088/012 to Professor M.R.C. McDowell.

crises in controlled experiments. It seems clear that crisis influences decision making in certain definite ways, and it would be useful to be able to compare and study models of the Decision Maker which attempt to account for such phenomena.

The aim of our work has been to develop a suitable methodology to test models of the Decision Maker. It is difficult to study decision makers in naturally-occurring crises in a controlled fashion: crises, by their nature, are often unexpected, they occur infrequently, and they take many forms. In such situations, important variables are not subject to control in the strict way we would like for hypothesis-testing. We wish to study decision making in environments in which variables can be manipulated to enable specific aspects to be examined. A gaming approach seems suitable for this purpose [4]. Our initial experimental work has been directed towards the development of a game in which a subject would be required to make decisions under experimentally controlled conditions, and in which it would be possible to analyse decision making in detail.

This paper is directed to those interested in research methods for investigating decision making, and it discusses some of our experience with board wargames for this purpose. We begin by outlining our initial experiments with a commercially-available wargame. From this, we derive general requirements for an experimental game and a gaming methodology for research into decision making. Next we describe briefly a game we have developed and a method for controlling it experimentally to satisfy our requirements, and we note possible extensions to other areas.

The conclusions we draw are based on our own specific research context. This is concerned with the ways in which decision makers' perceptions of the options and alternative courses of action open to them change under various conditions, particularly those related to crisis or stress and to different levels of data aggregation. However, our conclusions are not restricted to this context, and we believe them to have more general applicability to a wider area of decision making and gaming research. Some of these broader considerations are discussed in the final section.

## 2. Experiments with a tactical wargame

Wargaming has been used for the analysis of military situations and the training of commanders since the early nineteenth century [13], but it is only in the last 25 years that wargames have become widely available to the public [12], and a variety of commercial board wargames is now produced [10]. The first stage of our research programme was to assess whether commercial board wargames could be adapted for research purposes.

The game we chose to study was *Panzer Leader* [2], a wargame depicting tactical armoured combat in World War Two. This game is complex, with many types of fighting units and a variety of terrain. Its complexity was one of its most attractive features, suggesting that it might have potential for research. It was envisaged that scenarios could be designed which involved decisions concerning a wide range of alternative courses of action, offering many options to players. Palmer [10] regards *Panzer Leader* and its companion *Panzer Blitz* [1] as exciting and requiring a high level of skill. It was hoped that these features would encourage experimental subjects to become involved with the scenarios and induce them to play well.

Our initial choice of a gaming methodology was the Intelligence Man system developed by Sharp and Dando [11]. In this system, subjects believe themselves to be acting as advisers to real players, and they are led through identical series of predetermined moves. Subjects supply the experimenter with information about their views of the game via messages to the players whom they believe they are advising, and possibly by a real move which concludes the game. Sharp and Dando used their Intelligence Man system successfully with an adaptation of the commercial boardgame *Diplomacy* [3].

### 2.1. Recording a tactical game

A desirable feature of an experimental game is that it is possible to record the course of the game accurately and quickly, with a minimum of disturbance to the players. Our preliminary investigations concerned the feasibility of recording *Panzer Leader* games. Two games using one simple scenario were played, with four student players. The experimenters attempted to record the play on prepared forms.

The results indicated convincingly that recording *Panzer Leader* is not practicable. Either play must be

held up frequently, or there must be a long pause at the end of each turn while the recorder notes the outcome in detail. Neither alternative is satisfactory. Additionally, for several operations in *Panzer Leader* the player must explain to the recorder exactly what he is doing; virtually all combat requires such commentary.

Even when the time spent on the recording process exceeded tolerable limits and interference to the players was considerable, recording errors still occurred. It is clearly not possible to record *Panzer Leader* or similar games reliably without severely disrupting the flow of the game.

### 2.2. Intelligence Man system applied to *Panzer Leader*

Our next experiments were to assess whether the Intelligence Man system could be implemented using games like *Panzer Leader*. Intelligence Man worked for *Diplomacy* [11], but the kind of wargames we hoped to use differed from *Diplomacy* in many ways and the applicability of Intelligence Man could not be taken for granted.

The output of an Intelligence Man experiment is to a large extent obtained in the messages the adviser sends to the person he believes is playing the game, and therefore it is important to find out exactly what the subject, in his role as adviser, makes of his task. He may see himself as similar to a real player, and his messages may provide detailed suggestions concerning the microstructure of play; alternatively, he may take a wider view, and, freed of the constraints of organising the details of movement and combat, may point out a range of strategies and options. It is in this latter mode that the output of the game is most interesting for our research purposes.

Four students were used for an Intelligence Man evaluation applied to *Panzer Leader*. Two subjects played a normal game, while the other two acted as advisers to them. The advisers were separated from the players (all communications being in writing, and carried by the experimenters) and they were presented with an identical game. It is important to note that Intelligence Man, as used here, differs from conventional Intelligence Man experiments in that there was a *real* game going on — no deception was involved. Our purpose was to study the nature of the output of the advisers and their general attitude to the game. Their reactions to the specific scenario were not being tested.

Despite repeated discouragement from the experi-

menters, the advisers tended to become extremely involved with the detailed mechanics of the game. The game was viewed from a tactical rather than strategic viewpoint and there was little sign that the advisers had studied and assessed alternative courses of action. They were, in effect, thinking as frustrated players rather than as advisers. This was not the result for which we had hoped.

It seems an intrinsic quality of *Panzer Leader* and similar tactical games to encourage this mode of thinking. *Panzer Leader* was designed as a game of tactical level combat, and to expect subjects to look for broad strategies when it is clear that the outcome will be decided at a tactical level is perhaps unfair. Everything about the format of *Panzer Leader*, the wealth of terrain detail, the diversity of units, and the range of combat strengths, demands attention at a tactical level. We are now satisfied that this type of game has little to offer us in our experimental work.

### 3. Requirements for a research game

From our experiments with *Panzer Leader* we have derived general conclusions about the qualities an experimental game and a gaming methodology should have to be useful for research. Here we concentrate on the requirements specific to the game itself.

**G1: Simple, easily learnt rules.** The rules of *Panzer Leader* are very detailed. Even experienced players sometimes unintentionally broke them, without their opponents noticing. It is clear that the game rules must be much simpler than those of *Panzer Leader*, simpler, indeed, than those of all but the most elementary commercial wargames. Not only must the rules be simple to apply, but they must be easy to learn. Ideally, when the game simulates closely the environment in which the decision maker usually operates, the rules will be the normal "rules" of that environment, and so will be known to the player. Few formal game rules will then be required.

**G2: Speed of play.** It is essential that play in an experimental game moves quickly, otherwise the subject may become fatigued and lose interest in the game towards the end. (Usually it is towards the end of a game that we would hope to extract the most interesting information.) We feel that an experimental wargame should not last longer than 3h, and we regard 30 min as the maximum acceptable time for a turn, although it would be preferable if this could be cut by as much as half.

**G3: Ease of recording.** It should be possible to keep an accurate record of an experimental game. The format of the game should allow easy, swift recording, with the minimum of interference to the players.

**G4: Strategic aspect.** For our research, a strategic orientation is the most important property of the game, because strategies and options form the material which we would like the subject to consider. We would like a game which is entirely strategic, presenting no tactical aspects to the subject at all. He should adopt a strategic frame of mind naturally, prompted only by the format of the game; it should not appear to be a rather unconventional and contrived way of looking at the situation. Scenarios must be rich in strategic options, and there must be potential decisions for the subject to make. To some extent preparation of these rich scenarios could clash with the requirement that the game be simple; careful design of scenarios is necessary to reconcile this conflict.

**G5: Interest.** To be successful, an experimental game needs to hold the attention of the subjects for the duration of play. Thus, in designing a simple game, we must not allow the game to become too abstract and unrealistic. A strategically-rich game seems likely to hold the interest of subjects.

**G6: Natural data extraction.** As much as possible of the data we wish to extract from the experiment should be provided by the player in the natural course of the game. Intelligence Man achieves this elegantly: data is supplied by the advisers' communications to their (non-existent) players, and by their final move. We would hope for a similarly natural data extraction process, which would encourage the player to regard himself as a participant in a game rather than a subject in an experiment.

### 4. Requirements for a methodology

The requirements for a research gaming methodology have been discussed in detail elsewhere [5,6]. We state them here for completeness.

**R1:** The Player in a game must believe that he is an active Decision Maker.

**R2:** A game must contain sufficient complexity and detail to be accepted by the Player as realistic.

**R3:** The experimenter must be able to control a game so that he can replicate it exactly with different subjects.

## 5. The Organisational Control Game

The Organisational Control Game is a game which we have designed to satisfy our research requirements. It is a wargame in which the player acts as a Divisional Commander (Div Comd) in a Blue (friendly) force engaged in combat with a Red (enemy) force. Subordinate to the Div Comd are three Task Force (TF) Comds, and above him is a Corps Comd (Fig. 1). Detailed combat resolution occurs at the TF level; the Div and Corps Comds provide advice and orders and exercise overall control. From the Div Comd's point of view, the game resembles a traditional wargame or Command Post Exercise. However, all communications are transmitted by Game Control, and there is no direct contact between different levels of command.

The game uses the Superior Commander system. It is a single-decision game [4], in which Blue plays against a predetermined Red threat. It is also a single-player game; TF and Corps Comds are played by Game Control, and they respond in a preset way. However, the Div Comd does not know he is the only player, nor that the game is predetermined. He believes that he is an active decision maker in an interactive game involving other purposeful players. The game is complex and realistic, yet it can be controlled and replicated, and so detailed comparisons of the single decision from game to game can be made.

The Superior Commander has been described elsewhere [5,6], and we shall not discuss it in detail here. We note, however, that it satisfies the three requirements for a methodology R1–R3, and we shall elaborate on one aspect of it, the dummy task.

### 5.1. The dummy task

The main tasks of the Div Comd are to collect and collate information received from the TF and Corps Comds and from local reconnaissance units, and to issue advice and orders. He also has another task, keeping an adequate line of supply from his base area

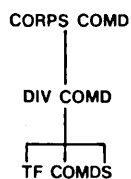


Fig. 1. Command structure of the Organisational Control Game. The player is a Div Comd.

to his forward TF's. This appears to be an important function for the Div Comd, as his combat guidelines indicate that the results of combat at the TF level may be affected by the provision of logistic support (see below). However, this is not the case; all combat is predetermined, and supply has no effect on the outcome. The supply task is a dummy task. (The allocation of air strikes was used as a dummy task in an early version of the game.)

The dummy task has important functions in the Superior Commander system. It helps to maintain the deception that the player is interacting and influencing the game (requirement R1), and it adds complexity and realism to the scenario (requirement R2), without affecting the replicability of the game (requirement R3). As we see below, it also assists the Organisational Control Game to satisfy some of the game requirements G1–G6. The concept of this kind of dummy task is believed to be new; it has been adapted in at least one other research game to maintain the involvement of players and to avoid loss of concentration [8].

### 5.2. Game requirements

We now show that the Organisational Control Game, with the Superior Commander system, satisfies the requirements G1–G6. The results are based on a series of games with six student players, and a second series of games with ten former senior military staff officers. A detailed discussion of some aspects of these games is given in [7].

G1: *Simple, easily learnt rules.* The design of the game is such that all combat is resolved at TF level. By removing the possibility of commenting on the detailed tactical aspects of the game, many of the problems associated with learning the rules necessary in a complex wargame are eliminated. The Div Comd's rules are extremely simple, and even an inexperienced player was able to grasp them very quickly. For example, in the first scenario the Div Comd is given the following guidelines:

Combat is resolved in detail at the Task Force level, and Divisional Commanders require no detailed knowledge of the combat rules.

An enemy regiment may be twice the size of a friendly battalion, but at the beginning of the game the strengths of enemy units not in contact with friendly forces are unknown. As a rough guide to the relative fighting strengths of the units, a battalion stands a good chance of defending success-

fully against an enemy regiment, but a larger enemy force will probably defeat it.

The defensive strength of a unit is increased if the unit did not move in the previous turn; this represents the ability of a unit to "dig in". Both the offensive and defensive strengths of units are severely reduced if they are not kept in supply. The strengths of units are increased if they occupy hill hexes or town hexes.

G2: *Speed of play.* Because individual units are represented as aggregates on the Div Comd's map, there is very much less work to do in assessing them and organising them in each turn. This reduces considerably the time taken for each move.

G3: *Ease of recording.* The Div Comd cannot move units directly, only indirectly through his orders to the TF Comds, and detailed combat evaluation only occurs at TF level. Very little recording is therefore necessary.

G4: *Strategic aspect.* The Organisational Control Game at Div Comd level is very different from *Panzer Leader*. Details of the tactical battle are eliminated from the Div Comd's map, and his view is in all respects strategic. He generally gives strategic rather than tactical advice to his TF Comds, and he is concerned almost entirely with strategies and possible courses of action, both of his own forces and the enemy. This approach is encouraged by the dummy task: the need to maintain logistic support requires the player to think closely about his own lines of communication, his future plans and the intentions of the enemy.

G5: *Interest.* The Div Comd's task in the Organisational Control Game, while conceptually simple and unencumbered by detailed rules, is interesting for a player. We base this conclusion on our observations of the players' concentration during the course of the games, and their comments during post-game debriefing. The students were very enthusiastic (they even designed their own multi-player hierarchical game, using ours as a model), while the military players were able to make detailed suggestions for modifications and improvements to the game. (For further discussion of this point, see [7].) The players claimed that the dummy task was one reason for their interest in the game; they thought it was important, they could control it directly, and, although it was straightforward, it did lead them to consider some wider strategic problems.

G6: *Natural data extraction.* In the initial versions of the game, Div Comds sent written orders and advice

to the TF Comds, and prepared written assessments for the Corps Comd. These reports, which provided the main experimental data, were a natural part of the game communications. However, the second series of games, with players with military experience, forced us to modify this procedure [7]. At the suggestion of the experienced players, all reports are now verbal ones, recorded on cassette tapes, and the reporting has been divided into phases, separating the advice and orders which are a natural part of the game from the assessments which are more obviously part of the experiment. The same information is sought, but the players are happier with this arrangement.

The dummy task may also be a source of experimental data. The allocation of supply and the emphasis placed on different aspects of it may reveal or confirm the player's intentions and priorities, again as a natural part of the game.

## 6. Discussion

Of the requirements we have considered for a research game, a strategic aspect is particularly important if the Superior Commander System is to be used. Only at a strategic level can the deception that the player is affecting the game be maintained, and hence the replicability of the game guaranteed. If a player is allowed to make tactical decisions, almost by definition these must be interactive in some sense, and so the player will expect to see some result. We provide a dummy logistic task to deceive the player that he is making meaningful decisions, but the deception could not be sustained if he was permitted, for example, to direct the movements of individual combat units. It seems likely that this is a general result for games using this method of control.

The hierarchical command structure in the Organisational Control Game allows the tactical and strategic aspects to be separated naturally. It would appear that the Superior Commander could be applied in other hierarchies, in areas where there can be a similar separation of tactical and strategic decisions. For example, the system could be implemented in the context of a business corporation (Fig. 2), where the operational (tactical) decisions are taken by Regional Managers (corresponding to the TF Comds), while medium-term strategic planning is undertaken by a National Manager (the Div Comd), under the overall guidance of a Corporate Headquarters (the Corps

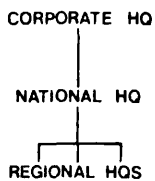


Fig. 2. A corporate business hierarchy suitable for the Superior Commander. The player would be a National Manager.

Comd). A hierarchy based on functional divisions is also possible in a business context (Fig. 3), where the functional groups might include areas such as production, sales, finance, distribution and R & D.

Most existing business games are interactive, concentrating attention on operational (tactical) decisions in a directly competitive market place, rather than on the longer-term strategic aspects of the firm. However, there seems no reason why these games could not be adapted to construct a game suitable for research into business decision making, just as we have adapted board wargames to develop the Organizational Control Game. With careful scenario design it should be possible to satisfy the requirements G1-G6, and the application of the Superior Commander system should be straightforward. The kind of dummy task that is used will depend, of course, on the specific game context and on the purpose of the research.

In this paper the only commercial wargame we have discussed in detail has been *Panzer Leader*. However, we have played a variety of board wargames, exhibiting many different features, and we have extracted those characteristics of physical design, layout and rules which are most suited to our needs. (In the wargame context, useful references for this

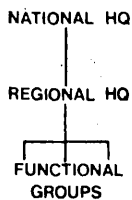


Fig. 3. A functional business hierarchy suitable for the Superior Commander. The player would be a Regional Manager.

purpose are [10] and [12].) We recommend that anyone seriously considering using a game for research, whether one that is commercially available or one, like ours, which is developed specially, should play a variety of commercial games to obtain a broad understanding of their capabilities. This is a general recommendation which seems as appropriate to games in a business context as to wargames.

We have shown that commercial board wargames, with the Superior Commander system, can be adapted to examine decision making scientifically, and we have indicated some properties we think a research game should have. Using these guidelines, it seems likely that other commercially-available games, and particularly business games, could be adapted similarly for research purposes.

#### References

- [1] Avalon Hill (Manufacturer), *Panzer Blitz*, Avalon Hill Game Company, Baltimore, MD (1970).
- [2] Avalon Hill (Manufacturer), *Panzer Leader*, Avalon Hill Game Company, Baltimore, MD (1974).
- [3] Avalon Hill (Manufacturer), *Diplomacy*, Avalon Hill Game Company, Baltimore, MD (1976).
- [4] K.C. Bowen, *Research Games* (Taylor and Francis, London, 1978).
- [5] D.F. Cooper, On the design and control of crisis games, *Omega* 6 (5) (1978) 460-461.
- [6] D.F. Cooper, The Superior Commander: a methodology for the control of crisis games, *J. Operational Res. Soc.* 30 (6) (1979) 529-537.
- [7] D.F. Cooper, J. Klein, M.R.C. McDowell and P.V. Johnson, The development of a research game, submitted for publication.
- [8] D.W. Daniel, What influences a decision? Some results from a highly controlled defence game, Defence Operational Analysis Establishment, West Byfleet, U.K., submitted for publication.
- [9] C.F. Hermann, ed., *International Crises: Insights from Behavioral Research* (Free Press, New York, 1972).
- [10] N. Palmer, *The Comprehensive Guide to Board Wargaming* (Arthur Barker, London, 1977).
- [11] R.G. Sharp and M.R. Dando, The Intelligence Man system: an approach to assisting managers in complex conflicts, *Europ. J. Operational Res.* 1 (1977) 275-288.
- [12] *Strategy and Tactics, Wargame Design* (Strategy and Tactics Staff Study No. 2) (Simulations Publications Incorporated, New York, 1977).
- [13] A. Wilson, *War Gaming* (Penguin, Harmondsworth, 1970).

# The Development of a Research Game

DALE F. COOPER

Department of Accounting and Management Economics, University of Southampton,  
Southampton SO9 5NH

and

JONATHAN KLEIN, M. R. C. McDOWELL and PHILIP V. JOHNSON

Department of Mathematics, Royal Holloway College (University of London), Egham TW20 0EX

## INTRODUCTION

A RESEARCH game has been developed to investigate the behaviour of decision makers. Previous papers have described both the Organisational Control Game, a hierarchical wargame in which the player has the role of a Divisional Commander, and the Superior Commander system used to control and replicate it.<sup>1-3</sup> In this note, some results are reported from two series of plays of the Organisational Control Game, the first with students as players and the second with former military staff officers, and their behaviour is compared.

The first series of games, with the student players, was used to test the Superior Commander system and to develop two different game scenarios, one featuring a defence, the other an advance-to-contact. The long term aim of the work was to test hypotheses about the ways in which a decision maker's perceptions change as the level of stress increases.<sup>4,5</sup> And this series of games indicated that the Superior Commander system permitted the collection of the data necessary for this.

In the second series of games, the players were ten retired military officers, all of whom held senior positions in U.K. Local Government at the time of the experiment. After the success with the student players,<sup>2</sup> serious problems were not expected with these experienced players, but as the games progressed, it became clear that important changes were required in the mechanics of the game and the design of the scenarios. It was not possible to use these games for testing hypotheses.

The importance of using players with relevant experience during the process of game development is clearly demonstrated.

## EXPERIENCED AND INEXPERIENCED COMMANDERS

The differences in the behaviour of the two groups of players can be grouped into three categories, although these categories are inter-related and should not be considered in isolation.

1. The military officers who played the game used their previous experience of command. They considered factors which they had found valuable in practice, even though these factors were not mentioned explicitly in the game rules or briefing. They injected their own ideas, assumptions and preconceptions into the game; to an extent they "built their own scenarios" using the one provided as a basis. The students, on the other hand, tended to accept the scenario as presented, probably due to their experience with commercial wargames (in which game rules are more rigid and there is less flexibility).
2. The military officers had a more developed knowledge than the students, performing equally well on both the defence scenario and the advance-to-contact. They were more conscious of the losses that were possible in each case, and they had a better grasp of the principles of war (although these were violated on occasion). The students preferred attack to defence, and showed little appreciation of possible losses. While they

- managed advance-to-contact quite well, they had almost no idea about the way to conduct a defence.
3. The experienced players viewed the problems posed by the scenarios as complex and ill-defined, with no "best" solution. In reports to superior and subordinate commanders in the game, they assessed situations more fully and formed better strategic views of the scenarios. In contrast, the students made rather cursory assessments, looking for unique "best" solutions. They tended to see situations in uncomplicated terms, viewing all data as known with certainty, with no "fog of war". They also tended to regard all data as equally important, whereas the experienced players were more selective in their approach.

Both groups of players exhibited a wide range of playing styles, and both groups appeared very involved in the game. The deception inherent in the Superior Commander system, which was known to work for student players,<sup>2</sup> was also achieved with the experienced players: they were convinced that the game was interactive. As noted above, the experienced players made assumptions which "expanded" the original scenario, perhaps to make it more lifelike; it is suspected that this process, together with the preconception that no attempt would be made to deceive senior managers who were doing the experimenters a favour by playing the game, is important to the success of the deception.

## DISCUSSION

The experience indicates that student players are very useful in the initial process of designing the general structure of a game, planning experiments and testing for gross breaches of validity, but experienced players must be involved in the final detailed design (see also Hartley *et al.*<sup>6</sup>). After each of the second series of games a debriefing session was held, when the players themselves made valuable contributions and suggestions for improving the realism and level of detail of the scenarios, the mechanics of representing units and transmitting information, and methods for extracting research data. The conceptual abilities and verbal skills of the experienced players were decisive in this process. It is believed that scenario creation should be a dynamic, interactive procedure, involving both the game designers and players with suitable background and experience, and that this procedure should be a continuous one. Each series of games then becomes a learning process for the game designers, who benefit by having their scenarios probed each time the game is played (see also Brewer<sup>7</sup>).

The simple but nevertheless important point demonstrated by the games described in this note is that *students do not behave in the same way in a game environment as experienced decision makers*. Therefore, it would be unwise to generalise game results obtained with student players to professional decision makers. Differences in behaviour are both quantitative and qualitative. Quantitative differences in the alternative options considered and in the levels of skill displayed were expected, but these were also important qualitative differences. The students approached the game differently from the experienced players, and exhibited a different range of decision making styles. The quantitative and qualitative differences are apparent in all three of the categories identified in the previous section.

Many behavioural studies in the operational research and management literature have used student subjects to derive conclusions about professional decision making. It is clear that extreme care is required in making such derivations. This point has been demonstrated before, notably in accounting and financial contexts.<sup>8-11</sup> Birnberg and Nath<sup>12</sup> identify two ways in which student and professional decision makers are likely to differ, viz. "common skills and experience" and "basic personality traits". These factors are likely to result in the quantitative and qualitative differences noted above. Students, having considerably less experience and training in the decision making environment,



display less quantitative skill when required to work in that environment. Their different personality traits, including cognitive styles, values and beliefs, are bound to have a qualitative effect on their decision making.

In general, qualitative differences in decision making styles are likely to become apparent in comparisons between any distinct groups of individuals. This is seldom explicitly recognised in operational research, although instances where it is implied can readily be found. Bowen,<sup>13</sup> for example, notes the limitations of having a Royal Navy officer representing an enemy submarine commander.

The recognition and study of different decision making styles is believed to be essential to the understanding of decision making. Single decision games provide a suitable environment in which to identify and measure the different dimensions of decision making that together make up the decision making style of an individual.

#### ACKNOWLEDGEMENTS

The authors wish to thank the students from Royal Holloway College who assisted in the first series of games. They also wish to express their gratitude to the former military staff officers who so generously gave the benefit of their time and experience in the second series of games. The research described in this paper is supported in part by the Ministry of Defence under Contract Number ER3'9'4'2088 12 to Professor M. R. C. McDowell.

A version of this paper was presented at the Third European Congress on Operations Research, Amsterdam, Netherlands, 9–11 April 1979.

#### REFERENCES

- <sup>1</sup>D. F. COOPER (1978) On the design and control of crisis games. *Omega* 6, 460–461.
- <sup>2</sup>D. F. COOPER (1979) The Superior Commander: a methodology for the control of crisis games. *J. Opl Res. Soc.* 30, 529–537.
- <sup>3</sup>D. F. COOPER and J. H. KLEIN (1980) Board wargames for decision making research. *Eur. J. Op. Res.* 9, to appear.
- <sup>4</sup>O. R. HOLSTI (1972) Time, alternatives and communications: the 1914 and Cuban missile crises. In *International Crises* (C. F. HERMANN, Ed.), Chap. 4. Free Press, New York.
- <sup>5</sup>T. W. MILBURN (1972) The management of crisis. In *International Crises* (C. F. HERMANN, Ed.), Chap. 11. Free Press, New York.
- <sup>6</sup>D. A. HARTLEY, P. V. JOHNSON, A. FITZSIMMONS, J. LOVELL, B. CHIPPENDALE and J. K. S. CLAYTON (1979) A case study on the development of the Home Defence training game HOT SEAT. *J. Opl Res. Soc.* 30, 861–871.
- <sup>7</sup>G. D. BREWER (1978) Scientific gaming: the development and use of free-form scenarios. *Simul. Games* 9, 309–338.
- <sup>8</sup>A. R. ABDEL-KHALIK (1974) On the efficiency of subject surrogation in accounting research. *Acc. Rev.* 49, 743–750.
- <sup>9</sup>B. ALPERT (1967) Non-businessmen as surrogates for businessmen in behavioral experiments. *J. Bus.* 5, 203–207.
- <sup>10</sup>R. M. COPELAND, A. J. FRANCA and R. H. STRAWSER (1973) Students as subjects in behavioral business research. *Acc. Rev.* 48, 365–372.
- <sup>11</sup>J. E. FLEMING (1969) Managers as subjects in business decision research. *Acad. Mgmt J.* 12, 59–66.
- <sup>12</sup>J. G. BIRNBERG and R. NATH (1968) Laboratory experimentation in accounting research. *Acc. Rev.* 43, 38–45.
- <sup>13</sup>K. C. BOWEN (1978) *Research Games*. Taylor & Francis, London.

# Assessing problem complexity

Jonathan KLEIN

*Department of Mathematics, Royal Holloway College  
(University of London) Egham, Surrey TW20 0EX,  
United Kingdom*

Dale F. COOPER

*Department of Accounting and Management Economics,  
University of Southampton, Southampton SO9 5NH,  
United Kingdom*

Received June 1979

Revised October 1979

The difficulty of objectively measuring the decision-making load imposed by a complex decision problem on a human decision maker is identified. It is suggested that a technique for such measurement would be useful in Operational Research. The secondary task technique, a method much used by psychologists and ergonomists to assess the mental processing load presented to subjects by a variety of tasks, is discussed. It is observed that the type of complex problem often of interest in Operational Research is characterised by the involvement of model building in the process of solution. An experiment is presented to compare the mental load presented by a number of different chess problems using a mental arithmetic secondary task, and it is shown that the method appears to be valid for the study of complex problems.

## 1. Introduction

Situations are often found in Operational Research in which it would be useful to have some objective measure of the complexity of a problem facing a decision maker. Different problems might be expected to present different *decision-making loads* to a decision maker, and it would be of interest to know how decision-making load varies according to the precise features of the problem. In particular, it might be useful to be able to identify those problems which present

decision makers with particularly high decision making loads. In the case of large-scale decision-making systems, this would allow the decision points at which satisfactory decision making is likely to break down to be located. These points are not necessarily easily identifiable, either by inspection or by theoretical analysis of the systems.

An objective method of assessing problem complexity might be extremely valuable in a variety of areas. In problems of medical diagnosis, for example, where diagnoses based on a limited quantity of data are often required in a short time, an objective technique would permit investigation of the way in which ease of diagnosis varies according to the quantity and quality of the data, its nature, and the way in which it is presented. Similar considerations apply to military decision-making in battle, where once again time is likely to be a severe constraint. What types of decision tend to overload military decision makers, causing their decision-making to deteriorate? By answering this question with some certainty, it might be possible to begin to devise suitable decision rules and decision-making systems which would avoid, to some extent, the worst effects of unexpected overload.

The authors' interest in problem complexity arose in the course of experimental gaming work [6,7]. A wargame was being used to study the decision-making behaviour of players under various conditions. Two different scenarios had been designed, and a method was sought to compare objectively the decision-making load imposed by them. The *secondary task* method of assessing mental processing load, a technique used by experimental psychologists and ergonomists, was adopted.

This paper discusses the secondary task technique in an Operational Research context. Section 2 outlines the theory of the technique and briefly describes some previous work in the field, while Section 3 considers the nature of the decision-making problems relevant to Operational Research. Section 4 describes an experiment involving this type of problem, and presents preliminary results. Conclusions are presented in Section 5.

---

The research described in this paper has been supported in part by the Ministry of Defence under Contract Number ERC/9/4/2089/012 DOAE to Professor M.R.C. McDowell.

© North-Holland Publishing Company  
European Journal of Operational Research 6 (1981) 243-247

## 2. The secondary task technique<sup>1</sup>

The secondary task technique is based on a model of the decision maker as a mental processing system of limited capacity (Fig. 1a). Kerr [10] identifies and discusses critically the salient points of the model and the assumptions implicit in it. According to the model, any mental load lower than the capacity of the system will undergo satisfactory processing (Fig. 1b), and errors due to overload will not occur. However, if capacity is exceeded, then overload occurs, and mental processing consequently deteriorates (Fig. 1c); additional errors in performance will occur in this case.

Consider now two tasks, the loads of which are to be compared (Fig. 2a). If neither task individually overloads the system, they will both be carried out properly, and no direct measure of load in terms of performance standard will be possible. How can the tasks be compared? In the secondary task technique, the subject is asked to perform a secondary task of constant difficulty at the same time as the main tasks (Fig. 2b). This task is such that the combined loads of the main and secondary tasks together exceed the capacity of the decision making system. Now overload does occur, and processing errors are observed.

Subjects are asked always to give priority to the main, primary, task, so that this task is not affected by overload. The deterioration of performance is directed to the secondary task. It is argued that *the greater the deterioration of the secondary task, the greater must be the mental processing load presented by the main task.*

Psychologists have used secondary task techniques successfully in many areas. In ergonomics, for example, secondary tasks have been used to evaluate man-machine systems: the loads imposed on human operators by different pieces of equipment built for the same purpose can be compared (e.g., [8]), and in this way guidelines for the redesign of systems to increase the effectiveness of the human operator may be provided. Brown [3] reviews the secondary task literature and describes the essentials of the technique. Rolfe [13] provides a fuller and more recent review of, in particular, the ergonomic literature.

<sup>1</sup> The secondary task technique to be described here is the *subsidiary measuring task* technique, as distinct from the *loading task* technique. The distinction, which it is not necessary to describe here, is clearly defined by, e.g., Brown [3].

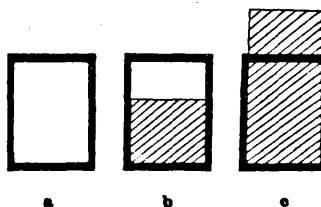


Fig. 1. (a) A mental processing system of limited capacity. (b) Loads lower than the limit are processed satisfactorily. (c) If the limit is exceeded, processing deteriorates.

A series of experiments on motor-vehicle driving illustrates the power of the technique. It seems reasonable to suggest that the load imposed by driving in heavy traffic is greater than that imposed by driving in light traffic, but how can this be shown in practice? Driving behaviour can seldom be seen to deteriorate, and, when it does, the results are often catastrophic. However, secondary tasks have shown the differences in driver load under different conditions, without adversely affecting driving. Brown and Poulton [5], for example, using a mental addition secondary task, compared driving in a quiet residential area to driving in a busy shopping area, and found that the error rate under the 'shopping' condition (7.3 mean percentage errors, eight drivers) was significantly higher than under the 'residential' condition (6.3 errors), and that both error rates were higher than a non-driving control (4.1 errors). The error rates appear to correspond to the difficulty of the primary task. Similar experiments have investigated fatigue effects in police drivers [1,2], and the training of bus drivers [4].

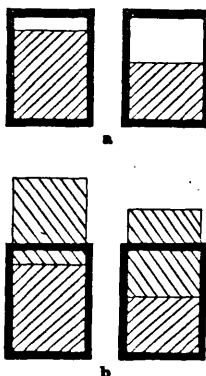


Fig. 2. (a) The loads due to two primary tasks, neither of which overload the processing system. (b) The loads of the primary tasks, combined with a secondary task of constant load.

### 3. The nature of complex decision problems

Operational researchers might be interested in tasks which differ in an important way from the types of task investigated by psychologists using secondary task techniques. In virtually all the studies reported in the psychological literature, the primary tasks can be categorised as short-term tasks in which solution models for decisions are already available to the subject (Fig. 3a). (Note that in the driving tasks cited above the aspects studied are related to the physical handling and control of the vehicle.) In other words, for a given set of stimuli, the subject already possesses the essentials of a suitable response. After perceiving the input stimuli, the subject consults his mental model of the situation, which he already has available in memory, and he selects the appropriate response. This process rarely takes more than a matter of seconds.

The types of task of interest here are different. Being those necessary to solve complicated problems, they are generally far more complex, possibly less easily formulated, and straightforward solution methods are unlikely to be immediately available. A different process takes place to solve the problem (Fig. 3b). Following the perception of the input stimuli, the subject has to *construct* a model of the situation. He then uses this model to select a response, as before. The important point is that *the subject must construct his model as an integral part of the complex decision process*. By constructing the model he is learning how to solve the problem.

In this type of problem, there is unlikely to be a unique 'best' solution. Indeed, it is often impossible to compare the quality of different solutions. Correctness may depend to a greater or lesser extent on subjective assessments as well as on objective factors. In general, solutions to complex decision problems cannot be considered as 'right' or 'wrong', and the

quality of the decision cannot be used as a guide to the load imposed by the problem.

### 4. A secondary task experiment

A preliminary experiment was designed to extend the secondary task technique to the study of complex decision-making problems. As experimental problems, chess problems were used. The degree of complexity of these problems, and the model building that their successful solution requires, are similar to those the imposed loads of which operational researchers might wish to assess. A useful additional feature is that in chess problems a best solution usually exists, and different solutions can often be compared quite meaningfully.

The subjects were six undergraduate students at Royal Holloway College. Preliminary work had shown that simple problems were required in order to be within the ability of the average student. A series of fourteen suitable problems [11] was compiled, in the majority of which a subject was asked to select the move he would make next to win as quickly as possible. A typical problem is reproduced in Fig. 4, in a similar form to that in which it was presented to the subjects. Players were allowed two minutes to attempt each problem. At the end of two minutes they were asked to write down the solution. The first five problems were 'learning' problems; only the results of the final nine problems were analysed.

The secondary task was run concurrently with the chess problems (and also by itself at the beginning and the end of the series of problems). A tape recording of pseudo-random digits, from two to nine, read

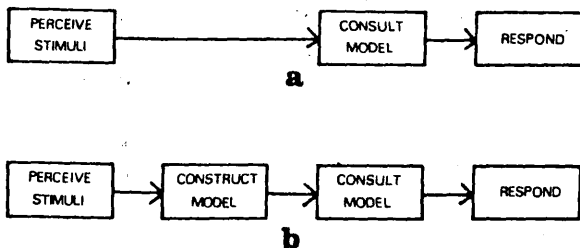


Fig. 3. Solution processes for (a) simple and (b) complex problems.

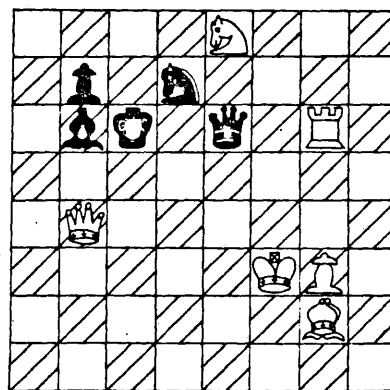


Fig. 4. Chess problem 6 in the experimental set. White to move, playing up the board. K to B4 gives a mate.

out at five second intervals, had been prepared. For the secondary task, subjects were asked to call out the sum of each pair of digits. For example, if the first two digits were 'two' and 'seven', the correct response would have been 'nine', and if the next digit was 'three', the next response ought to have been 'ten'. The number of mistakes and omissions in this task was recorded, and a total error score for each problem was obtained.

The problems were ranked for each subject according to the number of errors occurring in the mental arithmetic. With the assumption that the fewer errors that are made in the secondary task, then the easier is the main problem, this gives a rank ordering of problem difficulty. Agreement between the rankings for individual subjects was significant at the 2% level (Kendall concordance test,  $\chi^2 = 19.7$ ,  $\nu = 8$ ,  $p < 0.02$ , [9]). This indicates that the secondary task technique is measuring some property of the problems which is consistent from subject to subject. The individual rank orderings were combined to produce a total ranking of the problems in order of increasing complexity. This ranking, in which the total arithmetic task error rate for each problem has been normalised and summed over all subjects, so that the contribution from each subject to the overall result is the same, is shown in Table 1(a).

The individual rank orderings show a high degree of concordance, and it has been assumed that these orderings correspond to the difficulty of the problems. An independent test of the meaning of the orderings would be desirable. In the case of chess problems, such a test is possible. Most of the problems have one solution which is clearly better than all other possible solutions. Sixteen players in all have attempted these problems, which have been ranked according to the number of times the correct solution was given. (Some of the players attempted the problems as part of another experiment.) It is suggested that those problems which are solved most often present the least processing load, and those which are seldom solved present the greatest load.

The ranking of the problems by solution rate is shown in Table 1(b). There are significant similarities between this ordering and the secondary task ordering (Kendall's  $\tau = 0.63$ ,  $p < 0.05$ , [9]). Problem 14 was solved more often than any other problem, and the arithmetic task indicates that the load it presents is less than that presented by any other problem. At the other end of the scale, problems 7, 12 and 11 were solved least often, and they seemed to be the problems which presented the greatest load to the subjects. The five problems in the centre of the scales, however, display no apparent agreement in ranking.

Fig. 5 gives some suggestion of why this might be so. Each point on the graph represents a problem. The scale on the horizontal axis represents the number of times (out of a possible maximum of sixteen) that each problem was incorrectly answered. The scale on the vertical axis is the total arithmetic task error rate for each problem, normalised and summed over all subjects.

At the bottom left hand corner of the graph is problem 14, often solved, and with a low processing load. At the top right are problems 7, 11 and 12, seldom solved, and with a high load. The central group

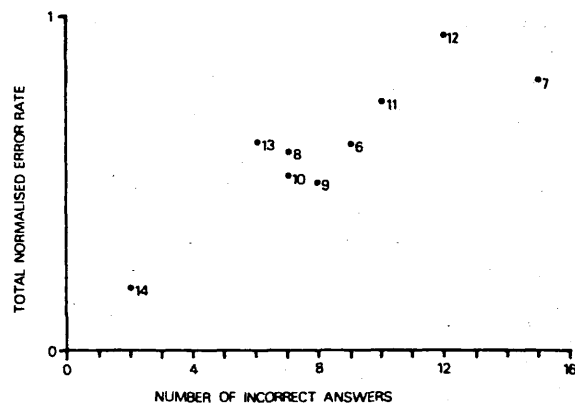


Fig. 5. Total normalised secondary task error rate plotted against the number of incorrect answers for problems 6 to 14.

Table 1

Problems ranked in order of increasing difficulty. Problem numbers in parentheses are of equal rank

Rank	1	2	3	4	5	6	7	8	9
(a) by secondary task	14	9	10	8	(6	13)	11	7	12
(b) by solution rate	14	13	(8	10)	9	6	11	12	7

consists of the other five problems. These problems cluster together on *both* scales. For these problems, the solution rates are all similar, and so are the arithmetic error scores. This indicates that the two rankings of these problems do not agree because the problems are of similar complexity. Neither ranking method can differentiate between these problems. This effect can also account for the difference between the two rankings of the 'high load, seldom solved' problems (problems 7, 11 and 12).

### 5. Conclusions

The points in Fig. 5 lie within a well-defined band across the graph, running as might be expected if the secondary task technique is measuring the load associated with the problem. This result indicates that this type of mental arithmetic secondary task is suitable for measuring problem complexity as intended.

A step that might be taken in the immediate future is to modify the arithmetic task in order to make the method more sensitive. It is felt that such modifications are possible. Eventually, the method might even be able to separate significantly those problems that appear in the central group of Fig. 5. In this case it would provide a better measure of problem complexity than the solution rate measure. Other kinds of secondary task may also be applicable [12].

The experiment that has been presented in this paper indicates that secondary task techniques can be used to measure the complexity of 'difficult' problems for which no obvious solution is available. As has been discussed, the chess problems with which this work has been concerned are similar in important ways to the complex decision problems with which decision makers are likely to be faced, and which arise frequently in Operational Research. We believe that the method has potential for extension to 'real' decision problems, the precise nature of

which may be difficult to define, and where unique 'best' solutions may not exist. We suggest that the secondary task method may provide a new technique for operational researchers to use in the investigation of a variety of complex decision making tasks.

### References

- [1] I.D. Brown, Measuring the 'spare mental capacity' of car drivers by a subsidiary auditory task, *Ergonomics* 5 (1962) 247-250.
- [2] I.D. Brown, Studies of component movements, consistency and spare capacity of car drivers, *Ann. Occupational Hygiene* 5 (1962) 131-143.
- [3] I.D. Brown, The measurement of perceptual load and reserve capacity, *Trans. Assoc. Indus. Medical Officers* 14 (1964) 44-49.
- [4] I.D. Brown, Subjective and objective comparisons of successful and unsuccessful trainee drivers, *Ergonomics* 9 (1966) 49-56.
- [5] I.D. Brown and E.C. Poulton, Measuring the spare 'mental capacity' of car drivers by a subsidiary task, *Ergonomics* 4 (1961) 35-40.
- [6] D.F. Cooper, The superior commander: A methodology for the control of crisis games, *J. Operational Res. Soc.* 30 (1979) 529-537.
- [7] D.F. Cooper, J.H. Klein, M.R.C. McDowell and P.V. Johnson, The development of a research game, *J. Operational Res. Soc.* 31 (1980) 191-193.
- [8] W.D. Garvey and F.V. Taylor, Interactions among operator variables, system dynamics, and task-induced stress, *J. Appl. Psychol.* 83 (1959) 164-171.
- [9] M.G. Kendall, *Rank Correlation Methods* (Griffin, London, 1975).
- [10] B. Kerr, Processing demands during mental operations, *Memory and Cognition* 1 (1973) 401-412.
- [11] J.H. Klein and D.F. Cooper, Chess problems for secondary task assessment (set C), Department of Mathematics, Royal Holloway College, University of London (1979).
- [12] J.A. Michon, Tapping regularity as a measure of perceptual motor load, *Ergonomics* 9 (1966) 401-412.
- [13] J.M. Rolfe, The secondary task as a measure of mental load, in: W.T. Singleton, J.G. Fox and D. Whitfield (Eds.), *Measurement of Man at Work* (Taylor and Francis, London, 1971).

# Cognitive Maps of Decision-Makers in a Complex Game

JONATHAN H. KLEIN

Department of Mathematical Sciences and Computing,  
Polytechnic of the South Bank, London

and

DALE F. COOPER

Department of Accounting and Management Economics, University of Southampton

This paper illustrates the use of a cognitive mapping technique to examine the behaviour and perceptions of individual decision-makers. A cognitive map is a representation of the subjective decision-making environment of an individual. Seven military officers each played two scenarios in a research wargame. Analysis of their communications in the game showed that individual players were remarkably consistent over the two scenarios, but their perceptions of their common decision-making environment differed noticeably. Differences related to the size and complexity of their cognitive maps, the detailed interpretation of the maps, the players' confidence and anticipation of the future and the way in which the maps were altered as time progressed.

## INTRODUCTION

THE AIM of Operational Research is to aid decision-makers. In the past, this has been achieved largely by the provision of techniques, together with guidelines for their application, directed towards the resolution of reasonably well-defined problems. In this area Operational Research has been unquestionably successful.

Recently, however, attention has been increasingly directed towards a different type of problem. It has now been widely recognised that many of the problems faced by decision-makers exist within complex, ill-defined and interconnected 'messes',<sup>1</sup> located in fundamentally turbulent environments.<sup>2</sup> Accordingly, much attention is now directed towards the problems of strategic decision-making and decision-making in conflict and crisis. It is recognised that purely analytical approaches are inadequate to deal with such problems, and the need for the development of a more general 'science' of decision-making has been identified.<sup>3</sup>

An important part of the development of a decision-making science is the provision of more detailed models of the individual decision-maker than exist at the moment. Psychological research provides a wealth of information about the behaviour of individuals, while political science and international relations have long been concerned with high-level decision-making, but it is only recently that *combined* approaches have been adopted. The interdisciplinary stance of O.R. should make it ideally suited to take an active role in this more unified approach to the study of decision-making. O.R. can both contribute to this area and learn from it.

This paper illustrates the use of a cognitive mapping technique to examine the behaviour of individual decision-makers. A cognitive map is a representation of the perceptions and beliefs of an individual about his own subjective world. The technique has been employed in the fields of international relations<sup>4,5,6</sup> and administrative science,<sup>7</sup> and recently there have been applications in O.R.<sup>8,9</sup> It seems likely that cognitive mapping may be a fruitful method of analysing decisions in a number of O.R. areas.

An earlier version of this paper was presented at the *Young O.R. Conference*, Coventry, December 1980.

In this paper, cognitive mapping is discussed in the context of a research wargame. The general gaming approach and the experimental structure are outlined in the next section. Following that, the methods of deriving cognitive maps from text are described, with examples drawn from the games. Results are then presented: these mainly deal with the initial perceptions of the players in the games but include a brief discussion of dynamic aspects and the way in which perceptions alter as the game progresses. Finally, conclusions are drawn about the usefulness and applicability of cognitive mapping, both for the type of decision-making experiment described in this paper and also in a wider O.R. context.

### GAMING APPROACH TO DECISION-MAKING RESEARCH

The cognitive mapping technique will be described in the context of a research wargame known as the Organisational Control Game. Both the game and its associated system of control, the Superior Commander system, have been fully described and discussed in the literature;<sup>10,11,12</sup> therefore, in this section of the paper only those features essential to the present work are outlined.

The Organisational Control Game is a board wargame in which the player has the role of Divisional Commander. He believes that he is part of a team, consisting of himself, one superior commander and two subordinate commanders, and he believes that this team is playing interactively against a purposive enemy. In fact, the enemy, and all players other than the Divisional Commander, are played by the game controllers; their actions are entirely predetermined, but the player does not know this. The important consequence of this deception is that different players can be led through the same sequence of pre-planned events in the game, so that it is possible to compare the behaviour of different players in exactly the same sequence of objective circumstances.

Seven serving army officers (labelled A, B, . . . G) played the game. At the time of play, all held the rank of Major. Each player played twice, on different occasions, once in a defence scenario (Scenario 1), and once in an advance-to-contact scenario (Scenario 2). The two scenarios depict quite different military situations; one question of interest in the work was whether this would lead to consistent differences in the decision-making behaviour of the players across the scenarios.

During the course of the game, in each time period, a player was required to make reports to his superior commander and to issue directives to his subordinate commanders. At less frequent intervals he was required to make comments to the game controllers (as researchers) on his decision-making in the game. All such communications were recorded on tape, and transcripts of these recordings were used as texts from which the games were analysed.

### COGNITIVE MAPPING

Traditionally, scientists (and therefore by implication most of those involved in O.R.) have been concerned with the real, objective world, where phenomena can be observed and measured. However, human decision processes never take place in this objective world. All that can be observed there are some of their effects. Human decision processes always take place within the subjective world of the individual decision-maker. Cognitive mapping offers a window on this subjective world. It provides a means of representing the way in which a decision-maker models his decision-making environment, in terms of the concepts he himself uses. The technique has been used to examine the causal belief systems of decision-makers in a variety of contexts, including O.R.

A cognitive map is derived from a *text*, a record of the verbal or written statements of an individual. From the text are identified the concepts used by the individual and the *causal relationships* between them. Only two types of relationship are considered in the application discussed here: positive and negative. A *positive* relationship occurs when a change in a predecessor concept causes a similar change in the successor: an increase (decrease) in the first causes an increase (decrease) in the second. With a *negative* re-



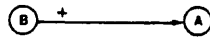


FIG. 1. Two concepts linked by a positive relationship. Text: "... morale is high, as they are advancing with little opposition ...". Concepts: A—high morale; B—unopposed advance.

relationship, an increase (decrease) in the predecessor causes a decrease (increase) in the successor.

An example will serve to clarify these definitions. Discussing a particular friendly unit, one player stated that "... morale is high, as they are advancing with little opposition ...". Two concepts may be identified in this text: 'high morale', and 'unopposed advance'. These concepts are linked by a positive relationship, since the player believes that an 'unopposed advance' causes or promotes 'high morale'. The two concepts, and the relationship between them, are shown graphically in Figure 1.

The coding procedure is generally more complicated than in this example. Relationships between concepts are often obscure and sometimes ambiguous, but with care large cognitive maps containing many concepts and relationships may be coded from a text and pictured graphically. Coding techniques are described in detail by Wrightson.<sup>13</sup>

In the gaming application described here, the texts used were the transcripts of the players' communications in the game. Maps were coded from these texts in two phases. The initial texts of each player—his first directives to subordinates and his first reports to his superior and to the game controllers—were used to construct maps of each player's initial perceptions of the game environment. Subsequent texts were then used to revise each player's map as his perceptions altered with the progress of time and the development of the game.

## RESULTS

Coding the initial texts yielded 14 cognitive maps (from the seven players in each of two scenarios). Broadly, analysis of the maps indicates marked differences between the players, yet, for any player, marked similarities across the scenarios. Figures 2 and 3 show the extreme cases for Scenario 2 (Players A and C), while Figure 4 shows Player C's map in Scenario 1. Only a limited description of the meaning of these maps is given with the figures, since in this paper the analysis is restricted.

### Concepts and relationships

The most obvious difference between players is the number of concepts identified in their maps. This ranges from 10 (Player C) to 26 (Player F) in Scenario 1, and from 8 (Player C) to 23 (Player E) in Scenario 2 (see Table 1). However, for four of the seven

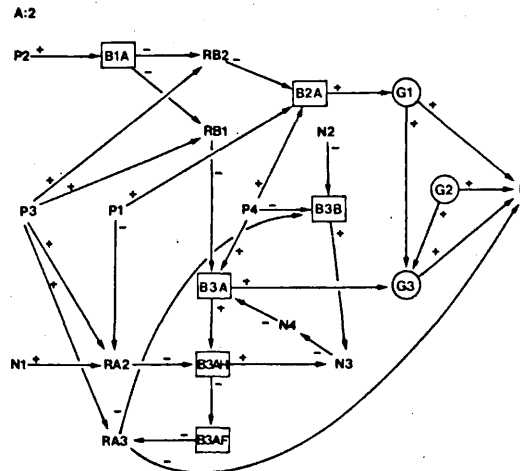


FIG. 2. Cognitive map of Player A in Scenario 2. Boxes indicate actions open to the friendly player; circles indicate his goals; U indicates his utility; other concepts are actions open to the enemy (initial character R), external factors specific to North and South areas (initial characters N and S) and other external factors (initial character P).

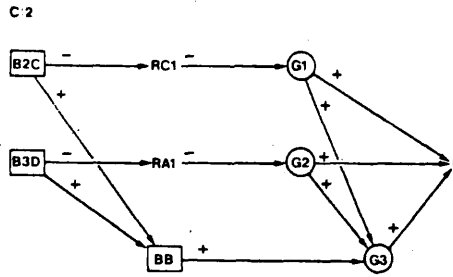


FIG. 3. Cognitive map of Player C in Scenario 2. The symbols are as described for Figure 2.

players, Players A, B, C and G, the numbers of concepts in the two scenarios is quite similar, despite the fact that they involve very different military situations. (In Scenario 1, the defence, the player is opposed by a superior enemy force, and he has little freedom of action; on the other hand, in Scenario 2, the advance-to-contact, the player has the initiative and military superiority.) Some of the decision-makers, therefore, appear to be modelling their decision-making environment using a relatively constant number of concepts. This number apparently has little to do with the objective situation and may represent some limit to the quantity of different concepts the decision-maker feels he can usefully cope with at any one time. This limit may vary quite widely between decision-makers. There is no suggestion, however, that the *quality* of decision-making exhibited by those players with fewer concepts is in any way inferior, at least in these games.

The number of causal links in the maps follows a similar pattern to that shown by the number of concepts. It is common to calculate the map density,  $D$ , defined as the number of links observed,  $L$ , divided by the maximum number of links possible between  $N$  concepts:

$$D = L/N(N - 1).$$

Numbers of links and densities are also shown in Table 1. With only three exceptions, the densities of all the maps lie between 0.05 and 0.1. The three maps with the largest densities are the three smallest maps—in fact, they are exceptionally small. The concepts in these small maps tend to be of central importance to the situation, and the decision-makers acknowledge many causal relationships between them. The larger maps all contain similar central sections, but there are also many more peripheral concepts, of limited influence and affecting only a small fraction of the others. The more concepts a decision-maker uses, the less each concept is likely to have a direct causal effect on any other concept. As might be expected, the central sections of different players' maps are frequently similar for a given scenario, with many concepts and links appearing in the same or slightly altered form in several players' maps, although this is not always the case.

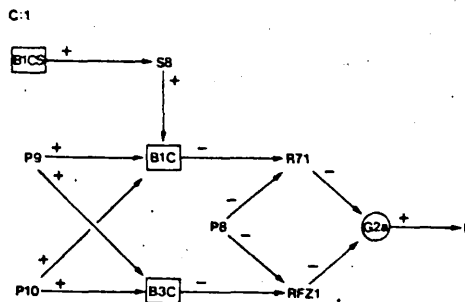


FIG. 4. Cognitive map of Player C in Scenario 1. The symbols are as described for Figure 2.

TABLE 1. CONCEPTS, LINKS AND DENSITIES OF THE COGNITIVE MAPS

Scenario	Numbers of concepts, <i>N</i>		Number of links, <i>L</i>		Density, <i>D</i>		
	1	2	1	2	1	2	
A	22	21	38	33	0.082	0.079	
B	22	22	26	29	0.056	0.063	
C	10	8	13	12	0.144	0.214	
Player	D	23	13	32	16	0.063	0.103
	E	16	23	21	39	0.088	0.077
	F	26	18	39	24	0.060	0.078
	G	21	19	27	24	0.064	0.070

*Differences in detailed interpretation*

Often, concepts that one player regards as fundamental do not appear on another player's map, and from time to time the same concepts can be found to be related in entirely different ways by different players. Figure 5 shows a specific example from Scenario 2. Case 1 illustrates the reasoning of most players: a friendly holding or delaying action will limit an enemy advance; if allowed to proceed, the advance would be detrimental to friendly goals; but a determined advance by the enemy will make the holding action more difficult. Case 2 shows the contrasting reasoning of Player G: a friendly holding action will cause the enemy to manoeuvre; this forced manoeuvre will ultimately put the enemy at a disadvantage and so will have a positive effect on the attainment of friendly goals. The positive as opposed to negative links in Case 2 reflect the fact that only this player felt that he had the initiative in this situation, a point he makes clearly in the text. This is a particularly fine example of the way in which one decision-maker can see the same objective situation in a completely different way from another.

*Confidence of the players*

It became clear on reading the texts that one of the main differences in the way players approached a given scenario was in the degree of confidence they displayed in their ability to achieve their goals. A subjective rating was made by the experimenters of the confidence of each player in each scenario (Table 2). This rating was compared with two quantitative measures, based on the cognitive maps and the texts.

In each cognitive map, paths of relationships leading from actions, events or circumstances to either an increased goal attainment or a decreased goal attainment were identified. Each goal was considered separately. Paths leading to a particular goal begin either with an action the decision maker can take or with some external factor over which the decision-maker has no control, but to which he usually assigns a value. In the first case, where a path begins with an action, it is assumed that the decision-maker will always act in order to increase his goal attainment—thus such a path is always goal-increasing. In the second case, paths may be goal-increasing or goal-decreasing, depending on the value

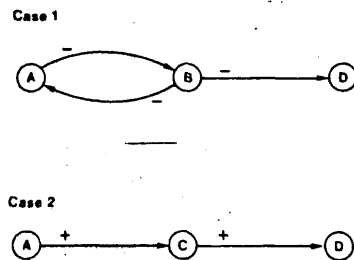


FIG. 5. Alternative views of the same objective situation. Concepts: A—friendly holding action; B—enemy advance; C—enemy manoeuvre; D—friendly goals.

TABLE 2. SUBJECTIVE RANKS, GOAL-INCREASING (G.I.) SCORES AND RANKS AND DESCRIPTION-SPECULATION (D.S.) SCORES AND RANKS FOR THE TWO SCENARIOS

Player	Order of play	Scenario 1				
		Subjective rank	GI score	GI rank	DS score	DS rank
A	1,2	5	0.493	4	0.602	4
B	2,1	4	0.294	7	0.456	7
C	1,2	2	0.714	3	0.821	1
D	2,1	6	0.417	5	0.560	5
E	1,2	3	0.875	1	0.647	3
F	2,1	7	0.311	6	0.549	6
G	2,1	1	0.800	2	0.661	2

Player	Order of play	Scenario 2				
		Subjective rank	GI score	GI rank	DS score	DS rank
A	1,2	6	0.300	6	0.622	4
B	2,1	3	0.231	7	0.660	3
C	1,2	2	1.000	1	0.779	1
D	2,1	4	0.778	2	0.587	5
E	1,2	7	0.538	5	0.557	6
F	2,1	5	0.769	3	0.537	7
G	2,1	1	0.722	4	0.762	2

assigned to the initial factor and on the number of negative relationships on the path to the goal. The number of goal-increasing paths for all goals divided by the total number of goal-increasing plus goal-decreasing paths defines a goal-increasing score (Table 2).

Intuitively, it seems reasonable to suggest that a high goal-increasing score might be associated with a high degree of confidence, since it indicates that the player perceives a relatively high number of factors leading to increased as opposed to decreased goal attainment. The goal-increasing score measures two aspects of the decision-maker's view of his decision environment. It reflects both his perceptions of external factors affecting his goals and also his freedom to act without external factors imposing constraints on him.

A content analysis of each text was carried out to count the numbers of words of 'description' related to the current game situation and 'speculation' concerned with possible future developments. A *description-speculation score* was defined as the number of words of description divided by the total number of words of description plus speculation (Table 2).

From Table 2 it can be seen that the players who were subjectively judged as confident are characterised by both high goal-increasing scores and high description-speculation scores. The rankings based on subjective judgement, the goal-increasing scores and the description-speculation scores show a high level of agreement: using a test based on Kendall's coefficient of concordance,<sup>14</sup> the agreement is significant for Scenario 1 at the 1% probability level ( $S = 212$ ) and only just fails to achieve significance for Scenario 2 at the 5% probability level ( $S = 152$ , critical value of  $S$  at 0.05 level = 157.3). Players with low description-speculation scores are those players who are thinking ahead. They consider more external factors which might adversely affect the attainment of their goals, and so they tend to have lower goal-increasing scores and they appear less confident.

#### Sequence effects

A marked sequence effect is detectable when scores on the first and second games played by a player are compared. Goal-increasing scores and description-speculation scores are generally lower on the second occasion an individual plays, irrespective of the scenario which he plays first. This suggests that on their second encounter with the game, the players may be taking a more serious and professional approach. By the time they played for the second time, the officers may have realised that the game was very similar in many ways to the Command Post exercises with which they were familiar, that it was

TABLE 3. SEQUENCE EFFECTS ON SUBJECTIVE ESTIMATES OF THE PLAYERS' CONFIDENCE

	First scenario played	Second scenario played
Confident players	C. E. G	C. G
Less confident players	A. B. D. F	B
Least confident players		A. D. E. F

not an 'amateur' game, and that it posed some interesting and non-trivial military problems. Subjective classification of players' confidence in the scenarios also suggests that a more cautious view is being taken on the second play (Table 3).

#### Dynamic aspects

As the games develop through time, the military situations change, the players' perceptions change, and in consequence the maps are amended. The rate at which the maps must be amended varies from player to player.

The way in which the cognitive maps of the players are amended indicates a tendency for some sections to become elaborated and considered in finer levels of detail, while other parts are lost from consideration. Which sections are expanded and which are lost vary for different players. Arguably, both scenarios develop towards crisis situations; the focussing by the players on subsections of the maps as time passes may be related to the general hypothesis that, in crisis, decision-makers tend to concentrate on certain aspects of their decision-making environment at the expense of others (see, for example, specific hypotheses cited by Milburn<sup>15</sup>).

## DISCUSSION

The results presented here show that the perceptions individuals have of a common decision-making environment differ in noticeable and sometimes surprising ways. In consequence, their decision-making behaviour might be expected to differ too. The differences that have been observed fall into distinct classes, related to the size and complexity of the cognitive maps, their detailed interpretation, the players' confidence and anticipation of the future, and the way in which maps are altered as time progresses. These differences occur precisely because the decision-makers are individuals.

Do these differences, though, indicate real differences between the decision-makers? In other words, how valid is a cognitive map as a representation of real cognitive processes? This question can be approached at two levels: first, the validity of a player's *assertions* as a statement of his true beliefs; second, the validity of a cognitive map as a model of the *mechanism* of thinking.

Concerning the validity of the players' assertions, an essential feature of the game methodology was the way in which most of the texts which were analysed were generated as a natural part of play. In fact, as far as the players were concerned, the game could not proceed without them. There was very little prompting by the experimenters, and thus the texts should reflect what the players themselves thought important. Additionally, the players were used to taking part in exercises of this kind as part of their normal peace-time military role. There is no reason to believe that they engaged in deliberate deception about their true views and aims. There may, however, have been unintentional biases in that players may tend not to articulate concepts and relationships which they have considered explicitly but which they have assessed to be in some way irrelevant, inconsequential or 'illegitimate'. To detect this kind of effect, other independent evidence is needed.

Regarding the validity of cognitive mapping as a way of modelling the mechanism of decision-making, there is some doubt. Certainly, in some cases, more sophisticated treatment will be required to provide adequate models of the structure and complexity of thought. However, despite the simplicity of its approach, cognitive mapping can provide a useful analogy, enabling predictive inferences to be drawn about the outcomes of decision-making processes.

Irrespective of questions of validity, it is clear that the cognitive maps described here have elicited features of the decision-making behaviour of individuals which are both consistent and capable of interpretation. The maps of particular players show marked similarities across the two scenarios, indicating that meaningful components of an individual's *decision-making process* are being detected, and these components are independent of the particular decision-making environment.

The experience of the authors, coupled with reports of related work, indicates that cognitive mapping and allied techniques can play a number of useful roles in Operational Research. A cognitive map may be viewed as an external model of a decision process, a model which can present the complex ideas and interrelationships perceived by a decision maker in a concise, tangible and manageable form. In this framework, a cognitive map is a means of *communication*, not only for the transfer of perceptions and ideas between different decision-makers, but also for aiding an individual decision-maker to elucidate his own perceptions.

One particular form of this communication is the process of *problem negotiation* between client and analyst described by Eden *et al.*<sup>9</sup> The use of cognitive maps to facilitate communication can also be envisaged in the *teaching* of decision-making skills and for *assessing* the ways in which decision-makers acquire problem-solving strategies as they learn about a new problem and its environment.

It is also felt that cognitive mapping has a role to play in *decision-making research*, as presented in this paper. The work that has been described here is essentially preliminary. It was an attempt to assess whether the combination of two comparatively new techniques—research gaming and cognitive mapping—was a worthwhile enterprise. The results have indicated that this joint approach can be a fruitful one.

It is appropriate to end with a warning. The experiment that has been described here was limited in size and scope. The players in the game were special, in that they were used to accepting a game world as one to be taken seriously. Nevertheless, although it cannot be claimed that the results obtained can be generalised, they do form a basis from which hypotheses can be generated for more extended experiment and testing.

#### ACKNOWLEDGEMENTS

The authors wish to thank: the seven military officers who played the game; the Defence Operational Analysis Establishment, West Byfleet, Surrey for its cooperation; Dr Philip V. Johnson, now with the Admiralty Surface Weapons Establishment, Portsmouth, who was one of the game controllers at Royal Holloway College; Professor M. R. C. McDowell and Mr K. C. Bowen (formerly Head of Research, Defence Operational Analysis Establishment), both of Royal Holloway College, University of London, Egham, who provided continuous encouragement and stimulation; and Ms Margaret Dixon, who prepared the transcripts of the games to a uniformly high standard from recordings of rather variable quality. The games were played at Royal Holloway College, and the work was supported in part by the Ministry of Defence under Contract Number ER3/9/4/2088/012 to Professor M. R. C. McDowell. Versions of this paper have been presented at the Fourth European Congress on Operations Research, Cambridge, U.K., 22–25 July, 1980, and the Young O.R. Conference, Coventry, U.K., 2–4 December, 1980.

#### REFERENCES

- <sup>1</sup> R. L. ACKOFF (1979) The future of Operational Research is past. *J. Opl Res. Soc.* 30, 93–104.
- <sup>2</sup> K. J. RADFORD (1978) Decision-making in a turbulent environment. *J. Opl Res. Soc.* 29, 677–682.
- <sup>3</sup> M. R. DANDO, A. DEFRENNE and R. G. SHARP (1977) Could O.R. be a science? *Omega* 5, 89–92.
- <sup>4</sup> R. AXELROD (Ed.) (1976) *Structure of Decision*. Princeton University Press, New Jersey.
- <sup>5</sup> G. M. BONHAM, M. J. SHAPIRO and T. L. TRUMBLE (1979) The October war. *Int. Stud. Q.* 23, 3–44.
- <sup>6</sup> J. A. HART (1977) Cognitive maps of three Latin American policy makers. *World Politics* 30, 115–140.
- <sup>7</sup> L. L. ROOS JR and R. I. HALL (1980) Influence diagrams and organizational power. *Admin. Sci. Q.* 25, 57–71.
- <sup>8</sup> C. EDEN and S. JONES (1980) Publish or perish?—A case study. *J. Opl Res. Soc.* 31, 131–139.

*J. H. Klein and D. F. Cooper—Cognitive Maps of Decision-Makers*

- <sup>9</sup>C. EDEN, S. JONES and D. SIMS (1979) *Thinking in Organizations*. Macmillan, London.
- <sup>10</sup>D. F. COOPER (1979) The Superior Commander: a methodology for the control of crisis games. *J. Opl Res. Soc.* **30**, 529-537.
- <sup>11</sup>D. F. COOPER and J. H. KLEIN (1980) Board wargames for decision making research. *Eur. J. Opl Res.* **5**, 36-41.
- <sup>12</sup>D. F. COOPER, J. H. KLEIN, M. R. C. MCDOWELL and P. V. JOHNSON (1980) The development of a research game. *J. Opl Res. Soc.* **31**, 191-193.
- <sup>13</sup>M. T. WRIGHTSON (1976) The documentary coding method. In *Structure of Decision* (R. AXELROD, Ed.), Appendix 1. Princeton University Press, New Jersey.
- <sup>14</sup>Sir M. KENDALL (1975) *Rank Correlation Methods*, 4th edn. Griffin, London.
- <sup>15</sup>T. W. MILBURN (1972) The management of crisis. In *International Crises* (C. F. HERMANN, Ed.), Chap. 11. Free Press, New York.