The Development of a Model of Multitasking Behaviour

by

Jonathan Miles Benson

University College London

Submitted for the degree of Doctor of Philosophy in the
University of London

December 1990
Abstract

The thesis begins with a consideration of what *multitasking* might mean, and derives a definition to be used in the rest of the thesis. The goals of the thesis are laid out - to develop a model, in the context of Cognitive Ergonomics, which has the properties of generality and utility. A consideration of this context along with the nature of models leads to the proposal of a heuristic strategy for developing a model with the desired properties.

A prototype model is proposed based on a consideration of literature drawn from Psychology, Computing, AI and Operations Research. The developmental strategy is then implemented through three successive iterations. Each iteration is observation of the multitasking behaviour associated with a real job, using concurrent verbal protocol techniques. The choice of jobs can be understood in terms of the development strategy. All the jobs observed possess certain key attributes, but crucially differ in their other attributes. The jobs observed are cooking, computer operations and railway signal control.

The final model represents multitasking as the allocation of resources to tasks over time. In this model, the person has two roles - as a controller and also as an effector. In the latter case, the person is modelled as a pair of resources.

The plausibility of the recruitment of the knowledge in the model for the purposes of design is then evaluated experimentally, providing a partial measure of the success of the developmental strategy.

The thesis ends with a summary and assessment of the model and of the strategy, together with a review of some additional areas of literature relevant to its final state.
## Contents

<table>
<thead>
<tr>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abstract</td>
<td>2</td>
</tr>
<tr>
<td>Contents</td>
<td>3</td>
</tr>
<tr>
<td>List of Figures</td>
<td>4</td>
</tr>
<tr>
<td>List of Tables</td>
<td>6</td>
</tr>
<tr>
<td>Acknowledgements</td>
<td>7</td>
</tr>
<tr>
<td>Chapter 1 Introduction</td>
<td>8</td>
</tr>
<tr>
<td>Chapter 2 The Method</td>
<td>18</td>
</tr>
<tr>
<td>Chapter 3 The Initial Model</td>
<td>41</td>
</tr>
<tr>
<td>Chapter 4 Study 1: Cooking</td>
<td>61</td>
</tr>
<tr>
<td>Chapter 5 Study 2: Computer Operators</td>
<td>92</td>
</tr>
<tr>
<td>Chapter 6 Study 3: Railway Signalmen</td>
<td>120</td>
</tr>
<tr>
<td>Chapter 7 Using the Model in Design</td>
<td>146</td>
</tr>
<tr>
<td>Chapter 8 Summary and Assessment</td>
<td>196</td>
</tr>
<tr>
<td>Bibliography</td>
<td>213</td>
</tr>
<tr>
<td>Appendix A</td>
<td>220</td>
</tr>
<tr>
<td>Appendix B</td>
<td>232</td>
</tr>
<tr>
<td>Appendix C</td>
<td>242</td>
</tr>
<tr>
<td>Appendix D</td>
<td>254</td>
</tr>
</tbody>
</table>
List of figures

Fig. 2.1. A Person as a Black Box.
Fig. 2.2. Diagram comparing a Person and a Model as Black Boxes.
Fig. 2.3. Schematic representation of the relationship of a model to different scenarios.
Fig. 3.1. Diagrammatic representation of the structure of the initial model.
Fig. 4.1. Plan diagram to show the general layout of the kitchen used in the study.
Fig. 4.2. Schematic representation of the Model's structure.
Fig. 5.1. Representation of the potential roles of the person in a (multitasking) job.
Fig. 5.2. The Structure of the Model
Fig. 6.1. Track and Platform plan of the terminus station controlled by the Signalman.
Fig. 6.2. Plan of the organisation of the station control room.
Fig. 7.1. Reproduction of the main screen of the carwash simulation.
Fig. 7.2. Reproduction of the screen used to support the Problem Representation.
Fig. 7.3. Reproduction of the screen used to support the Solution.
Fig. 7.4. Diagram of the experimental setup.
Fig. 7.5. Graph of the unadjusted time data for all groups.
Fig. 7.6. Histogram showing the average proportion of time spent gazing in five directions, for all four conditions, trial 1.
Fig. 7.7. Graph showing the average relative strategy time over four trials for the Base and PR conditions.
Fig. 7.8. Graph showing average relative free time (as a proportion of strategy time), over four trials for the Base and PR conditions.
Fig. 7.9. Graph showing average relative perceived efficiency, over four trials for the Base and PR conditions.
Fig. 7.10. Graph showing average relative perceived difficulty, over 4 trials for the Base and PR conditions.
Fig. 7.11. Average percentage of time spent looking in the direction of the PR screen in the PR condition.
Fig. 7.12. Graph showing the average relative strategy time over 4 trials for the Base and Sol conditions.
Fig. 7.13. Graph showing average relative free time (as a proportion of strategy time), over four trials for the Base and Sol conditions.
Fig. 7.14. Graph showing average relative perceived efficiency, over 4 trials for the Base and Sol conditions.
Fig. 7.15. Graph showing average relative perceived difficulty, over 4 trials for the Base and Sol conditions.
Fig. 7.16. Average percentage of time spent looking in the direction of the Sol screen in the Sol condition.
Fig. 7.17. Graph showing the average relative strategy time over 4 trials for the Base and PR + Sol conditions.

Fig. 7.18. Graph showing average relative free time (as a proportion of strategy time), over four trials for the Base and PR + Sol conditions.

Fig. 7.19. Graph showing average relative perceived efficiency, over 4 trials for the Base and PR + Sol conditions.

Fig. 7.20. Graph showing average relative perceived difficulty, over 4 trials for the Base and PR + Sol conditions.

Fig. 7.21. Average percentage of time spent looking in the direction of the PR and Sol screens in the PR + Sol condition.

Fig. 7.22. Graph showing the average relative strategy time over four trials for all four conditions.

Fig. 7.23. Graph showing average relative free time (as a proportion of strategy time), over four trials for all conditions.

Fig. 7.24. Graph showing average relative perceived efficiency, over four trials for all conditions.

Fig. 7.25. Graph showing average relative perceived difficulty, over four trials for all conditions.
List of tables

Table 4.1. Allocation of subjects to Job 1 or Job 2.
Table 4.2. Table of conventions and their meanings as used in all the protocol extracts in this and subsequent chapters.
Table 4.3. Summary of the BPs observed in this study, together with those from the literature review of Chapter 3.
Table 5.1. Summary of the BPs observed in this study, together with the complete set observed so far.
Table 5.2. The role of the person in a Job concerning more than one object.
Table 6.1. Summary of the BPs observed in this study, together with the complete set observed so far.
Table 7.1. The main JA differences between the carwash and previously observed jobs
Table 7.2. Table of processes in the different levels of service.
Table 7.3. Times and resource requirements of the various processes.
Table 7.4. The experimental design, showing screens used in the different conditions on trials 1 to 4.
Table 7.5. The configuration of a job in terms of cars.
Table 7.6. Strategy time data for Base and PR conditions.
Table 7.7. Free time data for Base and PR conditions.
Table 7.8. Perceived efficiency data for Base and PR conditions.
Table 7.9. Perceived difficulty data for Base and PR conditions.
Table 7.10. Direction of gaze data for PR condition on trials 2 & 3.
Table 7.11. Strategy time data for Base and Sol conditions.
Table 7.12. Free time data for Base and Sol conditions.
Table 7.13. Perceived efficiency data for Base and Sol conditions.
Table 7.14. Perceived difficulty data for Base and Sol conditions.
Table 7.15. Direction of gaze data for Sol condition.
Table 7.16. Strategy time data for the combined PR + Sol and Base conditions.
Table 7.17. Free time data for the combined PR + Sol and Base conditions.
Table 7.18. Perceived efficiency data for the combined PR + Sol and Base conditions.
Table 7.19. Perceived difficulty data for the combined PR + Sol and Base conditions.
Table 7.20. Direction of gaze data for the PR and Sol screens in the combined PR + Sol condition. Trials 2 and 3.
Acknowledgements

My greatest debt is to my two supervisors. Firstly to Andy Whitefield for his exemplary advice and support: I can’t imagine a more suitable supervisor. Secondly to John Long, for his encouragement and advice, but also for taking the risk on me in the first place.

Paul Reed and Stuart Bishop provided many stimulating discussions about computing as well as practical help with the experimental simulation.

Thanks are due to HM Systems PLC for the free loan of computer equipment used in the experiment reported in Chapter 7.

I could not have done it without the cooperation of the subjects:

Neal, Ross and the other operators of UCL’s EUCLID computer.

Dennis and the other signalmen on his shift in the signalbox at “a large terminus station in the north of England”.

The many friends, colleagues and strangers who cooked for me or who washed my imaginary cars.

Finally, the staff and other students at the Ergonomics Unit, University College London, provided a stimulating and supportive environment in which to work. I thoroughly enjoyed my time there.

This thesis is dedicated to the memory of Big John, who saw it started, which was enough.
1 Introduction

Prologue

This chapter outlines the area of interest which will be pursued in the thesis, and why
this is thought to be worthwhile. It provides an outline of the contributions of each of
the following chapters, and an initial conception which allows the derivation of a
definition of multitasking.

Contents

1. Overview

2. The Layout of the Thesis

3. Towards a definition of Multitasking

1. Overview

This thesis is concerned with the development of a model of multitasking behaviour.
The notion of ‘multitasking’ will be discussed in greater detail below, but the
following simplified example should serve to indicate the broad area of interest and the
importance of enhancing our current understanding of such behaviour.

There are many examples of jobs which could be thought of as entailing multitasking
behaviour. An easily comprehended example is that of a secretary or personal
assistant. Such a person’s job would typically include the preparation of documents
(letters etc), the management of files, receiving visitors and answering the telephone.
Instances of these tasks would constitute the multi-tasks, but such a collection takes on
an additional dimension when it is necessary to interleave these tasks. So, for
example, it might be necessary to pause in the middle of writing a letter to answer the
telephone, or the writing of a letter might take several days to complete because the
necessary information must be assembled and thus be carried out discontinuously with
other tasks. The list of such examples is extensive and in this case probably easily
imagined.
The point is that there has been much effort expended in trying to understand the behavioural phenomena associated with at least some of the individual tasks alone (e.g. Cooper 1983 on typing or Beattie and Barnard 1979 on telephone dialogues), but little or no research has been directed towards attempting to document any additional phenomena emerging as the result of the combination and interleaving of a number of such tasks.

The number of jobs entailing interleaving of tasks is already large. In an industrial society in the process of computerisation and automation, this might be expected to increase; the supposed reduced workload associated with the tendency towards supervisory roles could lead to the combining of responsibilities and thus tasks for reasons of economy. The effects of inefficiency are likely to be expensive in human (i.e. job satisfaction) as well as purely economic terms.

The lack of knowledge concerning the management or coordination of several tasks at once, coupled with its high incidence in the world of work and the potential impact of inefficiency in one form or another constitutes the rationale for the present study. Given this rationale, the context of the study must also be understood.

Whereas Psychology itself might be justly characterised as being interested in understanding behaviour as an end in itself, Cognitive Ergonomics (indeed Ergonomics in general) could be characterised as aiming to understand behaviour with a view to employing such knowledge to improve performance in the workplace. Although the possible applications of knowledge of multitasking or interleaving behaviour in the case of the secretary might seem trivial, it must be remembered that improvements in job satisfaction are a perfectly legitimate goal, especially considering the number of secretaries there are. It is perhaps easier to relate to the case of, for example, the solo pilot or the shift leader in a nuclear power station control room, where any negative consequences of multi-task interference might be sufficiently costly to warrant ergonomic intervention.

Although the content of this thesis is largely psychological in nature, it is so from the perspective of cognitive ergonomics. As such it is concerned with behaviour associated with the workplace and will choose to observe behaviour in as natural a setting as possible, sacrificing the fidelity and control normally associated with reductionistic psychological experimentation in favour of potentially richer observational data (see Chapter 2 for a lengthier discussion of this).

The aim of the thesis will be to develop an understanding of multitasking behaviour. This understanding will take the form of a model, which will be general - to the extent of being abstracted from any particular instance of multitasking. However, developing
such a model is not readily prescribed. For this reason, the meaning of *general* in the context of models and cognitive ergonomics will be explored and a rationale for the development of such a model constructed. To the extent that the model developed in the course of the thesis satisfies the initial aim, then the rationale is tested.

2. The Layout of the Thesis

The project reported in this thesis was concerned with the development of a model of multitasking behaviour. As such, the product and the process of the project constitute two equally important contributions. It is for this reason that the format of the thesis is as an historical account of the development of the model. Only in this way could the developmental process be adequately documented.

There is one potential disadvantage to plotting such a course of evolution, which is that the thesis will have to contain the imperfect early versions of the model that failed to survive. It is difficult to write such an account without feeling the need to change these early imperfections, which form such an important part of the development history. An effort has been made to minimise this.

The rest of this section is an overview of the contents of the thesis, and then the remainder of this chapter is given over to a discussion and derivation of an initial definition of multitasking.

Chapter 2 provides a discussion of models and generality in the context of Cognitive Ergonomics, deriving a rationale for developing such a model. This then forms the basis for the rest of the work reported in the thesis.

Chapter 3 takes the initial definition of multitasking and the stated context of the research (Cognitive Ergonomics) and identifies several potentially relevant areas of knowledge. This knowledge falls largely, but not exclusively, under the heading of cognitive science knowledge. The literature associated with these areas is sampled and reviewed to produce a specification for an initial model. Such a model is then described.

The emphasis, in developing the model in this project, is that it should be based as much as possible on observational data. The initial model presented in this chapter is therefore only a prototype, constructed to facilitate the analysis of subsequent observational data. The literature review upon which it is based is therefore not exhaustive.

Chapters 4, 5 & 6 are three observational studies and thus three stages in the development of the model. Each study concerns a different aspect of multitasking, as demanded by the development rationale of Chapter 2. Chapter 4 is concerned with
multitasking behaviour in cooking, Chapter 5 with such behaviour in the periodic backup job of a mainframe computer operator, and Chapter 6 is an account of the multitasking involved in the job of a railway signalman in the signalbox of a busy terminus station.

Note: all the chapters concerned with the development of the model (i.e. chapters 3-6, inclusive) will all follow the same basic format. The rationale of Chapter 2 treats both a person and a model as a black box, producing a certain output behaviour for a given input behaviour. The specification of the particular instance of multitasking constitutes the Input; the observed, or in the case of the literature review of Chapter 3, inferred, phenomena being the Output. The general scheme for such a chapter will then be to present the Input, in terms of a specification of a particular instance of multitasking, followed by the details of the data acquisition, and then discuss the Output as the observed phenomena. This is followed by an account of a model proposed to behave in such a manner. In short, the sequence will be Input; Study Particulars; Output; Modelling.

The modelling process is not totally data driven in this way, and each model will have its inherent shortcomings. These will also be discussed as appropriate.

Chapter 7 provides a test of both the model and the rationale in that it attempts to demonstrate the recruitment of the knowledge in the model to the design process. The chapter presents an experimental study in which subjects played a computer game requiring them to multitask. The model is used in the design of variations to the interface, thus providing the experimental manipulations for different groups. The recruitment of the model in this way is judged to have been partially successful. This is discussed.

Chapter 8 is the final chapter of the thesis. It starts with a summary of the model which has been developed, in its final state. This is followed by an assessment of the model as a product, with respect to the initial aims of the project, and then also the process used to develop the model. The rationale allows that the definition of multitasking will change as the model develops. Consequently, there are areas of e.g. the psychological literature additional to those reviewed in Chapter 3 which are now relevant. These are reviewed as they apply to the model in its final state. It is apparent that there is a degree of convergent evolution in the development of this model. The relationship of the thesis to Cognitive Ergonomics in general is also considered.

3. Towards a Definition of Multitasking

The word ‘multitasking’ will be taken to be a computer science jargon term. It is used in the current context since as a compound of ‘multi’ and ‘task’ it adequately captures
the essence of the class of behaviour of interest - namely being concerned with more than one task, and also because there is no other suitable candidate term.

The purpose of this section is to provide an initial definition of multitasking. It is thought that this definition will be altered as the project progresses and the understanding of the behaviour it encompasses is advanced. However, an initial definition is essential as it enables at least the following:

- a choice of instances of behaviour to study
- the identification of relevant areas of existing knowledge - in the form of, for example, the psychological literature
- the identification of analogous systems (in the general sense) for the purposes of modelling

To produce such a definition, the intention is to start with a lay definition, and progress on to considering some of the implicit terms (such as task and job) more carefully. What will need to be made clear however, is what exactly constitutes a task or an activity such that several together may or may not be considered to be an instance of multitasking.

The lay definition of multitasking, which gave rise to the project, is easily expressed as "doing more than one thing at once". Everyday instances of such behaviour are plentiful, and include cookery and the job of a typical secretary, for example.

It will be shown that there is not as yet a single, recognised, suitable definition of a task. Fleishman (1975) in fact suggests that we should not attempt to look for the definition of a task, but rather we should accept that many are possible and we should therefore concentrate on developing an adequate classification scheme for the purposes. The intention here is to review some of the various ideas such that the relative strengths and weaknesses of the position to be adopted here can be better understood. In particular, there are some general issues and then issues specifically associated with using hierarchical descriptions.

The first point of potential confusion is the mismatch between what the performer sees as the task, and what the task really ought to be. Miller's (1967) definition of a task as "any set of activities, occuring at about the same time, sharing some common purpose that is recognised by the task performer" would be one example of the former. Examples of the latter could be found in e.g. training manuals. Fleishman and Quaintance (1984) make the distinction between external and intrinsic definitions of tasks. Intrinsic tasks are defined with respect to the person's perceptions and needs, and are therefore subjective. An external definition is concerned with what is imposed.
on the performer. In Reason’s (1987) scheme, mismatches between the intrinsic and external tasks would be apparent as ‘mistakes’.

Some authors, at least implicitly, use the term task to refer to the external definition, and the term activity for the intrinsic definition - see for example, Cypher’s (1986) discussion of the organisation of activities from an introspective point of view (discussed later in this thesis). For the present, the terms activity and task will be taken to be synonymous, and for definitional purposes will centre on an external definition.

A common conception throughout linguistics and ergonomics (e.g. Task Analysis) is that it is possible, and useful, to describe action hierarchically (specifically, Hierarchical Task Analysis, Annett et al 1971).

A comprehensive example from linguistics is that of van Dijk (1980). It is presented below for this reason, since some of its facets illustrate many of the important issues, and also as an example of the general class of hierarchical analyses. However, as will become clear, it is not itself able to provide a suitable basis for defining multitasking.

What distinguishes van Dijk is his rationale, the basics of which are as follows. He begins by setting out what is meant by an action. It is not necessary to spell out this derivation in detail, save to say that it starts by considering ‘events’ in the world, and covers intention, purpose and even omission (which is an action in itself for him). He uses the term ‘macrostructure’ to refer to a higher level description of action, although to understand this, it is necessary to see how he derives such structures - in other words, how he proposes that the hierarchical tree be formed.

The rules for the derivation of a macrostructure fall into two parts. Firstly there are the rules that allow an action sequence to be abstracted from a “bundle of human activity”. A sequence is a set of actions that are related to each other, for example, temporally or conditionally; in other words that is ordered by certain relations. Examples of such relations would be: connection, which refers to the way that one action forms the necessary precondition for its successor; coherence, which entails, for example, that there is a common time and/or place shared by successive action; and orientation, which is concerned with the relationship of the goal of a particular action to those of the rest of the actions in the sequence. It is maintained that a sequence is a set of actions which are all performed by an agent with the purpose of realizing a goal that is the result (or the consequence) of the last action of the sequence. This goal (of the last action of the sequence) is the sequential goal. It follows that, indirectly, this sequential goal is shared by all the previous actions of the sequence. This sequential goal represents a strong coherence dimension to a sequence of actions.
After the abstraction of the action sequences, the second stage is then the application of Macro Rules, which allow macroactions to be derived from such sequences. A macroaction is thus a coherent sequence of actions (or macroactions) taken as a whole. Examples of Macro Rules for the derivation of higher level macrostructures from lower level representations of activity would include Deletion, in which aspects are rendered implicit and taken for granted, and Generalisation. In a way, these rules are echoed by Sebillotte (1988) when she states that asking why takes one up a level whereas asking how takes one downwards in a hierarchy of goals.

It would seem reasonable to try to define multitasking as being concerned with a second or subsequent task before the first one has been completed. Unfortunately, this still leaves unanswered the question of what level to cut across the hierarchical tree, i.e. whether it is better to conceive of multitasking as a large number of tasks defined at a low level, or fewer tasks described at a higher level. It is important to determine an appropriate level of description since too high a level may leave much interleaving of interest implicit at lower levels, and conversely, too low a level may introduce too much detail, obscuring the phenomena of interest.

This problem has been raised numerous times in the literature, with differing solutions. Annett et al (1971) view the problem of what constitutes an appropriate level as “one of the most difficult features of task analysis” (p 6). Card et al (1983) go as far as to suggest that it is impossible to determine the appropriate grain of analysis a priori. Gagné (1977) is a little more optimistic, and suggests that the degree of specificity of task description should be solved by considering the purpose of the description. This, as will become clear, is the generally adopted solution underlying the development of so-called stopping rules. Duncan (1975), for example, suggests that we should not attempt to define levels of description, since psychology is not able to support this, rather we should proceed from a general to a specific level of description with some rule of when to stop.

The original Hierarchical Task Analysis (HTA) of Annett et al (1971) incorporated what has become known as the ‘P x C’ stopping rule. This particular application of HTA was concerned with supporting training decisions, and a measure of whether it is worth redescribing a task is given by multiplying the probability that the person will get it wrong by the cost of getting it wrong. In practice, the P x C rule can be difficult to use because of the requirements of expert judgement and knowledge of the tasks in question (Stammers et al, 1990). There have been other suggestions for rules. Hodgkinson and Crawshaw (1985), combined the P x C rule with stopping when the implications for training became clear. Similarly, Shepherd (1976) combined the P x C rule with stopping when a satisfactory training method was already available. Shepherd and Kontogiannis (1987) advocated stopping at the point which describes
the operations through which the operator interfaces with the system - for example, opening a valve or throwing a switch. All of these applications have been in the area of training. Piso (1981) adapted HTA for analysis of complex process control tasks. He rejected the P x C rule in this adaptation as being inappropriate, and instead, continued redescribing the tasks until they were clear to both the operator and analyst.

What is clear here, is that none of the above are appropriate for the present needs of defining a level at which to stop - there is no intention, for example, to try to estimate costs and probabilities or error associated with each task.

In principle it ought to be possible to establish a criterial use of van Dijk’s macro rules instead. In such a scheme, the application of particular rules would constitute a criterion for deciding that a particular sequence had meaning in its own right. In practice this would require a very comprehensive hierarchical expression of the task to begin with. It is thought that the cost involved is unlikely to justify what would, after all, only be an arbitrary criterial solution.

For the present purposes, the hierarchical levels of description of action will be equated with timescale. Tasks at lower levels of description, i.e. further down branches of the tree, are likely to be of shorter duration than those which they constitute - i.e. the macroactions. It is proposed that the duration of the task will serve as a measure of its level of description, and to use this as the basis of a simpler criterion based definition of multitasking. It is obvious that activities can be interleaved at all levels, from years and months down to seconds and possibly less for individual movements. Whether there is interleaving or true parallelism at this lowest level depends to some extent on which psychological conception one favours.

However, it should to be possible at least to approximate the appropriate timescale for the current definition given its objectives. Given the context of Cognitive Ergonomics, the appropriate timescale will be taken to be of the order of hours and minutes. (This is not to say that Cognitive Ergonomics cannot legitimately be concerned with very short timescale activities, such as those involved in speech recognition for example.) The upper bounds of this region derive from the fact that Cognitive Ergonomics is primarily concerned with the workplace which is usually occupied for periods of around eight hours at a time.

It is proposed to use the term job to refer to a collection of tasks, and thus jobs of interest will have durations in a region extending from several minutes to several hours. This would be in agreement with Annett et al’s (1971) use of the word, where the assignment of a set of tasks to a person defines his job. A job is thus a person oriented concept.
It is important to note that there need not be the same relationship of a job to its constituent tasks as there is of a task to its sub-tasks. The latter depends on a hierarchical concept of action, as described above, whereas the former can be merely a collection of such tasks. However, it could be a collection of tasks at any level of description - it is therefore important to be clear about what constitutes a separate task within a job.

It would seem sensible to consider a job to consist of those tasks at the highest level of description. This of course ensures that a job consists of a set of tasks each of which addresses an independent goal. If a job is defined as such a collection of tasks, then multitasking could more obviously be defined in terms of a sequence of actions in which the achievement of such independent goals is interleaved. An example of such a job might be that used by Hayes-Roth and Hayes-Roth (1979), where the tasks would be such things as “have lunch”, “do shopping”, “visit...”, etc (i.e. all fairly independent). Choosing instead to define the level description of tasks in terms of timescale has the advantage that aspects of other jobs, at a lower level, can be considered to involve multitasking. For example, a process controller whose job might involve the tasks of “optimise production” and “rectify problems”, may exhibit interesting behaviour in the interleaving of subtasks at a level below one of these tasks.

To summarize, multitasking will be taken to mean being concerned with another task before the current task has been completed, where a task is defined in terms of the time it takes to accomplish. Using van Dijk’s terminology, multitasking would be manifest as a sequence of activities which do not necessarily share a common sequential goal.

The consequences of such a definition of a task within a multitasking job will depend on whether such a degree of detail is demanded for the intended function of the model which will be developed. It should also be remembered that the definition of multitasking is expected to develop as the model is developed and that this definition is merely for the purposes of an initial starting point. The reader is also directed to the discussion of task and hierarchies in the context of the acquisition of skilled or routine behaviour, in Chapter 8.

This chapter has described the intentions of the research presented in this thesis, together with an outline of the way in which it will be presented. The second half of the chapter has been concerned with a discussion of how to define multitasking, and an initial definition has been proposed. This definition is weak and somewhat arbitrary. It will become apparent that it does, however, serve its purpose of enabling some initial study. Chapter 5, which is concerned with the second observational
study, describes a revised and much different definition which remains in force for the remainder of the thesis.
Chapter 2

The Method

Prologue

This chapter presents a discussion of the context of the research - i.e. Cognitive Ergonomics, and also of models in this context. It develops a (heuristic) method for the development of a model with certain desired properties. The role of ecological investigation is discussed.

Contents

1. Introduction

2. Cognitive Ergonomics
   - Science
   - Engineering
   - Craft

3. What is a model?
   3.1. Model-1
   3.2. Model-2
   3.3. A Model as a Black Box

4. Why NOT construct a model?

5. What makes a good model?
   5.1. Correctness
   5.2. As a Scientific device
   5.3. As an Engineering device
   5.4. The above expressed as criteria
   5.5. Can one model be good in all respects?
   5.6. Summary

6. Generality
   6.1. What does ‘generality’ mean?
   6.2. Generality of Scenario (E) represented by the model
   6.3. Generality of Purpose of the model.
1. Introduction

The aim of this chapter is to take the goal of the PhD (to develop "A General Model Of Multitasking") and to specify more carefully what it means, in the context of Cognitive Ergonomics, so that a suitable method for attaining the goal can be formulated.

Although there are many definitions of 'model', it will be shown that there is a common theme, and where there are slight alternatives, a choice will be made such that a working definition of a model can be arrived at. The properties of such a model will then be considered, including its advantages over alternatives, its goodness, and so on. In a section which then follows, the idea of generality with respect to such a model will be considered. When all these points have been addressed and a workable view of a model has been established, a method for optimizing the development of such a model, given certain other criteria, will be presented. Finally, the role of ecological validity and different approaches to data gathering will be addressed. However, first of all it is necessary to state and define the context for the whole discussion, i.e. Cognitive Ergonomics.
2. Cognitive Ergonomics

The broad discipline of Ergonomics is characterised by Long (1987b) as one which is concerned with optimising the relationship between people and their work. This is achieved through using scientific knowledge and techniques. Subsequently, Long makes the distinction between Traditional Ergonomics, which draws on the scientific knowledge of Biomechanics and Experimental Psychology and so on, and Cognitive Ergonomics (CE), whose main influence is Cognitive Science. He suggests that the impetus for this new type of ergonomics was not the rise of Cognitive Science alone, but also the advent of computer technology.

It would be easy to misunderstand Long (1987b) and conclude that CE only addresses situations in which there is a computer - a position which is not accepted in the current context. The paper is probably better understood as a description of the origins of CE. It is interesting to note that Long ends his discussion by recommending CE as an approach for psychology more generally (because of the former's emphasis on behaviour in real-world tasks etc). It is plainly absurd to conclude that he intended that psychology should only be concerned with the behaviour associated with interaction with computers.

A discipline has arisen which is exclusively concerned with the ergonomics of working with computers - ‘Human Computer Interaction’ (HCI), and much effort has been directed towards establishing the nature of this discipline and how it should conduct itself. Although, as has been said, the current context is firmly in cognitive ergonomics, it is worth considering the relationship of CE to HCI, for the purposes of drawing on its conceptions.

Firstly, it is maintained that CE and HCI are not the same thing - although HCI might be regarded as a subset of CE. Long and Dowell (1989) present a characterisation of HCI which divides the subject into those aspects concerned with modifying the behaviour of the person in the interaction with a computer (in a work system), which they term Human Factors (HF), and similarly, those aspects which are concerned with modifying the interacting properties of the computer - Software Engineering (SE). There is an obvious correspondence between CE and their HF, or at least that portion of CE knowledge which intersects with so called HCI knowledge. Given this relationship, there would appear to be no reason not to borrow from current conceptions of the HCI discipline for the purposes of understanding CE.

From the writings of Long and Dowell (1989) and Long (1986 & 1987b), CE will be taken to comprise 3 disciplines: Science, Engineering and Craft.
• Science

For the present purposes, there are two features of the science paradigm as described by Long which are important:

a) scientific knowledge is required to be public - i.e. explicit.

b) the aim of science is to establish truths concerning the world, where ‘truths’ would be descriptive, explanatory and predictive.

• Engineering-

Engineering will be considered, as in Long (1986) to be the knowledge and practice involved in the application of scientific knowledge to the development of artefacts. There are similarly two important features of the engineering paradigm:

a) its knowledge must be public.

b) the aim of engineering is to produce a working artefact. This contrasts with the science paradigm in which the aim is for a proposition to be true.

• Craft

Craft differs from the above paradigms in that its knowledge is not required to be public. Otherwise it can be thought of as similar to engineering as its aim is the development of artefacts.

The term model will be discussed and defined below, but for the present it can be said that it will be taken to be an example of public knowledge (contrary to some of the uses of the term, as for example in “user’s mental models”). As a consequence, it only has a relevance with respect to the paradigms of science and engineering; craft will not be considered further.

3. What is a model?

This section attempts to put together a useful definition of a model. This will be done in general, and not with specific reference to the paradigms of science or engineering. That discussion will be left to a subsequent section which addresses what makes a good model. From reading the literature it could be said that there is a consensus on one point at least - the lack of a precise definition. [For example Marx (1976) defines a model as a particular type of scientific theory - but then Hawking (1988) describes a theory as a kind of a model.]
To achieve its goal, this section will draw mainly on two frameworks - Warr (1980) and Long (1987a). A framework is not itself a model, but rather an analytic structure which “characterises the objects of its scope” (Long 1987a), providing a basis upon which different models can be discussed and compared, and is thus exactly what is needed for an abstract discussion of models. The intention is to show that there is also a common theme, upon which it is possible to build a more precise definition for the purposes of this project.

Warr (1980) distinguishes two uses of the word ‘model’ - which he conveniently labels model-1 and model-2.

3.1. Model-1.

In non-scientific terms, an example of a model-1 would be a scale model of an aeroplane or train. The important property of these models is that they are direct representations of the object (i.e. plane, train or automobile). In scientific terms, such models would include most of what are called theories. “The models are systematic conjectures about a part of reality and the attempt to simplify that reality within a network of thought” (Warr, 1980, p295). Warr points out that the way the word ‘model’ is used by psychologists is often to signify a limited or provisional theory.

3.2. Model-2

A model-2 would be an analogy, imported from some other sphere, used to assist in thinking about the unknown or unfamiliar. An example of such a model-2 in psychology would be the use of a telephone switchboard to model attention (see Marx 1976). If models-1 are direct then models-2 are indirect representations.

This definition of a model accords well with the definition of a model as a particular type of theory by Marx. Thus for Marx also, a model acts as a guide to research: it is an analogy borrowed for the purposes of proceeding ‘as if’ it were the case in the new domain, and with no intention of changing the analogy with respect to its source (in other words, work on attention using a telephone switchboard model was not used to change how telephone switchboards work). Marx also makes the point that the term ‘model’ is used outside this particular definition where it takes on a broader definition as a theory (i.e. a model-1). There is then a general agreement between e.g. Marx, Hawking and Warr: a model is both a type of theory and at the same time, a theory is a particular sort of a model.

There is an important difference, agreed upon by Warr and Marx (if one aligns the latter’s definition of a model as a particular type of theory with the former’s), in that a model-1 is formed by examining the reality and then generating the model to interpret
what is known. This is in direct contrast to a model-2, which uses an analogy borrowed from somewhere to suggest ways of looking at and interpreting the reality. This is what is meant by describing the models as direct and indirect.

The relationship between models-1 and models-2 is not fixed. Whilst they both have their own uses independently, a model-1 in one field of research can be borrowed to become a model-2 in another field. Similarly, the distinction can also blur in the other direction. “The process of interbreeding which takes place during sustained application of a model-2 to a prior theory often begets a framework in which the origin of the components is no longer clear” (Warr, 1980, p299). The same structure may switch its class membership and its classification as a model-1 or a model-2 is really dependent on its function at the time.

Having established that there are two ways of looking at a model, and that this is broadly a consensus position, the need now is to consolidate this into a more precise definition. To this end, the following characterisation of models offered by Long (1987a) is introduced.

Long defines a model as a representation of an entity:

A model ‘M’ is a representation ‘R’ of an entity ‘E’, which he writes as:

M -> R(E)

The entity E can be thought of as having attributes (Ai...An), and the representation R of having attributes (ai...an). The set (ai...an) is a systematic reduction of the set (Ai...An) - in other words, the model only has some of the attributes of the thing being modelled. The reduction is systematic, and reflects the purpose of the model.

The full characterisation of models, by Long, is as follows:

“A creator C creates with purpose P a model M expressed as a representation R of an entity E for a utiliser U”.

Note that Long’s characterisation of a model is independent of whether the model is classified as a model-1 or a model-2.

(This characterisation will be used again below when the question of what constitutes a model is addressed, and also how a model may be general.)

The final contribution to a derived definition of a model is that of the Oxford English Dictionary. This provides admittedly a non-scientific, non-specific definition, but makes explicit one important feature. The definition is “[a] simplified description of a system to assist in calculations and predictions”. The important word here is
'system', and the important attribute of a system will be taken to be that it embodies the interactions of its constituent parts - it has a certain measure of gestalt.

To summarize the definition of a model which will be used in the rest of this thesis:

A model is a system, with certain attributes, which it has in common with another system. The model is thus said to represent that system. The form that a model takes will be related to its intended function, and its classification as either a model-1 or a model-2 could be taken to be indicative of its state of development (models-1 being more direct representations of the target entity and thus, for some purposes, more desirable).

3.3. A Model as a Black Box

It useful to consider a model as a ‘black box’ as well as all of the above. The apocryphal black box is a device which accepts an input and produces some output based on this. The point is that at this level of description, the mechanism by which it accomplishes this is left unspecified.

Before presenting a model as a black box, it is worth presenting a person as a black box:

![Diagram of a Person as a Black Box]

Fig. 2.1. A Person as a Black Box.

Notice that the stress would be on the lack of a 1:1 relationship between any one feature of the scenario and any one behavioural phenomenon. This is the point that was introduced above concerning viewing models as systems - the person is viewed as a system, so the output need not be related to the input in a simple manner.

The above diagram (Fig. 2.1.) can be modified to introduce a model (of the person - ie the person is Long’s Entity):
A Model and a Person viewed as a 'black box' process, taking the attributes of a Scenario as input and producing Behavioural Phenomena as output. The more nearly Phenomena M equal Phenomena P the better the model could be said to be.

Fig. 2.2. Diagram comparing a Person and a Model as Black Boxes.

Note that Fig 2.2. is a very simplified representation of a model; for example there is no stated purpose or utiliser. Also it is not the intention to imply a strict modelling of an individual person.

A model is only a restricted representation of an entity (in Long's terminology, its attributes are a subset of the entity's), and so the Phenomena 'M' can only be expected to be a subset of all the behavioural phenomena of the entity (ie person). The goal in constructing a model must be for the phenomena, with respect to the attributes being modelled, to be as identical as possible.

4. Why NOT construct a model?

There are many examples of models in Cognitive Ergonomics/HCI (see e.g. Whitefield, 1987), but it is worth asking whether there are any arguments against this practice. Such a consideration would at worst (i.e. the best argument) result in abandoning the idea of modelling, and at best (i.e. a poor argument) add weight to the case for modelling by virtue of forcing the argument to be explicit. The intention is that this section acts as a Devil's Advocate.

The only argument purely against models encountered, as distinct from arguments as to what would constitute good and bad models (which will be dealt with in the following section) is that of Kelvin (1980). Kelvin has a view of models which would conform to Warr's model-2 (i.e.analogy). He presents an argument intended to discourage the practice of such modelling (within limits) in psychology.
The first, and minor, strand he terms *parasitism*. This is an accurate use of the word since there is not intended to be any benefit for the science donating the analogy, but potentially suffers from being an emotive word. This strand reflects concern over borrowing from a well established science to gain respectability by association. This is indeed something which must be taken seriously; for example it is tempting to confer a greater level of importance upon a proposition expressed as a mathematical function over the same expressed in words. The second, and more important, strand concerns his view of the development of the sciences and ultimately psychology.

He asserts that the quest for systematic understanding reduces into a need to gain a systematic understanding of causation. In the early days, he suggests that man himself provided the model (i.e. analogy) for this understanding. The sciences (as we know them) only began to develop when they began to move on from such analogies and to consider the functional relationships observed between the phenomena in their respective domains. Having several independent sciences was necessary to provide the models on which psychology could initially be based. In this way he acknowledges the utility of models in the early stages of science or idea formation (he also acknowledges their pedagogical utility). His assumption that psychology needed the foundation of other sciences to borrow its initial models (analogies) is probably justified as an account of the late emergence of psychology relative to other sciences. The idea of using man as a model for man is somewhat circular. The point where his thinking departs from that which is being presented here is in his call to abandon modelling in psychology to achieve conceptual maturity. To try to force maturity on psychology this way is misguided; it is suggested here that it is better to think of the development of models-2 such that they eventually blur into models-1, as the route to independent theories.

It is assumed here that modelling is still a worthwhile pursuit in psychology (and thus in at least part of cognitive ergonomics). In any event, as Hargreaves (1980) observes, when science is going well, the philosophers of science tend to be ignored - maybe if modelling seems to be a useful enterprise, we should just get on with it.

5. What makes a good model?

A model could be considered ‘good’ in many different ways, depending on what is required of it. This will be discussed here in terms of two broad purposes - as a scientific device, and as an engineering device. Separately from this it is possible to ask whether it is, broadly speaking, correct (note then, that goodness is NOT equated with correctness). The issues of correctness will be discussed first, followed by the issues of goodness as a scientific device, and an engineering device.
5.1. Correctness

In the previous diagram, a model was represented in parallel with the entity it was representing - taking the same input (scenario) and resulting in a set of behavioural phenomena. The relationship of the behavioural phenomena output by the model to those observed to be 'output' by the person is the correctness of the model as a representation. In the best case, since it has already been said that the model is only a simplified representation of the entity, will be for Behavioural Phenomena 'M' to be a subset (rather than merely intersecting with) Behavioural Phenomena 'P'. In plain terms, this would mean that the model did not exhibit any behaviour over and above that of the person, but not vice versa.

5.2. As a Scientific device

In this section, the criteria which will be presented will be those that apply to scientific theories in general, and thus to the subclass, models-1. What makes a good model-2 will also be addressed.

A scientific theory has two purposes - it is both the embodiment of scientific knowledge (i.e. the product) and a vehicle by which knowledge is refined. A model will be a good one if it fulfils both purposes. If the goal of scientific enterprise is to establish truths about the world (Long 1986) then a model will obviously be considered a good one in the science paradigm if it is explanatory and predictive with respect to the behaviour of the entity it represents, i.e. if it is correct. However, whilst correctness is the ultimate goal, it is not necessary along the way. In the Popperian view of science, correctness is established by attempting to find fault with a model (or theory) - science advances, paradoxically, when a theory is found to be wrong. In this view of science then, there is another important quality of a model - that it be testable. This of course says something about the status of a scientific truth or fact - it only has that status until it is disproven. Once upon a time, the sun orbited the earth, which was flat and people would die if they travelled in excess of 25 mph: scientific fact.

The Popperian view of science is not the only one. Whilst the testing of scientific models by refutation is logically most desirable (see e.g. Marx 1976), everyday science often proceeds by finding supportive evidence (“survival by a thousand qualifications is what goes on”, Hargreaves, 1980, p308). This is a choice, and it should remain possible to refute a model, so it should be testable.

Besides being correct (or incorrect) and testable, a model should be simple - the principle known as Occam's Razor. Broadbent (1980) warns against the tendency to formulate over complicated models. The tendency is based, he suggests, in the
mistaken belief that complicated theories are more respectable, because theories in other, respected, sciences such as physics are complicated. The point is that theories in physics are complicated only because the simpler ones have been ruled out. Finally, a model in science is worthless unless it is communicable to the rest of the scientific community for criticism and further development (Wallis, 1980).

Models are different from scientific theories in general - they are also analogies. Within the current framework for thinking about models as analogies, it is possible to ask what would be a good analogical model? Such a model would have as many of the properties of the entity being represented, whilst at the same time having the minimum of additional misleading properties. This is most likely to be the case for a simple analogy over a complicated one.

5.3. As an Engineering device

The purpose of a model in the Engineering paradigm is to aid the design of an artefact. This would be a large responsibility for a single model alone. In practice models are only intended to be useful in more restricted design contexts. A model must therefore be the right tool for the job. The range of design problems which a model might be expected to be good for would be derived from a combination of its scope and its purpose. A model’s scope is what sort of behaviour it addresses, whilst its purpose is how it addresses this behaviour.

However, the question then remains as to what extent its recruitment to a design problem within this range is guaranteed to be useful. This might vary between the extremes of no guarantee, such that it only might be useful, and a full guarantee, such that it will definitely help. It is suggested that the latter would require a comprehensive framework for understanding different instances of design. Such a framework is not currently available and is beyond the scope of this thesis. Some alternative scheme is then required.

Given two models with equivalent scope and purpose, etc, it would seem easier to have more confidence in one which had already been successfully used than in one which had never before been recruited to design. A model might then have an additional status attribute indicating to what extent its usefulness in design had been demonstrated. Admittedly, there is no logical basis for believing that just because a model was useful before it will be useful again; this notion of a model’s status then has the same standing as many of the other proposals in this chapter - i.e. at a heuristic level. Such a demonstration can serve an additional purpose: Whitefield (1990) laments that whilst many people claim that their models would be useful in system design, this is seldom backed up with any detail saying how, let alone illustrated in practice.
An important consideration, which could be easily overlooked, is the cost of applying a model. Obviously a model will not be considered to be a good one if the cost of using it exceeds the benefits that result. Finally, a model should also be both communicable to, and usable by, its intended users.

It will be reasonably clear by now that there are many contributing factors to whether a model is a good one in an application sense, summarised best by 'what you want to do with it', and that there are interactions between these factors such that it will be difficult to say definitively that any particular model is a good one. (e.g. usability may interact with cost - see below). What is necessary in the minimum would appear to be a set of parameters accompanying a model which specify its intended scope, purpose and procedure. After this, some actual verification - i.e. documented use of the model is probably necessary.

Finally, whether the model in question is in fact better described as a model-1 or a model-2 (i.e. an analogy) is not relevant in this paradigm.

5.4. The above expressed as criteria

To summarize the above, a model would be considered good in the science paradigm if it were:

• correct (given what is known at the time)

• testable

• simple

• communicable

A model would be considered good in the engineering paradigm if it were:

• communicable

• usable by an intended user

• able to improve the design of an artefact (of specified class)

• not more costly than some other route (to the same state)

Both of these assume a degree of correctness of representation.
5.5. Can one model be good in all respects?

That the same representational structure can be a good model in all senses is possible since there are no obvious conflicts in the above criteria. It is, naturally, also possible for a model to qualify as neither.

The fact that engineering depends on science, but not the other way suggests that a model which is good in the scientific sense will at least have the potential to be good in an applied sense, but may not necessarily be.

For symmetrical completeness, it is necessary to consider whether it is possible for a model to be a good model in the engineering sense whilst not being so in a scientific sense. The position maintained here is that it is - an example would be Newtonian physics which is no longer thought to be strictly ‘true’ in the scientific sense (Hawking, 1988), but is used in nearly all engineering calculations because of its cost and usefulness compared to the theory of Einstein which is more strictly correct. There are situations where the Einstein model is much better - space flight, for example, where the accumulation of error due to a Newtonian model would be quite serious.

5.6. Summary

To summarize so far, a classificatory scheme has been presented which allows models to be divided into Models-1 (so called direct representations, which would include many scientific theories), and Models-2 (termed indirect, and covering those occasions in science where an analogy is used as a tool). The point was made that many Models-1 (i.e. theories) start out as Models-2 (i.e. analogies) and in the course of their development take on a meaning of their own. Arguments against the use of models were discussed, but it was nevertheless decided that the enterprise was worthwhile. The role of models as both scientific devices and engineering devices has been reviewed and expressed as criteria. It is suggested that the relationship between the criteria for the a model in the two paradigms does not preclude the same model functioning as a device in both.

6. Generality

This section addresses the question of what generality means, specifically with respect to the derived working definition of a model. The aim will be to arrive at a conceptualization which, with the additional constraints in the following section, can be used to specify a (heuristic) method to optimize the development of a model with the desired properties.
6.1. What does 'generality' mean?

Generality means different things to different people, and, as will shortly become clear, different things to the same person.

The derived definition of a model presented earlier relied in part on Long's (1987a) characterisation of models, expressed in its short form below. Marked on this expression are the three points where it is thought the term generality could apply:

\[ M \rightarrow R(E) \text{ for purpose } P \text{ and utiliser } U \]

This section will address each of the above in turn.

6.2. Generality of Scenario (E) represented by the model.

Long's (1987a) characterisation of a model as a representation of an entity assumes what is here called a scenario, i.e. it is a representation of the interaction between an object and its scenario (situation), specifically in this case, a person doing a job.

Generality across people is assumed by much of psychology - i.e. it is assumed that there is a degree of consistency across people, although the experimental practice of looking at behaviour across a group of subjects is a concession to supporting this assumption. There are branches of psychology concerned with the opposite, i.e. individual differences, but these will not be reviewed here.

The alternative is generality of scenario - and it is proposed that it is this which can be optimised, even if not guaranteed. This is possibly one of the more important categories of generalisation; after all, we are always trying to extrapolate our knowledge to new instances, thus Broadbent (1980) comments that the point of a model is to avoid the need for experimentation. (He is arguably assuming that generality of scenario is implicit within models, which, as will become clear, is not disputed here. The aim here will be to establish some basis for this such that it is not surprising when an extrapolation fails [or will it be more surprising?]).

6.3. Generality of Purpose of the model.

A previous section has described the practices of science and engineering and thus the purpose of models within each. It was noted that there need not be a clash in scientific purpose and engineering purpose for a model (although there could be). However, it is suggested that a given representation may not necessarily support more than a restricted set of purposes within a given paradigm. This may be called the representational problem, or 'tools for the job'. To take an analogy used by Marr (1982), there are many systems for representing numbers - Roman, Arabic and so
The point is that these representations are not equal with respect to the purposes they support. The Roman system of numbers does not support long division very well at all. The Romans never really developed much in the way of mathematics, however, they were good civil engineers, so it could be hypothesised that their number system did have some purpose. This problem of generality of purpose applies to scientific models as well as applied models. The level of description of a model can have important consequences for its purpose. It may well be possible (one day) to describe the functioning of a human being at the level of the interaction of molecules, but this representation would not realistically support discussion of, for example, perception of language. Possibly it could be done, but it would require a disproportionate amount of effort.

To summarize, it is suggested that the right tool for the job will be more efficient than a single general all-purpose tool.

6.4. Generality of Utiliser of the model.

Different users of models could be expected to require models to be expressed in different forms, even for the same purpose. Given a model of behaviour, for example, some users, with previous experience of Cognitive Science, would be able to use a model in a form which relied on such experience by utilising known concepts. Other users, without experience, would need the model to be expressed more completely. One could construct a model that was intended to be used by any utiliser, but as in the case of generality of purpose, it is a case of being efficient. To create a model usable by all, one would have to choose the lowest common denominator and concede nothing for the benefit of experts. This would be both excessively demanding on the constructor of the model, and it is felt, constrain some users unnecessarily.

6.5. The type of generality to be addressed.

Whilst it is possible to provide generality in all three ways, this PhD will only attempt to address Generality of Scenario. This is because a) there is a mechanism which allows, it to a degree (the following heuristic), and b) it is implicit in the subject that is to be modelled. The following sections of this chapter attempt to provide a mechanism for this. It has been stated above that a model’s scope and purpose should be explicit. To this list it is necessary to add the intended user. For the present model, this will be a behavioural scientist.

Unless otherwise stated, all following references to a ‘general model’ can be taken to refer to generality of instance.
6.6. Conceptual Variables and Generalization.

This section is intended to show that generality of scenario is already implicit in the way we talk about entities. The task of this chapter then becomes to make this explicit as well as to optimise it in some way.

Chapanis (1988) discussing generalization, makes the interesting and important point that all things bearing the same label are not necessarily equivalent. He takes the example of the study of fatigue - not all studies address the same phenomenon simply because their titles contain the word. The problem lies in that fatigue is what is known as a Conceptual Variable. Conceptual Variables cannot be studied directly - they have to be operationalized and operationalizations of the same Conceptual Variable may differ. It is this differing operationalization which he argues acts against generalization. The problem is easily illustrated by the timeworn scientists’ qualification: “ah, well, it depends what you mean by ...”.

Behaviour expressed as a conceptual variable is common to many different instances, so it is tautological to talk of a general model of a conceptual variable. To take an example, Multitasking is a conceptual variable, but one couldn’t develop a ‘General Model of Multitasking’ (where general refers to behavioural instance). Rather it would automatically be a general model of multitasking. Here lies the point that possibly Chapanis could have pushed further - he was right that because of the necessary operationalization, generalization does not follow to all instances, but it can be taken a step further to say that by constraining the operationalization (i.e. stating what one means by ...), then generalization across instances is possible (to a degree).

There can be no predictable, guaranteed, relation between the characteristics of one study and those of future studies (or applications) to which one might want to generalize. Whatever is done can thus only be an optimization, a heuristic, based on some criteria (such that its failure will be at least understandable). The heuristic method, presented below, for developing a model which is general across instances will be based on the idea of constrained, explicit, operationalization.

6.7. A basis for generalisation

The basis for generalisation of instance will be Constrained Operationalization, in other words, saying what is meant by a conceptual variable and only expecting generalisation if the definition fits.

A scenario will be equated with a job, and expressed as a set of attributes. A conceptual variable is thought of as relating a scenario to a particular set of attributes. Thus the same job would have many attributes, but different subsets of these would relate to different conceptual variables. As an example, consider the job of a sonar.
operator on a submarine. Some of the attributes of such a job would be time of day, target frequency, equipment characteristics (e.g. screen brightness) and level of general illumination. If one were interested in the conceptual variable vigilance, the attributes of interest would probably be time of day and target frequency. If, however, one were instead interested in visual reaction times, the one would be more concerned with the equipment characteristics and prevailing levels of illumination. Note that these attributes are not the same as those of Long discussed above which refer to the model/person and the input together, rather than just the input alone. The reason for this is that the present concern is with behaviour, which is exhibited when a person does something, and so it is only reasonable to expect that behaviour will vary (in some way) with the properties (i.e. attributes) of what people are given to do.

Any one scenario will contain a subset of attributes relevant to a given conceptual variable - so vigilance in terms of submarine sonar operations may not be precisely the same as vigilance in the case of a nightwatchman when expressed in terms of job attributes. The assumption is that a number of such attributes will nevertheless be common to both. Generalising a model across scenarios thus requires that the new scenario be expressible in terms of attributes already known. Obviously then, the larger the set of Job attributes addressed by the model the better - this will be the basis of the heuristic method for developing a model with this type of generality. In short, such generality could be summarised as generality by commonality.

6.8. Generality breaking down

Generality could be said to have broken down if the model no longer produces a behavioural output which agrees with the behaviour of the person (see the original 'Black Box' diagram of Fig 2.2). This could occur for two reasons:

a) the input being outside its limits (this ought not to be surprising)

b) the model itself being wrong, but the input supposedly within its limits. This might be the case if the model is internally incorrect - analogies break down sooner or later. The job specification is in terms of attributes, but the model is a system which has to represent their interactions. It is entirely possible that this particular constellation of attributes (and thus set of interactions) has not been met before, and so might not be represented correctly in the model. This is fine from the point of view of the model as a scientific device since it is an accepted part of the development process, but in the context of application, there is a need to optimise a model's useful life against this.

Minimising both of the above factors requires a development program which draws on as many different, partially overlapping instances as possible, plus a requirement
that a model come with an explicit definition of its conceptual variable which it
purports to accept as its input, in terms of attributes.

Note that the precise definition of a conceptual variable can be expected to develop
alongside the model which refers to it.


Fig. 2.3. Schematic representation of the relationship of a model to different
scenarios.

Figure 2.3 is intended to represent the scheme for generalisation of models based on
the commonality of Job Attributes as discussed. This diagram will be returned to
later, when the approach to developing such a model is presented (Section 8).

7. Generality in the context of Cognitive
Ergonomics

The proceeding discussion of generality has been mainly concerned, at least
implicitly, with generalisation in the context of application. However, an earlier
section described how this only formed part of the larger context (and the context of
the present research) of Cognitive Ergonomics. If the purpose of engineering is the
design of artefacts, and engineering knowledge is what supports this, then generality
of instance is the goal of being able to design artefacts to solve the problem in more
than the original domain. In science, it has been argued that the purpose of a model
(pedagogy aside) is to encapsulate some truth about the world in a way which allows
that truth to be build upon. Generality then would be both the ultimate goal (i.e. a
theory which accounts for everything), and an assumption underlying the model.
Testing this assumption would be one route to achieving the goal. A general model,
as defined, in the context of Cognitive Ergonomics should then be expected to be
general in both of the above ways.
8. A (heuristic) method for developing a model with the properties outlined above

This section presents a method whereby a model might be developed with the property of generality as discussed. The method draws on the model for generalising models as presented in figure 2.3 (Section 6.8). The basic notion is that models can, and need to, be developed in two directions, referred to as horizontally and vertically. In this diagram, jobs (scenarios) are represented as sets of attributes, within a universe of all possible attributes. Similarly, models are mapped onto this universe as the sets of job attributes for which they represent behaviour. Different scenarios will share some attributes, and so will intersect. The horizontal development of a model refers to its modification to apply to a novel domain whose set of attributes only partially intersects with those of the model. The set of attributes addressed by the model is thus extended to become (theoretically) the union of both sets. In this way the scope of the model is extended - its set of attributes is now likely to intersect (more fully) with a greater number of domains.

Vertical development of the model refers to its refinement, within its scope. The term vertical arises from considering the series of sets of attributes (i.e. scenarios) addressed by the model as layered on top of each other, at 90° to the plane of the universe in the diagram. Vertical development, concerned with enhancing the robustness and detail of a model, is thus based on those aspects of a particular job which are repeated between different occasions. For clarity, vertical development further subdivides into three aspects, depending on the way the data affect the model, referred to as Debugging, Consolidating and Unpacking. The following sections briefly outline how each of the four aspects of model development are supported by data.

8.1. Horizontal

Behavioural phenomena new to the model, probably associated with any novel job attributes of the particular instance would be relevant here. It is of course possible that the horizontal extension of the scope of the model does not require any changes. Since it is important to practice vertical development within the intended scope of the model, and this scope may well change as a result of horizontal development, it seems only pragmatic to advocate doing the horizontal part of any development cycle before the corresponding vertical development.
8.2. Vertical - Debugging.

Debugging addresses the question of whether there are any behavioural data contrary to that predicted by the model. Behavioural data not falling into this category must be at least consolidatory (see below). This is really a question of the internal consistency of the model, the importance of which is recognised by Chapanis (1988). It is possible (even probable) that during development of the model, a job (i.e. a set of attributes) will be found for which the model does not produce the same phenomena as the person being modelled. In such a case, the model is obviously not fulfilling its objective. It is important to note that this must be with respect to its intended scope and thus a new task for which the model fails must fall within this scope before it is proper to make any modifications to the model.

8.3. Vertical - Consolidating.

The credibility of the model increases with each successful application to a new task. This is the sort of science which is logically contrary to that advocated by strict Popperianists. It could be said that consolidation is really what is left after debugging, and might therefore be considered redundant. However, it is thought that such positive support for the model (rather than evidence against it) should be stated explicitly.

8.4. Vertical - Unpacking.

It is likely to be the case that each cycle in the development of the model will emphasise different aspects of the model. The resulting data can be used to specify a structure or a mechanism more precisely.

Whereas the two other categories of evidence which come under vertical development (debugging and consolidating) are alternatives to each other, unpacking could really be thought of as a sub-category of consolidating. This is because data which are relevant to consolidating may then be of further use in unpacking (and similarly, it is not possible for data to be useful for unpacking without fulfilling a consolidatory role at the same time). Unpacking assumes consolidation. Different instances of behaviour are likely to yield data with different content biases. Unpacking takes advantage of this to use data which are consolidatory at one level, but which have available more detail, to enhance the level of detail of some of the structures and mechanisms in the model. However, rather than present it as such a sub-sub-heading, it has been decided to class it as an equal activity under the heading of vertical development.
The above model development method needs to be harnessed to a data gathering strategy which supports it - this has already been discussed (above) and entails the observation of behaviour in partially differing scenarios.

9. The nature of the data collection

This section considers the nature of the different ways of collecting data concerning human behaviour. It will be taken that there are broadly two alternatives, although it is acknowledged that these might more properly be thought of as lying at opposite ends of a continuum. The aim is to assess these two extremes and to justify the choice of one of them. The section is organised with a discussion of their merits *per se* to begin with, followed by considering each with respect to the proposed developmental method. A choice is made which is used in subsequent chapters (4, 5 and 6).

The first extreme would be characterised as reductionistic experimentation, and is typified in much of the research of Experimental Psychology. Briefly, it centres around the idea of eliminating or controlling as many of the variables in a situation as possible in order to gain an understanding of just one or two of them. Such research methods are closely associated with the Popperian view of the advancement of knowledge by refuting hypotheses.

The other extreme would be characterised as *ecological* research. This paradigm seeks to emphasise the richness of the world, but in so doing forsakes a great deal of the experimenter’s control over a situation, and thus the data gathered. Rather than hypothesise and test, this type of research would be more likely to rely on less structured observation.

There are proponents of this alternative. Brehmer (1984) discusses what has been termed ‘Brunswikian’ psychology, after its originator. This view distinguishes between an encapsulated psychology, concerned with the organism - i.e. the psychological machinery without reference to the environment, and a functionalistic psychology. Encapsulated psychology is considered to have missed the point, in that psychological mechanisms have a specific use and must be studied in use. The need to study the organism in a context gives rise to functionalism:

“Consequently, the primary datum for psychological analysis should be achievement of goals under natural circumstances and further analysis should be directed at finding the strategies and tactics employed by the organism to reach its level of achievement” (Brehmer, 1984, p384).
Much the same distinction is apparent in e.g. Chapanis (1988). He distinguishes between basic and applied research, where the former is equatable with traditional psychology laboratory research. The latter is not strictly the same as functionalism in that he uses the word ‘applied’ simply to indicate the degree to which the experimental conditions match some current problem. Nevertheless, there is a feeling that this has some of the flavour of more ecologically valid research.

The choice is then between a functionalistic, ecological, approach, which forsakes control but promises to deliver a richer and fundamentally more relevant set of observations, and the traditional experimental paradigm. The latter’s hypothesise and test mechanism will suffer from the limits of the experimenter’s imagination in the initial stages when little is known. The same is not true of less structured observation.

9.1 Real-World observation

Several authors (e.g. Baddeley & Wilkins, 1984) have made remarks to the effect that one of the important benefits of practical studies is that they throw up new phenomena. In terms of the developmental method outlined in this chapter, this means that such studies are likely to contribute well to horizontal development, and thus to the final scope of the model.

Considering the contribution of such a study to the vertical development of a model, such data can be used in all three sub-categories. It can support debugging, but only in a passive sense. It can certainly support consolidation, and it is probably quite good for unpacking for the same reasons that it is good at horizontal development (because it turns up new things).

9.2 Manipulated, experimental research

This type of study might be expected to be less able to feature in the horizontal development of the model, for reasons already stated. Similarly, its contribution to unpacking might also be restricted. This type of data gathering is, however, likely to be much stronger when it comes to debugging a model. Rather than being limited to passive debugging, it would be possible to contrive a situation to answer a particular question, and thus actively debug the model. There is no reason why experimental research could not contribute to the consolidation of a model, although it should be remembered that such development is typically in the form of supportive evidence and thus contrary to the usual approach.
In conclusion, it is the intention for this project to rely on observational data gathered in an uncontrolled, ecological, manner. This is partly a matter of choice, but also partly justified by the slightly more favourable balance of contributions such data is able to make to the development of a model according to the proposed method.

10. Summary

This chapter has reviewed what it means to model behaviour in the context of Cognitive Ergonomics. From this consideration, a heuristic development method has been proposed. This method forms the basis for the research to be reported in the remainder of the thesis, and is reviewed in the final chapter.
Chapter 3

The Initial Model

Prologue

This chapter provides an account of the generation of an initial model of multitasking behaviour. Rather than basing this solely on observation, the behavioural phenomena are dictated by a review of relevant literature. The selection of this literature is in turn directed by a consideration of the preliminary definition of multitasking provided in Chapter 1.

Contents

1. Introduction
2. An Initial Set of Job Attributes
3. The Literature
   3.1. Network Planning
   3.2. Planning Behaviour
   3.3. Attention & Performance
   3.4. Multitasking
   3.5. Interruption
4. Summary of Behavioural Phenomena
5. The Model
   5.1. Introduction
   5.2. The Two Foundations of the Model
      5.2.1. Planning part
      5.2.2. Execution part
         5.2.2.1. Microsoft Windows
   5.3. Structure
   5.4. Mechanism
   5.5. Problems with the model
   5.6. Summary
1. Introduction

This chapter is the first step in the realisation of the developmental strategy outlined in the previous chapter. It will take the initial definition of multitasking from Chapter 1 and identify relevant and potentially useful areas of knowledge. The literature in these areas will then be reviewed, with the dual objectives of identifying an initial set of behavioural phenomena which will need to be addressed by the initial model, and also identifying concepts which can be recruited to this initial model.

Rather than attempt an exhaustive review of the literature, a prototyping approach will be adopted, in which enough literature to provide a reasonable basis for forming the initial model will be considered. It is acknowledged that normally there might be a risk of repeating what others have done associated with such a limited review. However, in the present case there is in fact very little directly relevant available literature. The main aim of this thesis is to develop a model based largely on observation of behaviour in the world - the initial model only enables this in the context of the proposed developmental strategy. It is envisaged that the definition of multitasking will develop alongside the model and thus new areas of relevant literature will emerge. To this end, there will be a further, post hoc, literature review in the final chapter.

Following the review of the literature, the target behavioural phenomena will be summarized, and an initial model presented. This will be based on conceiving the problem in two parts - one concerned with planning and the other with execution - and thus the model will be based on this dual foundation.

2. An Initial Set of Job Attributes

The method described in Chapter 2 requires that instances of multitasking be expressed as sets of attributes, termed Job Attributes, or JAs. The purpose of this section is to take the simple definition of multitasking provided by Chapter 1 and derive an initial set of such JAs.

Obviously, the prime attribute for a job to be considered as multitasking is that it should feature the interleaving of tasks, for whatever reasons. Let this then be the primary JA:

• There is interleaving of tasks

This alone might be too broad, so a second JA will be used as a qualification:

• The person has some degree of familiarity with the individual tasks.
This JA is associated with the (informal) observation that the majority of tasks performed in the workplace are ones with which the person is familiar, and the widely accepted belief in psychology that there are likely to be qualitative differences in the performance of the same task, depending on the degree of familiarity. The alternative would be to ignore any such distinction, which might result in familiarity effects obscuring the phenomena of interest, or to concentrate just on the combination of unfamiliar tasks. As stated, it is taken to be the case that this would not be representative of jobs in the workplace, and therefore contrary to the primary concern here.

The third, and final, member of this set of initial JAs is concerned with realising that the world is a constantly changing and sometimes unpredictable place and thus specifying that any coordination of multiple tasks will need to be flexible and adaptable, possibly at short notice. A typical instance of the above would be the interrupting telephone call which requires immediate and unplanned-for attention. Thus the final JA can be stated as:

- The multitasking situation is not entirely predictable and may change at short notice.

3. The Literature

This is not a critical literature review; it is directed at identifying interesting and potentially relevant behavioural phenomena (BPs), and so might seem unusually selective. The literature to be presented is drawn from five areas. With the exception of Network Planning, which derives from the discipline of Operations Research, the areas or categories are derived from Cognitive Science in general (see Chapter 2 for a discussion of how this forms the major knowledge base for Cognitive Ergonomics) and various sub-disciplines such as Psychology and Artificial Intelligence (AI) in particular. Since there is no obvious mapping to the set of JAs, the simplest and least confusing way of presenting this diverse collection of observations is by the five categories in turn. Immediately below is a brief justification for the inclusion of each of the five areas. Following this, each is addressed in full. The section ends with a summary of BPs which will then be addressed by the derivation of an initial model.

Network Planning. Network planning is one aspect of a branch of science known as Operations Research. Operations Research originated in the Second World War and was concerned with applying scientific knowledge, not to the development of new technologies, but rather to the improvement of the utilization of existing ones. It has grown to be a large discipline, encompassing many different areas, one of which is concerned with project management, and might be termed 'Network
Planning'. This aspect of Operations Research is included for its concepts and terminology, which will be recruited to the initial model of multitasking.

**Multitasking.** Although it would be possible to ignore any existing research into multitasking, and develop a model completely independently, this is not, in principle, the approach which will be used here. However, in practice, there is only very little such research and its contribution will be correspondingly small.

**Planning Behaviour.** Given the criterion that there should be some degree of familiarity with the individual tasks which are interleaved, it is further assumed that any interleaving will be coordinated in some way, rather than being purely reactive. Such coordination has been termed planning, and has been the subject of much research. Some of the observations emerging from this research are presented.

**Attention & Performance.** Multitasking is taken to be the realisation of parallel tasks as a single serial stream of behaviour. Research in psychology which has addressed the limits of the behavioural capacity of the person has typically been put under the heading of ‘Attention and Performance’. This area of the literature is reviewed as a basis for understanding the nature of this seriality.

**Interuption.** One of the initial JAs specifies that there are likely to be changes to the set of tasks being interleaved at short notice. Although this will have an effect on the planning and coordination of the interleaving, it may also, in some circumstances, be considered as an interruption. From everyday experience, interruptions may be expected to have a negative effect on performance. This section is included to attempt to qualify this.

The above categories are very similar to those suggested by Miyata and Norman (1986) in their discussion of multitasking. However, the intention here is not to review these categories in the same way, or to the same depth since they were pursuing a different goal. The context of Miyata and Norman’s review is that of a guide for (computer) system designers, and given the constitution of such a population, amounts to little more than a constrained introduction to psychology. The present review, in contrast, will take much of psychology for granted (e.g. the psychology of memory), and will not get into research on ‘action’, for example.

### 3.1. Network Planning

Network Planning is a convenient, although not necessarily universally accepted, generic term for a collection of particular approaches (Smith, 1971). Historically, its basis is in two techniques, PERT and CPM, originating in the late 1950s. PERT stands for Programme Evaluation and Review Technique, and was conceived to manage the development of the Polaris missile. CPM (Critical Path Method)
originally addressed chemical plant overhaul. The common objective for the two techniques was to optimise the length of time a project would take. Since then, this technique has been shown to be widely applicable, and able to be used to optimise a project with respect to any one objective (e.g. time, cost, etc). Optimising a project with respect to all of its objectives has proven less achievable in practice.

A Network Planning exercise starts with the construction of the network. This is a set of nodes, representing events, linked together by arrows, representing activities. It is then necessary to calculate the earliest and latest times for each event. From these figures, the event slack can be derived as the difference between the earliest and latest times of the event. A critical event is one for which there is no slack in the time at which it can occur. A critical activity would be one which fell between two critical events, and which was equal in duration to the difference in time between these two events. The chain of critical activities through the network, from the start to the finish, is the sequence of longest duration and as such governs the duration of the project. This sequence is termed the critical path. Only if the time taken by activities on this path is shortened can the total project time be shortened. Shortening the duration of activities elsewhere on the network has no such effect.

Such a network, however, is only (at best) half the story, since it only considers the duration and sequential dependencies of activities. A subsequent stage is required to produce a schedule from such a plan. At this point, the resource demands of the activities must be considered with respect to availability. Knowing the time windows (i.e. the difference between the earliest start and the latest finish) of a pair of activities competing for a particular resource allows one to determine whether they can proceed in parallel. If they cannot, then the resource could be said to be overcommitted and require either levelling (taking advantage of under-use elsewhere), or more of that resource. Resource levelling is usually performed on an ad hoc basis.

Most of the jobs which could be described as multitasking under the current definition have many of the attributes of the projects typically analysed by network methods, albeit on a much smaller scale. They consist of tasks, with sequential dependencies and resource requirements. Perhaps a major difference is in the fact that instances of multitasking may not have clearly defined start and end points, but rather are more arbitrary episodes in an ongoing stream. A definite start point and end point is a requirement in network analysis.

There is no intention to perform formal network analyses in the course of this thesis, primarily for this reason, but also because it is not thought that the effort involved would show any appreciable return for the purposes of understanding multitasking behaviour. Some of the concepts of network analysis will nevertheless be recruited informally to the model of multitasking. Chief among these will be the concept of
critical activities and the critical path. Without constructing a network it will not be possible to specify fully the critical path. However, it is often possible to judge informally which activities control the overall job time.

3.2. Planning Behaviour

A great deal of planning research has been conducted in the context of AI and has been concerned with the development of computerised planners. Some of this research has been concerned with developing optimal planning systems, whilst other research has been aimed at mimicking human decision making and thus advancing our understanding of it. In the present context, the interest is clearly biased towards this what and why of planning behaviour, and away from the how (i.e. discussion of backward and forward chaining etc).

Hayes-Roth and Hayes-Roth (1979) is a useful example of planning research in AI. They present observational data gathered from a planning exercise plus a cognitive model of planning, using the mechanisms of the Hearsay-II system. The specific details of the study and the model will not be discussed here. There are however, several more general points which emerge from their paper. Rather than propose a hierarchical, successive refinement model of planning, they choose to model planning as an opportunistic process. A plan could thus be conceived of as existing in a 2-dimensional space. One dimension would be time, the other some measure of completeness or degree of detail. Importantly, they observe that planning does not necessarily proceed strictly forward in time, but rather a partially incomplete plan may consist of some complete (i.e. specified in detail) elements interspersed (on the time dimension) with other elements specified in less detail.

Although Hayes-Roth and Hayes-Roth acknowledge that planning alone is only half the story, the other half being the implementation or execution of the plan, they do so in a way common to much AI planning research. This typically sees planning and execution as two separate and successive events, i.e. a plan is something which is created in its entirety beforehand, and then executed. It is worth stating, in this respect, that their own model produces a plan which cannot be realistically executed. Fortunately, there are more recent movements against this divorce of planning and execution and possibly one might speculate that such an unrealistic plan would be avoided in this way (because it would be constantly revised during execution).

The theme of the construction of a plan bound up with its execution is taken up by Suchman (1987). Central to her argument is the notion of situated action. For her, all actions are situated - that is to say they exist only in the context of a particular set of concrete circumstances and because these circumstances can never be fully anticipated the actions are therefore essentially ad hoc. Such a belief reduces the role
of plans in behaviour from complex constructions to simpler orientation devices and similarly reduces the effort an organism might be expected to expend in devising them.

The two above points - that it might be more accurate to think of plans as simple, partial devices, the construction of which is intimately bound up with their own execution - are echoed by Young and Simon (1987). They introduce the concepts of horizontal and vertical incompleteness in plans. A vertically incomplete plan is described as one in which the level of description is still higher than the primitive actions which the executor can carry out and thus requires further refinement. Horizontal incompleteness in a plan refers to the possibility that not all steps of a multistep plan may be specified. In general, they state, "a partially incomplete plan can exhibit both kinds of incompleteness, freely intermingled".

Whilst one of the reasons for incompleteness in plans is thought to be the unpredictable nature of the world, they suggest that another reason for favouring partial plans is the limited mental workspace capacity of the human (see discussion of Attention and Performance in Section 3.3 below).

Johannsen and Rouse (1983) proposed a new concept in planning in terms of a dichotomy between what they term event driven and time driven planning. Their proposals arise from experimental work using aircraft pilots in a simulator, in which the subjects were exposed to various missions of controlled character. The main measure taken was a verbal self rating of depth of planning on a scale of 1 to 5, every 30 seconds. Subsequent analysis of this data with respect to the nature of the mission at the time was interpreted in terms of the two types of planning. Using the frames model of Minsky (1975), they describe time driven planning as monitoring the execution of a script - so this type of planning might be expected to account for increased depth of planning as an anticipated event was approaching in time. Event driven planning, on the other hand, is described as updating a script, or creating a new one. This type of planning would be expected to occur in response to an unanticipated event.

Finally, in respect of plans, it is worth considering the contribution of Byrne (1981). The assertion of this paper is that plans exist in memory (possibly therefore one might use the term script or frame), but constitute more than just a set of instructions for some complex act, but also serve as a mental representation whereby associated information may be accessed. This conclusion is based on the observation that when, for example, housewives are asked to list the ingredients for lemon meringue pie, they do so in the order that they are used - the inference being that subjects are mentally executing the relevant plan in order to access the desired information. Thus information associated with the performance of a plan may be bound up with the
mental representation of that plan, and may be accessed by the mentally simulated execution of that plan.

To summarize the important points of this admittedly limited sample of the literature on planning behaviour, there are suggestions that a plan is not necessarily a very low level specification (in terms of actions) even in its most complete form, and is likely to be constructed in both time driven and event driven (opportunistic) manners such that it will probably be both horizontally (time dimension) incomplete and vertically (detail) incomplete. Further, it is acknowledged that plans can be remembered from one occasion to another, and therefore need not be constructed afresh but rather modified as necessary. Finally, plans may be mentally executed, and this may serve as a means of recalling information associated with that plan.

3.3. Attention & Performance

The goal of this section is to provide a working approximation of the serial nature of the human which underlies the need for a set of parallel tasks to be realised as an interleaved sequential stream. It is acknowledged, however, that such seriality is not necessarily absolute, and it may be more precise to think in terms of the constraints on performing activities in parallel. The attention and performance literature is vast and so some means of achieving this desired approximation, without an exhaustive review, is necessary. This is provided by Reason (1986), in the form of a ‘consumer guide’ to what he terms Framework Models of Performance.

He contrasts such framework models with local models in psychology in terms of many characteristics. For example, he suggests that whereas local models tend to be predictive and refutable by experiment, framework models tend to be descriptive and more subject to paradigm-shift, and if the primary function of local models is to disambiguate theoretical questions, then that of framework models is more directed towards emphasising the general points of agreement between models and supporting research aimed at limiting the generalisations. The characteristics of such framework models, particularly the broad agreement on certain points, suggest that they should be considered further in the current context.

Reason reviews six of these framework models, which will not be repeated here. Out of this, he concludes that there are several areas of commonality, particularly relevant here are the concepts of a) restricted workspace, and b) controlled and automatic modes of processing.

The concept of a restricted workspace refers to the observation that people are only able to be mentally concerned with one task at a time, whether it be rehearsal for remembering, or visual search of novel stimuli. For example, a great deal of research into the properties of this system has been conducted using an interference paradigm,
in which the subject is encouraged to try to attend to a second task simultaneously with the first. The finding is that performance on both tasks suffers in such circumstances; the inference being that one task is interfering with the other’s use of a restricted workspace. An everyday example might be trying to listen to two separate conversations at once. This property is typically modelled in terms of a structure, under the name of ‘Working Memory’, however, to avoid association with particular uses of this term, this thesis will recruit the concept under the name of ‘Mental Workspace’.

The other relevant concept, that of controlled versus automatic modes of behaviour, refers to the need for the above Mental Workspace to be involved in controlling behaviour. Much of our behaviour is carried out in a so called automatic mode, without the influence of the Mental Workspace, in which we are not consciously aware of each action performed and each piece of information heeded. Such behaviour is typical of tasks which have been practised and become routine. In contrast to this is a controlled mode, typified by novel or difficult tasks. In this mode, we are consciously aware of our movements and our decisions to take particular courses of action. Controlled mode behaviour is often slower than automatic behaviour, but has the advantage of being very much more flexible - automatic, learned routines are thought to be fairly specific, whereas behaviour in controlled mode is thought to be guided by the circumstances in real time. This coordination of guidance is provided for by the restricted capacity Mental Workspace.

From the earlier discussion of the limited capacity of this structure it follows that a person could only realistically be engaged in one task at a time which required such control. The supposed benefit of evolving automatic processing for routine behaviours is that it frees up the precious Mental Workspace for other things, generally what might be termed parallel mental activity. The limitation of being concerned with just one task at a time then applies to the use of the Mental Workspace in a controlled mode, rather than strictly to the person as a whole. This issue will be taken up later when the initial model is described.

3.4. Multitasking

There is very little reported research which specifically addresses the area of multitasking itself, even in its broadest definition. A particular exception to this is Beishon (1969). This paper reports an observational study of an ovenman - the person responsible for managing the ovens in a bakery. The job of this particular ovenman can be considered to involve multitasking since he is required to interleave the baking of a number of different batches of cakes in three different continuous ovens. It could in fact be described as possessing all three of the JAs specified at the beginning of this chapter.
The reasons for presenting this study in relative detail are twofold. Firstly, as an observational study, it reveals several potentially interesting behavioural phenomena, and secondly, Beishon's goal was to produce a model of the ovenman's behaviour. This model will be described below, but will not form the basis for the present model, but will at least provide a contrast.

The two most basic, obvious behavioural phenomena which can be gathered from this study are that people can and do interleave their activities, and also that they do anticipate and plan ahead. Additionally, Beishon provides a classification of interruptions into two groups according to the immediate origin of the trigger. The majority of interrupts have external triggers; that is an event in the world directly or indirectly causes the ovenman to suspend his current task and take up another. Examples of external triggers are given as noticing a new batch of cakes at the mouth of the oven or being told of something by a fellow worker. The former would be an example of an indirect external trigger whilst the latter would be a direct external trigger (although depending on what the workmate told him). In contrast to these externally triggered interruptions, there are those for which there is no such discernible trigger. These abrupt changes in ongoing task for no apparent reason are considered to result from some internal trigger. The principal example of such would be the occasional monitoring and inspection of a batch of cakes in an oven by the ovenman.

The structure of Beishon's model is in terms of a set of procedures, and in this way it utilises a computer program metaphor. The majority of the procedures are the low level routines: Beishon proposes eight of these, one for each of the identified stages in the baking process, so for example there are separate routines for recognition and oven allocation. These routines are linked together by the Main Procedure which seems to specify the relationships between the different routines. Finally, since many of the steps in the main procedure are conditional upon events in the outside world (i.e. the bakery), the main procedure is dependent on the Executive Routine.

The Executive Routine is responsible for controlling the sequence of events in the ovenman's behaviour. Whilst this is probably the most interesting element in the current context, it is also an area which Beishon finds difficult to model. The Executive Routine is described as having four responsibilities:

1. Entering and receiving call-ups for future entry into a specific routine.

2. Handling externally initiated interrupts.

3. Searching through future expectancies for anticipatory activities.
4. Keeping and updating a list of activities to be done next.

The last of these (maintaining a *to-do* list) is its most important function.

In addition to the above procedures Beishon proposes firstly a Lookup Table in memory, which would contain much of what the ovenman needs to know to do his job, since he seems to have this information readily available. Secondly, he acknowledges the need for some sort of perceptual mechanism, but leaves this unspecified.

The full mechanism imposed on the above structure will not be discussed here, only that which is associated with the executive routine - the structure of most interest since it controls some of the sequence of behaviour. The remaining observed changes in activity correspond to the basic routines discussed above and are thus fairly easy to account for, since such changes are implicit in the main procedure. Of the four listed responsibilities of the executive routine, two would seem to be more important, namely the handling of externally initiated interrupts and keeping and updating a list of activities to be done next. The other two activities could be considered to be implicit in maintaining and updating a to-do list. In the case of externally triggered interruptions, the observation is that if the trigger is of sufficient priority, then this causes the relevant activity to become the current activity.

Internally triggered interruptions are taken to imply that there is some internal call-up which alerts the ovenman at some point in the future. This, however, implies in turn that the only reason the ovenman is internally interrupted is because he set up such an intention in advance.

The ovenman is also considered to construct a plan to guide his sequence of behaviour, although only for the forthcoming 30-60 minute period. The observation that much behaviour is not in fact directly determined by external events is taken as evidence for the existence of a plan, and one of the hypothesised roles of the executive routine is "handling and resolving the often conflicting demands placed on the operator by the requirements of the production process", and thus constructing the plan.

The only other contribution to this section is that of Cypher (1986). This is also a description of multitasking behaviour, although in far less depth than Beishon, and with different aims. The interesting observation raised by Cypher, and indeed also alluded to by Beishon, is that in some instances, the sequence of tasks performed can be accounted for by a "while I'm at it” rule. This is intended to account for those occasions when people do something that they had not necessarily intended to do at that moment, merely because they find themselves to be conveniently situated for it.
The reason for not using Beishon's model any further in this project is largely because it is a model specifically of an ovenman's behaviour. This is both directly and indirectly a problem. Directly it is a problem because it is the goal of this project to develop a model which is a) general, and b) just concerned with the multitasking elements of a given job (the extent to which these two are the same is arguable). The indirect problem is the level of description. Beishon's model utilises quite a low level of description, where the model is expressed in terms of individual decisions. The suitability of this depends on the intended purpose of the model - that of the model to be developed here has already been specified. It was felt that the level of detail in Beishon's model is lower than necessary for the current purpose, and also possibly too ambitious given the other constraints on development. Note, for example, that it has already been stated that the focus is on the phenomena associated with a planner (the what and the why) rather than the detailed mechanism (the how). It was decided that the benefits of starting afresh outweighed those of persevering with the existing model and possibly having to try to extract just the general features.

One of the main benefits of starting afresh in this manner is being able to integrate some of the later developments - in the understanding of planning for example (see above). Beishon's model is quite old in this respect, and although it has a planning element, it was felt that this should play a more prominent role.

To summarise the contributions of this section, it has been possible to identify several potentially interesting BPs. These would include the observation that people multitasking can and do plan or coordinate their activities, but at the same time may choose to do something opportunistically in passing. Finally, it is important to consider two different sources of interruption.

3.5. Interruption

Whereas there is a great deal of literature concerned with, for example, attention and performance, the same is not true for interruption. Beishon (1969), discussed above, provides a classification of interruption, where this is some unexpected deviation in the ongoing sequence of behaviour. The type of interruption discussed in this section corresponds most closely to his external category.

Field (1987) is a report of an experiment which shows that interruption does actually affect performance of the interrupted task when it is resumed, something which he says "has never been shown experimentally". He adds that he is unable to show any effect of task complexity.

The only other contribution to be found is that of Gillie and Broadbent (1989) who provide a deeper investigation into several factors which might be expected to affect
post interruption behaviour. They conclude that the nature of the interruption task is important, but that the opportunity to control the precise point of interruption is not important. Strictly, they do not show this; rather they show that providing an opportunity to rehearse before commencing the interrupt task has no effect. It is as much as can be concluded for the present to say that interruption is likely to affect performance, and that the nature of the interruption may well be important. More than this seems unclear.

4. Summary of Behavioural Phenomena

The following is a list of some of the behavioural phenomena which, given the literature reviewed above, might reasonably be expected to be present in some instances of multitasking behaviour. There are several reasons why such a set may not correspond with what is actually observed. Firstly, the literature may be inaccurate (given the level of detail and the sampling, this is thought to be unlikely). Secondly, the literature may have been wrongly interpreted either in itself or with respect to multitasking. Thirdly, the observed behavioural phenomena may (are likely to) depend on the method of looking - so, for example, some may only be identifiable via a concurrent verbal protocol, whilst others might be transparent to such data capture.

• Interleaving

It can safely be assumed, from e.g. the observational data of Beishon, as well as from general observation, that people are capable of, and are likely to, interleave the performance of what might be considered as different tasks.

• Can only think about a single task at a time

Attention, in its psychological sense is taken as being a unitary property of the individual, and one of the limiting constraints on the performance of parallel tasks.

• Can only do a single task at a time

For the present purposes, it will be assumed that just as a person’s mental capacity is limited to one task at a time, so is a person’s physical capacity. Admittedly, this is probably a limiting case, but it is thought to be not unreasonable given that a person has only one pair of hands, which are typically used together.

• Parallel mental activity (i.e. of automatic skills)

It is possible for a person to be mentally concerned with a different task from the one with which they are physically concerned given certain considerations about the degree of experience with the physical task.
• Interruption may be disruptive

Interruption, in the form of unexpected events, requiring a person to change the task with which they are concerned, is likely to affect the performance of that task in some way.

• Planning ahead (i.e. coordinated interleaving)

It is likely that if a number of tasks must be interleaved, then there will be a degree of anticipation and forward planning.

• Plans likely to be partial, and incomplete in terms of both time and detail

Any plans or strategies constructed to coordinate the interleaving of activities will not be expected to be exhaustive specifications, but rather more general guides for behaviour.

• Planning may be time or event driven

Such incomplete plans will necessarily have to be made at least more specific as the time for their execution approaches, but one of the assumed reasons for incompleteness is that the premises upon which the plan is constructed are likely to change (i.e. an event happens), thus forcing the plan to be changed.

• Some tasks may be done in passing

Both Cypher and Beishon observed that a task might be performed “while I’m at it” rather than at some future planned point.

• Some information is bound up with remembered plans

It is expected that some of the information relevant to performing a task, and thus potentially relevant to planning the performance of that task in the context of other tasks, will be implicit in the stored plan for performing that task, and thus may only become known during the performance of the task.

• Some information is directly available from Long Term Memory (LTM)

In contrast to the above, people can be quite reasonably expected to have a great deal of information about a task readily available in memory for recall as necessary.

• Plans may be mentally rehearsed/executed

It should be possible for a person to think through the performance of a given task - this is a possible mechanism for recalling information associated with a task.
5. The Model

5.1. Introduction

This section takes the JAs and BPs identified above and proposes a model which, taking the JAs as input, has the capacity to produce at least the majority of the BPs as output. Although this section will present the model in terms of its structure and then the mechanism supported by this structure, the model implicitly includes far more than this. For example, it implies a particular way of viewing the world; in this case the view of tasks and resources adopted from the above discussion of Critical Path techniques in Operations Research. These assumptions of the model will be made explicit as necessary.

5.2. The Two Foundations of the Model

One of the conclusions which could be drawn from the foregoing discussion of planning, and indeed also the model of Beishon, is the necessity to consider planning and execution together when modelling behaviour. This duality is not visible in the structure of the model, but does exist in terms of the two analogies recruited for the model. These two foundations will be discussed in turn below, following which, their embodiment in the model will be presented, first structurally, and then mechanistically.

5.2.1. Planning part

It was felt that constructing a plan from a set of constraints might more generally be thought of as solving a problem to produce a solution. The specification of the tasks to be done, and the available resources then constitutes the problem. The solution thus becomes the plan for interleaving these tasks.

Since the problem will be both incompletely specified and constantly changing, the solution could be thought of as the most appropriate for the problem as it is currently specified, but not necessarily complete either.

5.2.2. Execution part

Beishon's approach used a computer program analogy as a basis for his model, and he had to invent his own basis for managing the separate tasks. Since then, there have been developments in computing which suggest a more appropriate analogy.

5.2.2.1. Microsoft Windows

Windows is a program environment, written by Microsoft Ltd, for the IBM PC family of computers. Its most obvious feature is that it provides a graphical WIMP
interface for an otherwise character and command line orientated machine. Less obviously, it implements multitasking on a computer which would otherwise not permit it. The IBM PC family of computers, along with most micros, have only a single processor and typically run only a single program at once - which then has exclusive use of the processor.

This is analogous to the unitary nature of attention described above, and since Windows provides a mechanism for sharing such a processor between several programs, it seemed plausible that the same mechanism might provide an appropriate way of modelling the sharing of a person's attention among several tasks.

In Windows (see e.g. Petzold 1988), each program is associated with a separate task. To manage these tasks, Windows utilizes the object metaphor typical of Object Orientated Programming (OOP). Each program or task is therefore an object (there can also be objects at lower levels than this, i.e. within a program). In this sense, an object is a self contained entity which performs certain actions in response to messages. Messages are the mechanism by which Windows achieves the sharing of the processor.

In essence, each task/program/object has a message queue, into which the system puts messages addressed to that object. If a program has messages in its queue, then it is permitted to run (i.e. it gets the processor) in order to process the messages until the queue is empty. When the processor is then no longer required, it will be allocated to some other program so that it may process its pending messages. In practice, the amount of execution in response to a given message is kept deliberately low by the programmer so that no single program appears to have a monopoly on the processor for very long.

This type of multitasking is known as non-preemptive. This means that once a task has the processor, it has it for as long as it wants - it must surrender it voluntarily to another task rather than have it snatched away. Alternatives in computing might include a strict circulating of the processor on a time basis - so each task would get it for a milliseconds, whether it needed it or not, or a more sophisticated preemption scheme which might distinguish tasks of different importance. The advantage of a non-preemptive system, such as that used by Windows, is that it should be more resource efficient, since the processor is only allocated where it is needed. One of the hidden costs is that the individual tasks must cooperate as described above.

For the time being, this will be the only aspect borrowed from Windows - also, this will be approximately the level of detail which will be used.
5.3. Structure

The main structural elements of the model are indicated in Fig 3.1. In the model, each task is represented by a separate Task Unit - there are four of these illustrated in the diagram, on the left hand side. Given the Windows analogy, these are objects, and are likened to separate programs.

Continuing the OOP metaphor from Windows, each Task Unit encapsulates both the relevant procedural knowledge (i.e. the instructions or script) for carrying out the task, plus a certain amount of explicit, declarative knowledge associated with that task. Such knowledge might include things like the resource demands of some parts of the task, or the duration of some of the stages. This explicit knowledge is thought of as being similar to, and serving the same purpose as, Beishon’s Lookup Tables. Importantly though, this is thought to be only a subset of all the possible information which is known about an activity, the rest being implicit and only available by further consideration of the procedural knowledge. In this way there is no single repository in the model labelled LTM. General knowledge, not associated with any particular task, is taken for granted in the model.

Shown on the right hand end of each Task Unit are two Message Queues. In Windows, each program (i.e. Task Unit) has a single message queue which allows the unitary processor to be shared around. From the above consideration of attention
and performance, it has been proposed that there are two such unitary properties in the person. The first of these is the Mental Workspace, and as such is most nearly analogous to a computer’s processor. The other of these will be termed the Performance System, and will serve as a convenient expression of the person’s ability to effect something in the world. The simplifying assumption is that this also is unitary in capacity, and hence there are two message queues - because there are two essentially independent unitary processors which must be shared. Neither the Mental Workspace nor the Performance System is directly represented in the model, rather both are implied by the existence of the two queues.

In the diagram, all messages pass through a central point. This pathway will not be elaborated on, save to say that it has the potential of being unreliable or noisy should this be necessary (to accommodate, for example, forgetfulness or otherwise imperfect behaviour).

On the right of the diagram are three boxes labelled the Problem Representation, the Solver, and the Solution. The Problem Representation contains the specification of the current interleaving problem - i.e. the tasks to be done together with some indication of both the resources they will require and the resources available. The Solver is responsible for taking the current problem and producing a solution. The Solution is then a specification of what order the tasks will be done. The current bias is towards understanding the planning aspects of multitasking in terms of what the plan looks like, rather than how it is constructed. It is for this reason that the Solver is represented in the diagram as a black box, and it will be taken as such - i.e. there is no current intention to account for its contents.

Certain structures are taken for granted; for example, the perceptual and physical effector systems are not represented.

5.4. Mechanism

The mechanism described in this section, in association with the above structure, is able to provide at least a basic account of all the BPs summarised in section 4 (above). However, the account provided is quite shallow, owing to the feeling that some experience and observation of the actual BPs is necessary before any more detail is added to the account of the model.

The general mechanism is that messages get sent between the different structures, some requesting information, some supplying information and others instructing a Task Unit to perform some action (either mental or physical). Such command messages will originate in the Solution, from its time based plan. This plan is thought of as the best solution for the problem as it is specified (see section 5.2.1.). If the problem changes - because some new information enters the Problem
Chapter 3

Representation, then this Solution is invalidated. An invalid solution is taken to be a sufficient trigger for the Solver to act, thus revalidating it. In this way, the Solution is kept as well specified as possible.

At the beginning of a Job, the specification of the job will put the necessary Task Units in place, with the possibility of adding more when necessary. At this point, both the Problem Representation and the Solution would be empty and thus nothing will be happening. The empty Problem Representation is conceived of as sending out messages requesting information from each of the Task Units. Any of the Task Unit’s explicit knowledge is made available to the Problem Representation. For access to any information implicit in the procedural plan associated with a Task Unit, a message must be placed in the queue such that the Task Unit mentally executes this code, in the manner of Byrne described above. Such Thinking Ahead can take place at any time, and may, for instance, take the form of a Parallel Mental Activity, in that it happens whilst the person’s Performance System is engaged independently.

The determinant of the behavioural sequence, then, is information entering the Problem Representation. Doing things in passing might be explained by a change in the Problem supplied by the perceptual system such that a re-solving creates a plan in which an activity is moved forward and done immediately. Similarly, other externally triggered interruptions would be the result of a change in the contents of the Problem Representation. Internally triggered interruption follows the same mechanism also, with the exception that the source of the information entering into the Problem Representation is not so clear. In the case of a spontaneous change of task, a source of the triggering information might be the knowledge implicit in one of the Task Units. In the case of periodic monitoring, as observed by Beishon, the intention to monitor could be set up in advance, in which case it might be thought of as appearing in the Solution.

That interruptions might be disruptive, could result from the interruption altering or even destroying the contents of the Problem Representation such that the person has to stop and think where they were up to before continuing. More knowledge of the effects of interruption is necessary here. This sentiment applies more generally - i.e. it is thought that some experience and observation of the actual BPs is necessary before any more detail is added to the account of the model.

5.5. Problems with the model

The above description is necessarily incomplete; there are many points which are glossed over and some properties which could already be argued to be inadequate. The goal is not to produce a description capable of implementation as a computer program, for example, but rather it is part of the strategy to produce a prototype.
which may be flawed, but which can be used as a reasoning device. The potential problems with this model include:

• The interruption mechanism. This seems inelegant, but it is probably not worth investing effort here until more is known of the phenomena to be produced.

• Task Units. The criteria for breaking a job into its constituent tasks are fairly arbitrary. Consequently, these model structures are also fairly arbitrary and thus areas of potential weakness.

• The unspecified nature of the messages.

These problems, and others, can be addressed in several ways in subsequent generations of the model. For example, the level of description of the current model is arbitrary, and may not be compatible with that supported by the chosen method of observation and data analysis. If the data supports a low level of detail, then it may be possible to increase the detail and precision within the model, alternatively, it may be necessary to de-emphasise certain specific aspects in favour of more general functionality.

5.6. Summary

The model presented in this chapter is based heavily on an analogy drawn from Microsoft Windows. It represents individual tasks as objects, in the OOP sense, and provides for a non-preemptive sharing of two unitary capacities (the Mental Workspace, and the Performance System) by the use of messages. The coordination of activities is represented as a Problem and a Solution with the concept of invalidity driving the maintenance of the Solution.

The next chapter is an account of an observational study, providing a set of BPs which is used to develop the model. The development reported goes some way towards addressing some of the flaws and weaknesses of the current model.
Study 1:
Cooking

Prologue

This chapter is the first full cycle in the development of the model. It takes the prototype model proposed in the previous chapter and consolidates it in the light of observations made in a real world instance of multitasking - cooking. The scope (in terms of JAs) is not extended in this cycle, but the model itself undergoes a certain amount of development.

Contents

1. Introduction
2. The (Abstract) Job Specification
3. The Study
   3.1. Specific Job Details
   3.2. Subjects
   3.3. Data Collection
   3.4. Using Concurrent Verbal Protocols
   3.5. Analysis Method
4. Results
   4.1. Horizontal
      4.1.1. Initialisation Heuristic
      4.1.2. Free Time
      4.1.3. Default Activities
   4.2. Vertical - Debugging
      4.2.1. Task switching
      4.2.2. Scheduling is Imperfect
   4.3. Vertical - Consolidating
      4.3.1. Information is implicit in stored ‘plans’
      4.3.2. Behaviour is largely controlled by a schedule
      4.3.3. Mental Workspace & Performance System can be concerned with two different tasks
      4.3.4. Monitoring
4.4. Vertical - Unpacking

4.4.1. Form of the expressed schedule

4.4.2. Preserving the Schedule

5. Summary of Behavioural Phenomena

6. The Model

6.1. The (dual) foundations of the model

6.1.1. Execution

6.2. Structure

6.3. Mechanism

6.4. Summary of the Model

6.5. Problems with the Model

1. Introduction

This chapter provides the first cycle in the development of the model prototyped in the previous chapter. Its product is a model with much in common with that in Chapter 3, but which has evolved in certain important aspects as a result of the observed phenomena.

The evolution of the model is as a consequence of the behaviour observed in cooking. Cooking, as a job, was chosen for several reasons. Already stated is the preference for real-world, ecological observation over a more contrived laboratory job. Cooking presented an example of this which had the additional benefit of having a large potential subject population (most people can cook to some degree), many of whom could be expected to have had a relatively large amount of practice.

The other advantage of choosing cooking for a first attempt at observing multitasking behaviour, is that it is readily comprehended, both from the point of view of the amount of immersion in the material demanded by the chosen, loose method of analysis, and also from the point of view of presenting the subsequent data. It is thought that progress would have been held up had this first study been concerned with a less accessible instance of multitasking.

2. The (Abstract) Job Specification

The cooking job used in this study can be specified in the abstract, in terms of its Job Attributes (JAs). This is necessary to understand the contribution of this particular cycle to the scope of the final model.

Rather than go into a commercial or even a domestic kitchen as a ‘fly on the wall’ and simply observe whatever happened to be going on, it was decided to gain a little more control, and thus aid the subsequent analysis, without surrendering too much realism. To this end, subjects, who were not professional cooks, were asked to cook a
particular meal (just the main course), for one person, in a setting which was as nearly domestic as possible. The meal was chosen to be both realistic and believable, and at the same time present a reasonable multitasking challenge.

At this point in the development of the model, its scope was expressable in terms of just three JAs (see Chapter 3). There was a degree of flexibility in the precise format of the job in terms of the particular details of the meal to be cooked. This was utilised to choose a meal whose creation fell within the scope of the model. This first observational cycle, then, mainly explores the existing scope of the model and does not contribute greatly to its extension. It is, however, the case that some interesting features of this job only become apparent in retrospect, from the point of view of subsequently observed jobs.

The first JA is that there is interleaving of tasks. This is certainly satisfied in the present case. One of the constraints of cooking a meal is that the constituent elements should be ready at approximately the same point in time. These elements often have to be cooked in quite different ways, therefore entailing a degree of coordination. In the study, this coordination requirement was enhanced by the restricted availability of the necessary resources (e.g. saucepans).

The second JA from the last chapter concerns the degree of familiarity with the individual tasks to be interleaved. A suitable expertise was achieved by choosing tasks requiring only a reasonable level of cooking ability, and at the same time attempting to attract subjects with such an ability. The exact combination of tasks in the job, however, was contrived such that it was unlikely that any subject had previous experience of the exact set.

Finally, the definition demands that the multitasking situation be not entirely predictable (and thus with the potential to change at short notice). In fact, the job was such that there were no contrived interruptions and it was therefore fairly predictable. However, there was still ample scope for minor interruption and some unpredictability.

It will become apparent, by contrast, in later studies that two other attributes of this particular instance of multitasking could have been of interest. These are that it is not 'open ended' i.e. the job is self contained, and the observation is not of an arbitrary period in a longer ongoing job, and also that there is no externally imposed time deadline. In this sense, it could be considered that the critical path is not of a fixed length, as it is in some other jobs.
3. The Study

3.1. Specific Job Details

Two different Jobs were used in the study, different subjects being assigned to one of the two (see below). Both jobs, i.e. meals, had the same general goal, which was that they should be cooked in the shortest time possible, assuming also that the final product be edible. Both jobs shared the same constraints such that they satisfied the same definition of multitasking (see above).

Specifically, it was thought that there should be some parts of at least one of the tasks which would require monitoring, and that the various tasks should compete for resources in some way. At the same time, the different tasks had to be of adequate duration such that any competition was realistic, but without making the attainment of the overall goal impossible.

The first job (Job 1) required the subject to prepare a meal consisting of a fillet of baked fish, boiled new potatoes, mangetout, and a cheese sauce. The fish in this case was either Haddock or Cod, depending on availability.

The second job (Job 2) consisted of preparing and frying a trout in butter, with roast potatoes and boiled carrots.

The jobs were conceived, for the purposes of analysis and modelling, as being split into tasks concerned with the preparation and cooking of the separate ingredients.

Subjects were given all the necessary ingredients and resources (tools) to prepare their particular meal. Both the saucepan and the cooker resources were constrained to affect the job. There were only three saucepans, and also only three rings on the stove. Whilst there were only ever three tasks requiring a saucepan and a ring, both of the latter varied in size considerably (one each of large, medium and small) such that some were more suitable for some tasks than others. In this way, although there was no absolute competition for resources, there was competition in terms of suitability.

The observations were conducted in a small semi-commercial kitchen (normally used for the preparation of sandwiches). This was the only available location suitable for observation, and was sufficiently close to a domestic kitchen in terms of its appliances. The basic layout of this area is illustrated in Fig. 4.1.
The general procedure was that subject were shown around the kitchen and it was made sure that they were familiar with the oven, for example. All the necessary resources and ingredients were laid out for them on one of the worktops. There were two instruction sheets (see Appendix A), the first one dealing with the general aspects of the study and instructions for the verbal protocol, the second being more specific and detailing the meal to be cooked. After answering any queries, subjects were left to proceed as they wished.

### 3.2. Subjects

A total of nine subjects were observed in total, consisting of 5 female and 4 male subjects. All were in the age range of 22 to 35, and all professed to be keen and competent cooks. None of the subjects were paid for their participation, but were offered the opportunity of eating the results of their labours.

Subjects were allocated to one of the two conditions on an alternating basis, starting with subject S1 on Job 1. It was necessary to break this sequence for subject S7 owing to the unavailability of suitable fresh ingredients that day. For convenience, the allocation scheme is summarised in the following table (Table 4.1.).

![Plan diagram to show the general layout of the kitchen used in the study.](image)
Table 4.1. Allocation of subjects to Job 1 or Job 2.

<table>
<thead>
<tr>
<th>Subject</th>
<th>Job</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1</td>
<td>Job 1 (Haddock/Cod)</td>
</tr>
<tr>
<td>S2</td>
<td>Job 2 (Trout)</td>
</tr>
<tr>
<td>S3</td>
<td>Job 1</td>
</tr>
<tr>
<td>S4</td>
<td>Job 2</td>
</tr>
<tr>
<td>S5</td>
<td>Job 1</td>
</tr>
<tr>
<td>S6</td>
<td>Job 2</td>
</tr>
<tr>
<td>S7</td>
<td>Job 2</td>
</tr>
<tr>
<td>S8</td>
<td>Job 1</td>
</tr>
<tr>
<td>S9</td>
<td>Job 1</td>
</tr>
</tbody>
</table>

3.3. Data Collection

Data were collected in the form of a video recording of the entire session together with a separate audio recording of the concurrent verbal protocols. The researcher was always present throughout the experimental session to answer any queries about the tasks, to take any additional notes, and to prompt the subjects to keep talking if and when necessary.

It was necessary for subjects to be fitted with a small portable tape recorder in a pouch on a belt around the waist. This enabled a reasonable voice recording to be made whilst they were free to move around the kitchen.

3.4. Using Concurrent Verbal Protocols

The primary source of data for the development of the model is in the form of statements taken from concurrent verbal protocols. However, there are contentious issues surrounding the use of verbal data in psychology. The purpose of this section is to review these issues such that the status of the data to be used can be understood. There are basically two separate areas to be discussed. Firstly, whether it is at all appropriate to collect verbal data, and secondly, given that one accepts that it is, what are the constraints on collection and inferring behaviour.

The major argument against relying on verbal data is provided by Nisbett and Wilson (1977). They provide a comprehensive review of the literature along with some of their own evidence, and conclude that people have little if any introspective access to higher cognitive processes. For example, subjects are reported as variously unaware of the existence of a stimulus (which has influenced a response), unaware of the existence of the response, or unaware that a particular stimulus has affected the response. Any reports that people attempt to make, they suggest, are likely to be
based on causal theories or judgements of plausibility. Naturally, this may sometimes lead to ‘accurate’ reports about cognitive processes, but not as a result of the simple accuracy of introspection. Other researchers have found the accuracy of verbal reports to be doubtful. For example, Berry and Broadbent (1984) report three experiments relating the performance on a cognitive task with the subjects’ explicit reportable knowledge associated with the task (accessed via written post task questionnaire). They report that performance on the task improved with practice as expected, but that there was no similar improvement in the ability to answer questions about the task. They are able to produce the opposite effect by using verbal instructions (i.e. with no effect on control performance).

There seem to be two main arguments against the position taken by Nisbett and Wilson (1977). The first is exemplified by Smith and Wilson (1978) and centres on methodological and theoretical issues (but see also White, 1988). The second is the position adopted by Ericsson and Simon (1980, 1984). It is worth first mentioning the simpler distinction made by Norman (1983), who proposes that the tendency of subjects to generate reasons for their behaviour are more associated with retrospective verbal reports and that online, concurrent, verbal protocols should be less susceptible to this.

Smith and Wilson (1978) criticise Nisbett and Wilson for wanting to use the existence of both correct and incorrect subject verbal reports to support their position. They point out that they are implicitly using an ‘impossible’ criterion for introspective awareness in the experiments that they review - that subjects are aware of what psychologists systematically (and effectively) hide from them by experimental design. Smith and Wilson offer the conclusion that to claim that introspective access to higher cognitive processes is never possible is too strong. Instead they maintain that the aim should be to consider the conditions which govern such access. Indeed, Nisbett and Wilson do state that it should be possible to provide reliable verbal reports in some situations (p 245). Ericsson and Simon (1980, 1984) provide a model of cognitive processing which allows them to reason about the ability of subjects to verbalise reliably. Fortunately, the conditions of successful verbalisation they propose are in agreement with those of Nisbett and Wilson.

Ericsson and Simon establish an information processing model of cognition, not dissimilar to the general, framework, models discussed by Reason (see Chapter 3). They then discuss verbal reporting in the context of this model, which enables them to make explicit and understand the various assumptions. Central to their model is the concept of heeding, also termed attending to, but understood by Ericsson and Simon as being in Short Term Memory (STM). They assume that information in STM can take many forms (verbal, spatial, etc), but to be verbally reported, must be
transformed to a verbal code, and further that only information heeded (i.e. in STM) can be reported in this way. This raises two points - i) what information is in STM, and ii) whether it needs to be recoded.

In this light, they distinguish three levels of verbalisation based on the intervening processes between the attention to the information and its delivery as the spoken word. Level 1 verbalisations are direct, i.e. they are a report of the contents of STM, which are assumed to be in a verbal code such that no transformation is necessary. Level 2 verbalisations are similar to level 1, in that the contents of STM are reported 'as is' except that they may be other than a verbal code and therefore may need to be transformed. The final classification is level 3 verbalisation. In level 3 verbalisations, there may need to be recoding as in level 2, but in addition there may be some intermediate scanning/filtering or generative processes. An example of a level 3 verbalisation would be one where the subject was required to report only a subset of heeded information, or alternatively report information which would not normally be heeded (such as reasons for actions). The assumption is that the level 1 and 2 verbalisations do not change the sequence of heeded information and therefore interfere less with actual behaviour.

Ericsson and Simon’s arguments apply to verbal data in general, which can take many forms, only one of which - the concurrent verbal protocol - is of interest here. They comment that this form of verbal report might claim to be the closest reflection of the ongoing cognitive processes, since the states of STM are verbalised directly by thinking aloud. To encourage such reports to be of level 1 or 2, they suggest that a general verbalisation instruction should be used (where the subject is asked to report everything, even if they do not consider it relevant) rather than a specific probe. The value of information which is volunteered without probing is potentially higher than that which is obtained by probing since in the latter case, the probe might lead the subject into a particular, possibly inferred, answer. Finally, on the basis of the role of STM (the Mental Workspace) in controlled and automatic modes of behaviour, and the relation of these modes of behaviour to skill level, it is to be expected that there might be a difference in the content of verbalisations of novice over skilled subjects.

Given, then, that there are types of behaviour which subjects ought to be able to report accurately, it is necessary to consider the actual collection and what effects this can have, both on performance and the data themselves. Bainbridge (1990) underlines the importance of considering the social situation surrounding the data collection. Asking people to talk to themselves is unnatural and can be uncomfortable - one of her suggestions is that getting two people to work together may result in more natural protocols, a suggestion also supported by O'Malley at al (1984). People often find it difficult to admit their weaknesses, which may be of major interest in a
protocol, and so steps could productively be taken to minimise any such tendency, for example by the establishment of a level of trust between the researcher and subject. Equally important would be the domain knowledge of the researcher (Bainbridge, 1979). This would obviously be important at the analysis stage, but during the collection phase might influence what the subject felt could be taken for granted or might lead to them feeling the need to go into too great a level of detail.

Such points emphasise the lack of a clear division between effects on performance and effects on the nature of data collected. In all cases, however, there is a need to ensure that the method of data gathering is not itself affecting the data, i.e. the observed behaviour (more than might be reasonably acceptable).

In the Ericsson and Simon model, the effects of verbalisation on performance will depend on the level of the verbalisation. From Bainbridge (1979) it is possible to conclude that the effects of verbalising (concurrently) on many tasks should be nil or actually positive, a position borne out to an extent by the results of Berry and Broadbent (1984).

Summarising the position so far, it seems clear that given certain constraints, it is possible to successfully collect verbal protocol data from subjects with some confidence that these data bear a relationship to the underlying cognitive processes. The next step is to understand what can be done with these data, and how.

It is necessary to make some assumptions about nearly all sources of data in order to be able to interpret them. This is especially true for verbal data. The extent and importance of these assumptions depend on what one intends to do with the data. Verbal protocol data cannot be used to test theories of mental behaviour since there is no strict correlation between behaviour and report - we may be conscious of the results of mental operations, but not necessarily how the judgements were made (Bainbridge, 1979). However, it is suggested that as a source of evidence, it correlates sufficiently with observable behaviour to be useful for some things, for example, hypothesis generation. In fact, it is suggested by Ericsson and Simon (1984) that if the intention is to use verbal reports only to generate ideas and hypotheses, then there is very little need to be concerned with the methodological issues which would be more applicable in the case of testing or verifying such ideas.

In passing, it is encouraging for the present purposes that Bainbridge (1979) remarks that scheduling decisions, i.e. decisions about how to make best use of several machines, appear to be described in verbal protocols more explicitly than dynamic control decisions.
A verbal protocol, once collected, can be analysed in a large number of ways. Analyses can be exhaustive classifications of the phrases of the entire protocol, from which quantitative analyses can be attempted. Alternatively, the analysis can be less formal. Bainbridge (1979) makes the useful point that if a behaviour occurs a number of times in a protocol, but is only imperfectly reported each time, it is often possible to construct a fuller picture by combining reports.

In the light of all of the above, the verbal data acquisition in this project can be suitably constrained. Firstly, the data will be used to generate hypotheses about behaviour, rather than to test a priori hypotheses, although there will be an a priori interpretation framework in the form of the current version of the model of multitasking behaviour. Such an interactive form of data analysis places fewer demands on the nature of the data. To avoid affecting the behaviour under observation more than is absolutely necessary, a general verbalisation instruction will be used, thus increasing the probability of level 1 or level 2 verbalisations. The only qualification to this is that subjects are encouraged to talk in a future tense. It is assumed that this is a cue to aid verbalisation rather than one which will result in generalisation, but this is arguable. Any probing or encouragement for subjects to speak is kept to a minimum, and confined, as far as possible, to statements which are not leading. An effort will be made to make the subjects feel comfortable and to reassure them where necessary that anything they say will not be made available to, e.g., management where it could be used against them. It will also be apparent that the three different jobs observed in the course of this research are such that the researcher could easily assimilate the necessary domain knowledge. As discussed, this has benefits for both collection and analysis. Issues of interpretation are provided for by presenting the data together with its interpretation for the reader. This of course makes the exposition longer, but more understandable. Finally, the development methodology has built into it a validation by replication scheme. This does not necessarily neutralise any interpretation bias, but does help in the area of subject bias (for example in inferring answers).

### 3.5. Analysis Method

In the event, the data from the video recordings were not used (and video recordings were not made in the subsequent studies). The main emphasis is on the verbal protocol data. Each protocol, which was between 45 minutes and an hour in duration was transcribed in full. A complete sample protocol from this study is provided in Appendix A.

Analysing the protocol data required both several passes through the data, and that the researcher had a good understanding of the ongoing tasks. There was a great reliance
on intuition and judgement, and also the notes made by the researcher at the time. The early passes over the transcripts - indeed the act of transcription itself - served to suggest categories of interesting phenomena. Subsequently, as many instances of each of these categories as possible would be found in each of the transcripts. Further passes over these excerpts consolidated each category by removing the weaker member, and also allowed for categories to merge or split. By this stage it was possible to relate the categories to the existing set of BPs, and create new BPs as necessary. To give some indication of the utterances used in the analysis and those ignored, the full example transcript is annotated further (see Appendix A).

It should be noted that no quantitative measures of performance were attempted. It was deemed sufficient if the meal was cooked and edible at the end. All meals were edible, and the subject had the opportunity to eat the meal that they had prepared.

4. Results

This section presents the observed behavioural phenomena, mostly in terms of excerpts from the verbal protocols. The phenomena are grouped according to their contribution to the development of the model, as discussed in Chapter 2. There are differences between subjects, both in the degree of verbalisation and in the way the tasks are performed. Nevertheless it is possible to find an acceptable level of agreement across subjects. Wherever possible, sample statements are included from more than one subject to show this.

Table 4.2. shows the conventions used in the protocol excerpts. In addition to the excerpts given in this chapter, a full transcript is included in Appendix A (and similarly for Chapter 5, one is to be found in Appendix B, and for Chapter 6, Appendix C). Where this subject's protocol is referenced in the chapter, it is given a superscript indicating the line where it occurs.
Table 4.2. Table of conventions and their meanings as used in all the protocol extracts in this and subsequent chapters.

4.1. **Horizontal**

Horizontal development is concerned with the extension of the scope of the model. The Phenomena in this section, then, are ones which the model has difficulty addressing. They differ from those which appear under the Vertical-Unpacking heading in that the latter are not considered to be directly the result of the JAs particular to the current job.

4.1.1. **Initialisation Heuristic**

In the model, the schedule is derived from a problem representation built up from data. More of this data becomes explicit as the tasks progress, either mentally or in reality. A question then arises as to how an initial schedule is arrived at, which allows the job to start as soon as possible, and provides something which can be amended. There are several of examples of what one might tentatively label heuristics to provide this initial foot-hold.

It should be noted that a good initialisation heuristic should have the same effect as working out the critical path in full - i.e. it should produce a schedule that can be used as a starting point, but without the associated overhead. There are two variations on this theme in the protocol data. Firstly there is simply identifying the longest activity and assuming this will be the critical path (which is not necessarily the case):

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fred</td>
<td>All personal names have been changed</td>
</tr>
<tr>
<td>(Fred: )</td>
<td>Comment by 'Fred' - i.e. not the main subject</td>
</tr>
<tr>
<td>( )</td>
<td>Comment by the researcher</td>
</tr>
<tr>
<td>[ ]</td>
<td>Additional explanation of the ongoing behaviour</td>
</tr>
<tr>
<td><em>italics</em></td>
<td>Emphasis of a sub-section of an excerpt</td>
</tr>
<tr>
<td>...</td>
<td>Indicates part of an excerpt missing either because it was unintelligible or irrelevant</td>
</tr>
<tr>
<td>S9</td>
<td>The subject number from which the excerpt is taken</td>
</tr>
<tr>
<td>S9 124</td>
<td>The superscript number is the line number of the excerpt in the full transcript given for this particular subject in the relevant appendix</td>
</tr>
</tbody>
</table>
The potatoes will take a bit of time. But the fish takes 25 minutes, so I'll get on with the fish first."

The variation on this theme is exemplified by Subject 1, who chooses to do an innocuous task while deciding what to do:

"First thing I'll do while I'm thinking about which order to do it is do the potatoes - it doesn't matter how far in advance I wash these."

Notice that the ongoing task (washing potatoes) is one which is probably able to run in automatic mode requiring only the Performance System, leaving the Mental Workspace free to plan.

**4.1.2. Free Time**

One of the things that was not really anticipated was that there would be occasions when the person would have nothing to do. Although this cannot be directly related back to the Job Attributes (probably because of their level of description), it is nevertheless attributed to the particular instance of multitasking and thus included in this section.

Interestingly, the recognition of such free time (by the subjects) implies the existence of some overall plan of what will be done at what point. The following examples illustrate this phenomena:

"Good, now there's actually going to be erm, quite a bit of spare time, cos the new pots and the cheese sauce will take about the same time to cook. I won't even have to start doing that"

"I need a book. ... I don't think there is anything else I can do at the moment."

"And everything I think is ready, so I'm going to take 5. Start on the boiling of the mangetout at about 5 past, and whilst the mangetout is doing, I can do the cheese sauce, because that's really quick, so I'll just take a little bit of a break."

**4.1.3. Default Activities**

In the cooking job presented to the subjects, it transpires that there was an additional, implicit task of lesser importance than the rest; tidying-up or housekeeping. It was observed that Ss would engage in this activity when there was nothing else for them to do, as a sort of default - hence the title of this phenomenon. There was no suggestion that such a task would be scheduled in advance. Considering Free Time (above) in the context of such default activities, the former would emerge when any possibilities of such default activities had been exhausted. This might, however,
depend on the tidiness habits of the subject. There are no protocol data available to support this observation.

4.2. Vertical - Debugging

Vertical debugging considers addresses the problems the model has, within its scope. The phenomena in this section are those which are incorrectly predicted by the model.

4.2.1. Task switching

In the previous chapter, the execution element of the model was based around an MS-Windows analogy. Two main properties of Windows were recruited. Firstly the object metaphor, with its encapsulation of data and procedures, and secondly its non-preemptive multitasking mechanism.

It was also stated in the previous chapter that one of the potential problems with the model as it stood was the inelegant interruption mechanism. This can now be elaborated upon, in the light of this first set of observational data. Interruption is one instance of the general class of task switches. Such behaviour has now been observed, and it is thought that the data in the form of verbal protocol statements do not provide the necessary (low) level of description.

To justify a non-preemptive scheme, it would be necessary to specify at least some of the messages upon which it is based and so capitalise on the particular features of such a system over any other. One of the properties of a non-preemptive system is that tasks can only be switched between the processing of two messages. Turning this round, it means that messages must specify behaviour at the level at which task switching occurs. To give an example, if it were observed in one of the above tasks that subjects always finished everything to do with washing the potatoes before moving on to cleaning the fish, and then finished cleaning the fish before moving on again, then there would only need to be messages of the order of clean the fish and wash the potatoes. Instead, it is observed that task switches occur much more often, in the middle of tasks described at such a (relatively) high level. To persist with the message mechanism, it would be necessary to be able to describe tasks at a much lower level, probably almost in terms of actions, at a level where it can be considered to be atomic. It is not possible to support this using the verbal protocol data.

In terms of the developmental strategy (Chapter 2), this is included as an example of Debugging in Vertical Development, since it is obviously a deficiency in the model, but is thought to be generally within its scope.
4.2.2. Scheduling is Imperfect

In the model, each stream of activity is conceived as a separate entity - a Task Unit structure, which has the potential to complete that task independent of the other tasks. It is the central controlling structures and the schedule that are thought of as providing the multitask context for the individual tasks. There is evidence to suggest that this is the case. For example, take the task of cooking a vegetable such as mangetout. This could be broken down into the sequence of activities “Wash & Chop” then “Boil for five minutes”. If one were doing this task on its own, then there would be no reason not to boil them as soon as they had been washed and chopped. However, in a multitask context this might not be the best thing to do and might result in the task being completed too early. This is illustrated in the following protocol examples.

S8 [Has just finished scrubbing his potatoes:]
“Right, so I'm turning the spuds on at a quarter to 7, they should take about 20 minutes, so targeting about 5 past 7 - But I'll leave them for a bit because the oven still isn't hot enough - it's taking a long time, isn't it? ... OK, well the spuds I reckon should take about 20 minutes so, yes, so I'll hang fire on the spuds for a bit”

S7223 And it's about a minute past 7, so I probably ought to turn the carrots down again, having just turned them up. (Are they boiling?) They're not boiling yet, no, but I want to slow them down now, because I haven't started cooking the fish yet”

Whereas the above illustrate the imperfect nature of the schedule with respect to timing, the same is true of resource allocation. Resources are items, tools for example, that are required by the person for some period of a task, after which they may be used in a different task. In the example of cooking, resources might include saucepans and stove rings.

S8 “Put that on a low heat (Is there any reason why you chose to use that back ring for doing the sauce?) Yes, I mean its a kind of lowish - that's kind of not as strong as these ones, so, I mean its only a sauce, so it doesn't need to be fiercely heated, so I'm using this one because it doesn't matter if its not very strong.”

However, had he objectively considered all the variables, he could have avoided the particular consequence:

S8 “It's a bit silly doing it on the back, because you've got the steaming potatoes in front of you.”

In a similar vein, when choosing which of the three saucepans to use in a particular task, S9 makes a reasoned decision:
"I should have used the other saucepan, I'll have this for the mangetout. I'm not being very clever here. OK, (Why are you going to do that?) Because it's better to use the little one for the smaller quantity"

However, this decision fails to take the other tasks into consideration, with the following consequence:

"the cheese sauce. Ah now, what I've done is I've used - I would have liked to have kept this saucepan for the cheese sauce, quite frankly, because it's more manageable to keep lifting on and off the heat."

- the result is that she takes the water heating for the mangetout out of the smallest pan and uses this pan for the cheese sauce.

### 4.3. Vertical - Consolidating

Vertical consolidation considers those BPs which are addressed satisfactorily by the model and plays on the commonality of some JAs between observed instances to provide support for the model.

#### 4.3.1. Information is implicit in stored plans

The model maintains that some information may be stored implicitly in a task's script. This information becomes explicit when the particular script is executed. However, it is thought possible to access this information in advance by mentally executing the task. For example, the following excerpt occurs as the subject is adding cheese to the sauce in the pan, i.e. as a product of real execution. The absence of mustard is not itself important for multitasking, but it is taken as an example of something which could have been brought up earlier and maybe acted upon.

"I notice there's no dried mustard which I might put into a cheese sauce, normally to save putting so much cheese into it - but seeing that we've got so much cheese, I don't think I'll need it - I'll just go overboard on the cheese."

It is possible to find other cases in the protocol data where subjects do think ahead, or mentally execute the task, and so made explicit important information. E.g.:

"I mean a lot of the recipes you sort of know them in your head and you sort of plan them from ingredients that you know you have - like its just occurred to me now that you may not have any butter, for example, is this true? (There's some there) Oh, that's good - but that would have been a problem for example in that sometimes I might plan to do it this way and you might not have the stuff ... Yes, well it had
occurred to me earlier actually, but I made a contingency plan that I could use a little bit of oil if it came to it.”

S5 “so the potatoes ... I'll put on at about 10 past, after cutting them up. Do I have a kettle? (No), so I'll have to boil the water as well.”

S7 [whilst scrubbing potatoes, i.e. as a parallel mental activity]
“I'm now wondering about the trout, and as it was actually bought for me, I'm wondering whether or not I'll actually have to gut it.”

Also less important information becomes explicit. This may be a reflection of implicit information stored in plans, or it may be more an example of more general association. In fact, irrelevant associated information of the general form “My Granny would have...” was very common in the protocols:

S3 “If there was a steamer, then its much better, because then you get the - you don't lose the vitamins as much as you do this way.”

4.3.2. Behaviour is largely controlled by a schedule

First of all, it can be said that ALL the subjects voiced some sort of schedule, i.e. an intended sequence of (sub)tasks. After this it needs to be established that they engaged in activities in this order (which is only broadly true given the frequency of unforseen activities and/or the high level nature of the schedule) and that they did not repeat the decision making at the changeovers between activities.

For example, the following is a statement made by subject S7 very near the start of the exercise

S74 “OK, well the first thing that I've got to do is turn on the oven, for the roast potatoes for later on, and the next thing that I've got to do is get the potatoes ready so that they can be par-boiled, and I'm going to need some hot water for the potatoes so I may as well put that on next, erm, and I can get the carrots ready at the same time as I'm getting the potatoes ready, and I don't need to do anything to the fish for the moment. So that's what I'm going to do.”

The actual order of activities performed, in the short term that the schedule seems to address, is as stated - turning the oven on, boiling some water, then preparing the potatoes and carrots.

Alternatively, subject S8 make the following statement of intent - i.e. his proposed schedule, early on:

S8 “The potatoes will take a bit of time. But the fish takes 25 minutes, so I'll get on with the fish first.”
And, true to intention, he completes the preparation of the fish, and then decides to get on with the potatoes:

| S8 | “- so I’m now going to go into the potatoes, that’s the next thing that’s er going to take a bit of time.” |

In the model, planning is considered to proceed in an opportunistic manner, where the plan is not constructed strictly forward in time. Also, the final plan is considered to be partial in that actions are ultimately situated in some real, partly unpredictable, context. This is borne out by the data. For example, the following subject has already voiced a high level schedule which involves doing the potatoes followed by the fish. At this point, she has done the potatoes and is now about to consider the fish in more depth before proceeding:

| S771 | “And I’ll just put them to one side and wash the knife, in case I need it in a minute, which I probably will. Right, I’m just going to check now what I’m going to do with the trout now I think.” |

Subject 2, on the other hand, has an *a priori* intention to put the potatoes in the oven and then turn over the the fish. In the event he does it the other way round:

| S2 | “Now this is a dodgy little exercise, ’cos I think these are going to spit, hang on, in fact I’m going to turn over the fish first before I put that in.” |

Given that there is a schedule, it is worth considering whether it is efficient; whether or not the subject is aware of a critical path goes someway towards indicating this. The critical path is that sequence of activities which determine the time that the whole job will take. One of the general instructions on the sheet given to subjects at the start of the experiment was that they should attempt to cook the meal in the shortest time possible, in addition to specifying that all the constituent parts should be ready together. To satisfy this criterion it is necessary both to identify the sequence of activities, and to start this sequence first. The other activities can then be arranged around it. Subjects varied in their appreciation of the critical path, which, assuming a perfect scenario, was as follows:

**Task 1:** Pre-heat the oven - Bake the fish

**Task 2:** Pre-heat the oven - Roast the potatoes
Some identified this straight away, e.g.:

<table>
<thead>
<tr>
<th>S6</th>
<th>[Task 2]</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>“Well I mean, when I look at the menu, my first, erm, product, if you like, is to put them in order of preference, so that the potatoes are obviously erm are the first thing I'll tackle”</td>
</tr>
</tbody>
</table>

The above was one of his earliest utterances after reading the initial instruction sheet. ‘Order of preference’ is taken to mean the chosen order of doing. Shortly after this, and as his first activity:

| S6  | “right, well I'm going to need an oven, so the very first thing I'm going to do is get that oven on, get that heated.” |

Other subjects only realised the Critical Path much later on, e.g. Subject S3:

This subject starts off by preparing the vegetables, and then prepares the fish, after which she is ready to cook it and turns the oven on - i.e. the first activity in the critical path. Shortly after she realises her error, and finds that indeed the job time is determined by the critical path and that she has time on her hands.

| S3  | “I'm not quite sure, er the sauce is going to be too cold I think if I do it now, so I shall leave it as that. And the potatoes are going alright, so I should have put the oven on first. Well, I'm not quite sure what I'm going to do now, since I've got most of the things ready.” |

4.3.3. Mental Workspace & Performance System can be concerned with two different tasks

This is not something which is readily evidenced by verbal protocols, since although they may represent the behaviour of the Mental Workspace, the location of the Performance System has to be filled in separately. Several instances of parallel mental activity have already been described above in other contexts. This section presents some others:

| S1  | “So while I'm scrubbing the potatoes, I'll just read over what I'm going to be doing with the baked fish.” |

and similarly, whilst scrubbing the last potato, but before finishing:

| S746 | “and this is my last potato, so I'm now beginning to think about my carrots and how I want to do those” |
Associated with the notion of not thinking about what you are doing, or even the cost of not doing so, is the prediction of certain classes of errors on the basis of models of attention in action (e.g. Reason, 1986). Such models maintain that the execution instructions / script / schema which runs off automatically, will follow a default set path or set of values based on past experience - i.e. learning. There will be choice points in such a sequence where the behaviour could be tailored to a specific occasion. However, this is taken to require the intervention of the Mental Workspace. An error - i.e. an action not as planned - results when an intervention is missed, for example because the Mental Workspace is engaged elsewhere in some parallel mental activity. There are examples of such slips in the protocol data:

<table>
<thead>
<tr>
<th>S1</th>
<th>“- I’ve actually made sauce for two rather than for one, out of habit,”</th>
</tr>
</thead>
<tbody>
<tr>
<td>S2</td>
<td>“er in fact I’ve done too many potatoes, sorry about that, unthinking,”</td>
</tr>
</tbody>
</table>

Note that it is not possible to illustrate satisfactorily the relevant decision points and show that the Mental Workspace was indeed otherwise engaged at the critical moment. Rather, the above are offered as examples of what would be expected given the proposed model.

4.3.4. Monitoring

Monitoring is not strictly an expected behavioural phenomenon. However, it is the only observed manifestation of two other BPs - namely doing something in passing (“While I’m at it”), and internally triggered interruptions.

Interrupting an ongoing task to check on the state of another without any external trigger or event is very common. E.g.:

<table>
<thead>
<tr>
<th>S1</th>
<th>“Just check the oven, see if that’s ready for the fish yet.”</th>
</tr>
</thead>
</table>

There is also evidence that there is some intention to monitor in advance:

<table>
<thead>
<tr>
<th>S2</th>
<th>“I’m not sure what the temperature is going to be like on this one - so I’ll just have to check that they’re simmering in a few minutes.”</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>“Need to check the potatoes from time to time as well.”</td>
</tr>
</tbody>
</table>

And true enough:

<table>
<thead>
<tr>
<th>S2</th>
<th>“The er, yes, I was going to check those potatoes.”</th>
</tr>
</thead>
</table>

However, the intention to monitor (or not) is not reliable. Take the following example, in which the person decides that it is no longer necessary to monitor the
state of the potatoes, i.e. he is content to go into what could be described as a feedforward mode:

S5 “Checking the pan for the potatoes again - I think I can ignore that now because it seems to be boiling at a constant speed, so that should be ready at half past, along with the fish.”

but later he checks them anyway:

S5 “Checking potatoes again ... (Were you really concerned about the potatoes not boiling, or did you just fancy having a look?). I was just having a look.”

Given the phenomenon of monitoring in this way, a further complication arises in terms of a tendency, for example, to monitor several items in succession. This is assumed to result from their physical proximity:

S1 “Let's have a look and check that the mangetout are coming up to the boil nicely, that's good, the potatoes are boiling well.”

There were only very few observed examples of externally triggered interruption, which may be a result of the particular job, or of the data capture and analysis. Of those which were observed, the following is a good example:

S2 “That fat is starting to smell very hot - I'm a bit worried about that. I think I'll turn down this slightly”

4.4. Vertical - Unpacking

This section presents those observations which, whilst agreeing with the model, also present a challenge in the form of increasing the detail of behaviour which must be addressed.

4.4.1. Form of the expressed schedule

It was observed that when subjects expressed their schedules, they did so in one of two ways (although this could vary within a subject - see later). Some subjects ordered their forthcoming activities with respect to clock times, i.e. absolutely. The alternative was that the ordering of tasks was expressed as relative to each other.

The latter case was probably most common. For example:

S1 “So the first thing is to get the oven on ... the new potatoes and the cheese sauce will take about the same time to cook. I won't even have to start doing that until the fish has gone in. ... and just before I'm putting the cheese in I'll put the mangetout in. New potatoes can go on as I start making the sauce”
and later the same subject continues:

S1 “I’ve got to remember ... and to turn the water up for the potatoes so it’s boiling before I put the pan of water on for the mangetout, ... I’m debating how long the fish will be cooked for and how long it will take the sauce, I think I should put the fish in before I start cooking the potatoes or start doing the sauce, although the potatoes will probably go on fairly soon afterwards.”

The fact that this subject, along with others, does not use an absolute schedule, does not mean that she has no conception of the time that it will take, as the following statement illustrates:

S1 “Check what time it is - just after 20 past 4, so it should be ready at about a quarter to 5 - oh plenty of time.”

A good example of a subject using a real time schedule is S5 (although S2 and S8 are similar). This subject begins by establishing the critical path:

S5 “Erm, well, what I’m going to do first is to do the thing that’s going to take the longest - which is obviously the fish”

and then refining this:

S5 “OK right I’m going to wash the - no, first I’m going to put the cooker on”

at which point it starts to emerge that she intends to rely on clock time:

S5 “Cooker set at 375F and timed to go on at 5 o’clock.”

Shortly after this, she states a more comprehensive sequence of tasks on a real time basis:

S5 “Right, that’s going to go in at 10, 5 past. ... If I put the fish in at 5 past 5, the meal should be ready or the fish should be ready to eat at half past, assuming it takes 25 minutes, so the potatoes, which should take 20 minutes should go on - well 15 to 20 minutes, I’ll put on at about 10 past, after cutting them up. ... Right, the beans will take 5 minutes to cook, so I’ll ignore them”

Later, there is evidence to suggest that she is indeed using such an absolute schedule:

S5 “Right, the water for my potatoes is getting near boiling and its nearly time for the fish to go on.”

However, it is observed that even subjects who construct their high level schedule based on clock time, can end up, later on, with a relative schedule.
S5 "What do I still have to do? And while the fish and the potatoes are cooking, I'm going to make the roux and just before I come to the end of making the roux, I'll put the peas on because I like them al dente."

One can speculate on the relative advantages and disadvantages of one form over the other. For example, it is possible that an initial order of cooking based on clock times is easier to construct, since it can be done with only minimal cross referencing between tasks. Such a schedule may then reveal the points where such cross referencing is necessary.

4.4.2. Preserving the Schedule

It is often the case that an activity turns out to need more or less time than has been scheduled for it (in other words, the schedule was inappropriate or wrong for some reason). When this happens, there are two things that a person could do to rescue the situation. One of these is to reschedule everything else around this new time estimate, the other is to try to preserve the existing schedule and amend the new required time to conform. The protocol data suggest that the latter is the case wherever possible. E.g.:

S1 "the potatoes are taking longer to cook than I thought they would do, so that's going to set my timing off."

S2 "Well these potatoes are looking underdone, I'm not very happy with them at all, so I'm going to stick them back in and turn it up as high as it will go."

S7244 "I think the potatoes are on time, but everything else is late - well no, the trout is late. ... The potatoes look excellent, and I'm going to turn them down I think at this point, because I don't want them to do too much more, so I'll turn them down"

S8 "Just stick a fork in the potatoes and see how they are doing. And they're rock hard, good. And they've got 2 minutes to cook completely. So we might have to adjust our final ..., well I underestimated the time it might take for them to do. I thought they might take 20 minutes, but they are actually quite big new potatoes so they could well take 25 to 30 minutes to cook through properly. ... So what I'm going to do, I'm going to cut them in half now because I don't want to wait. ... this is like a quick remedy, to speed things up so that they're in time with the fish, because I don't want to overdo the fish."

There are no examples of the alternative behaviour, which would involve the person noticing that the scheduled time was wrong and rescheduling all the activities.
In the model, this tendency is obviously a property of the Solver, but it could be thought of in at least two different ways. It could be a by-product of a partial, opportunistic, planner. Such a planner would, by nature, tend towards localised planning rather than global scheduling, and so a plan is more likely to be amended locally in the observed fashion.

Alternatively, one could imagine that it might be desirable for a planner to have such a bias towards local modifications built into it. Such a heuristic would function to economise on the effort expended by the Solver, given the assumption that making local changes requires less effort than globally rescheduling everything. Of course, there is likely to be an element of real world constraint such that, for example, it would be nearly impossible to affect the progress of some of the other tasks, and thus the only hope is local.

5. Summary of Behavioural Phenomena

The behavioural phenomena which a model must now address can be summarised as the phenomena from the previous development cycle, unless explicitly varied, plus the following from this study:

• The order of tasks is largely controlled by an *a priori* schedule. This schedule is incomplete and actions are ultimately situated.

• Start-up can be heuristically aided, and so quicker.

• There is a tendency to make local corrections or alterations to a plan.

• Scheduling is imperfect - i.e. not conducted objectively.

• A high level schedule may be expressed, and it is therefore assumed constructed, on a relative or an absolute basis.

• There is parallel mental activity (i.e. the Mental Workspace and the Performance System are dissociable).

• Some important information is at least accessed in association with the (mental) execution of a particular task.

• There may be activities of low importance which are not scheduled, but instead are accommodated in spare moments.

• Remaining spare moments are recognised as such (Free Time).
• Monitoring happens both as a consequence of advance intention (which may be unreliable), and in passing.

The following table (4.3.) is included as a synopsis of both the previous set of BPs and those of this study so that the reader can easily see the amount of overlap.

<table>
<thead>
<tr>
<th>Previous BP set</th>
<th>BPs from this study</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Interleaving</td>
<td>✓</td>
</tr>
<tr>
<td>• Planning ahead</td>
<td>✓</td>
</tr>
<tr>
<td>• Limited capacity of the Mental Workspace</td>
<td>✓</td>
</tr>
<tr>
<td>• Limited capacity of the Performance System</td>
<td>✓</td>
</tr>
<tr>
<td>• Parallel mental activity</td>
<td>✓</td>
</tr>
<tr>
<td>• Interruption may be disruptive</td>
<td>✓</td>
</tr>
<tr>
<td>• Partial, incomplete, plans</td>
<td>✓</td>
</tr>
<tr>
<td>• Some tasks done opportunistically, in passing</td>
<td>✓</td>
</tr>
<tr>
<td>• Some information bound to plans</td>
<td>✓</td>
</tr>
<tr>
<td>• Free time</td>
<td>✓</td>
</tr>
<tr>
<td>• Default activities in spare moments</td>
<td>✓</td>
</tr>
<tr>
<td>• Relative or absolute expression of the schedule</td>
<td>✓</td>
</tr>
<tr>
<td>• Initialisation heuristic</td>
<td>✓</td>
</tr>
<tr>
<td>• Local correction tendency (preserving the schedule)</td>
<td>✓</td>
</tr>
<tr>
<td>• Scheduling may be imperfect</td>
<td>✓</td>
</tr>
</tbody>
</table>

✓ BP present, blank cell denotes BP not observed.

Table 4.3. Summary of the BPs observed in this study, together with those from the literature review of Chapter 3.

6. The Model

This section presents the second generation of the model. It starts with a consideration of the analogy in use, describing a new approach in one respect. This is followed by the model itself as first a structure and then the mechanism which the structure supports. The model is then summarized and some of its problems discussed.
6.1. The (dual) foundations of the model

In the previous chapter, the model was described as having two main foundations, planning and execution. These two elements persist in the model to be presented. However, one of these, execution, will be based on an analogy drawn from a different system, since the one described previously has been shown to be inadequate for the purposes. The other element, planning, stays essentially the same.

6.1.1. Execution

In the section covering Vertical-Debugging above (4.2.), the problem of the level of description supported by the data with respect to non-preemption was discussed. The problem is that the data will not provide for messages to be specified at the low level demanded by the observed changes of task. It would be possible to persevere with the current analogy, but the model would become inelegant, and inelegance can mean unnecessary assumptions and side-effects. Consequently, it is necessary to identify a new mechanism upon which to base this part of the model.

A major new operating system was announced by Microsoft (the same company responsible for Windows), at about this stage in the development of the model. This system, known as OS/2, provides a graphical, WIMP, interface on similar technology, with a similar object metaphor for the programmer. The crucial difference is the mechanism by which is shares the unitary processor amongst competing tasks. The published account of OS/2 (Letwin, 1988) describes a preemptive multitasking mechanism built around threads of differing priorities. As much of this as is relevant is described shortly. It is conceivable that OS/2 might have been recruited at an earlier stage had it been available.

A multitasking mechanism based on preemption is one in which the shared processor can be snatched away from a task and given to another without regard to the ongoing task. Simple time-slicing would be one example of a preemptive mechanism. However, the particular preemptive scheme used in OS/2 is more sophisticated, and geared towards ensuring that the processor spends most of its time where it is actually needed (something which is not true in the case of time-slicing).

To achieve this, it introduces the concept of threads. Computer programs consist of many sequences of instructions, typically conditionally linked together. A particular path through such a set of sequences at run time is referred to as a thread. Note that it is possible for there to be more than one such thread associated with any program at a time. It is important to remember that the actual sequence of instructions which makes up a thread is likely to be only fully specifiable post hoc (because of the conditional branches).
To the OS/2 system, then, multiple tasks are represented as multiple threads, where each thread must be given the attention of the processor in order for that task to make progress. This sharing is optimised using a system in which there are different categories of thread, together with priorities associated with each one. There are three main categories of threads, with one category being usefully divided into two subcategories.

Being an interactive system, the thread categories are based on the assumption that the threads associated with the programs currently in use by the user, i.e. in the active screen group, are the most important, and therefore should have the highest priority. The first category is the General category, which in turn consists of the two subcategories a) Foreground and Interactive, and b) Background. The majority of threads fall into one of these two groups. The Foreground category are those threads associated with the currently active screen group, with the Interactive classification referring to the particular thread which is receiving user input. This thread has a higher priority than the other foreground threads, which in turn have higher priorities than the background threads. The latter are associated with tasks which are not part of the user’s active screen group.

The remaining two classes are more specialised. The Time-Critical category is for threads which must be run at particular intervals, for instance, and the Low Priority category is exactly what it says - a category for the threads of the lowest importance.

The allocation (and preemption) of the processor is the responsibility of a scheduler. The scheduler is able to modify the priorities of threads in the general category as necessary (see below), and is responsible for preempting the processor from a thread to ensure that the highest possible priority thread is always running.

In general, the interactive thread will have the highest priority and thus the processor whenever it needs it. After that, other foreground threads will be allocated the processor, accordingly, followed by any background threads. Any low-priority threads will mop-up any leftover processor time if there are no other categories of higher importance.

Time-critical threads periodically raise their priority high enough to preempt the processor, but with the caveat that they must not hold on to it for very long, otherwise the interactive thread will appear disrupted to the user.

Simple priority allocation in this way is not sufficient. Without any extension, the above scheme would mean that once the high priority interactive category thread gained the processor, it would only lose it to a time-critical thread. To overcome this, OS/2 introduces the concept of blocking. A thread blocks when it must wait for,
e.g., a device operation (reading from disk, user input, etc). A blocked thread surrenders the processor, which thus passes to the next unblocked thread with the highest priority. In this respect it is similar to a non-preemptive scheme. As soon as the blocked thread becomes unblocked again, however, the processor is snatched back by the higher priority thread.

While the above is only a limited account of the relevant parts of OS/2, leaving out many details which are not thought to be pertinent, it necessarily covers more than will be recruited to the model for reasons of providing a consistent account. The intention is to use the concept of threads and blocking, with differing priorities, and also to an extent, different categories of threads (in particular the special classes, such as time critical threads).

6.2. Structure

![Diagram of Model's Structure]

Fig. 4.2. Schematic representation of the Model's structure.

From the diagram, it can be seen that much of it remains the same. Both the Problem Representation and the Solution are unchanged. The Solver is drawn in grey, rather than as the strictly black box of the previous model. This is deliberate and is intended to convey the idea that the model now incorporates some notion of of the contents of the Solver in the form of heuristics. At this point at least one such heuristic is proposed - to do with enabling a quick startup. In addition to this, the property of tending to preserve the schedule, i.e. to make local alterations, might be conceived of as a heuristic (see discussion above in Section 4.4.2.).
Centrally, there is now no longer any depiction of message pathways. Although the non-preemptive, message based, multitasking mechanism has been dropped, the object metaphor in general has been retained. There is still a role, then, for messages, but they are no longer central to the allocation mechanism. Instead, there is a box labelled Scheduler. This is drawn in black, implying that its contents are not known, and not necessarily of interest. It is the structure responsible for acting on the basis of the schedule contained within the Solution and preempting/allocating the Mental Workspace as necessary. (The role of the Mental Workspace and the Performance System will be discussed further below.) In the model, the schedule, and the scheduler are depicted separately. Partly, this is due to a suspected misnomer on the part of the designers of OS/2, since it is implied that their scheduler does not allocate priorities, but merely acts upon them. In the present model, the priorities are calculated by the Solution, and implemented by the Scheduler. It is doubtful whether there is anything to be gained in keeping the Scheduler and Solution apart, and whether this will persist in future generations.

On the other side of the diagram, much also remains the same. There are still Task Units, which admittedly still leaves the model open to the criticism of the arbitrary definition of tasks. There are no longer any message queues associated with each Task Unit, for the reason already given, but the major change is the incorporation of threads. In the above description of threads in the computer case, they were described as a sequence of instructions.

A thread in the model is equated with the mental representation of the plan or script associated with the particular task (and therefore Task Unit). In this way it is most closely affiliated to procedural knowledge, but is also related to other, declarative, knowledge in the manner of Byrne (1981) (see Chapter 3, section 3.2.). They are depicted in fig. 4.2. as the wavy lines in the Task Units. The important point is that some Task Units are drawn as containing two such threads. The intention here is to convey the idea of one thread being executed for real, and a separate thread being mentally executed.

Like the previous model, there is no explicit representation of either the Mental Workspace, or the Performance System in the figure (fig 4.2.). The Mental Workspace is equated with the computer’s processor, and is thus the device which is allocated and preempted. In contrast to the previous model, the simplest representation of the Performance System is not on a par with the Mental Workspace, but is rather as the equivalent of a hardware device. This would fit the OS/2 analogy easily in terms of blocking.
6.3. Mechanism

This section will not provide a complete account of the mechanisms of the model, in as much as parts of it are similar to the previous generation and have already been described. The aim will be to concentrate on the points emerging as a result of the changes which have been made.

The mechanics of constructing a schedule are much as before. The invalidation-revision relationship of the Problem Representation/Solver/Solution remains. Solving the problem to create the solution, however, is now potentially supported by one or more heuristics. An example of the latter would be the case of starting up, when it was observed that subjects were able to get on with a task without (or indeed whilst) working out what in fact they should be doing.

Each Task Unit, and thus each task, is represented in terms of a single thread. In the diagram, one of the Task Units is represented as having two threads. The idea of a second thread will be discussed below. The primary thread would be of a general class, and at any moment in time, would have its priority determined by the schedule (i.e. the Solution). The scheduler has only to ensure that the highest priority thread possible has the attention of the individual (i.e. the Mental Workspace). Task switching now follows as a consequence of a change in the solution - which in turn is a consequence of a change in the contents of the Problem Representation.

The second thread is the mechanism of mental execution or thinking ahead. It is represented as a separate thread for two reasons. The first is that this is the manner in which the analogy allows for two separate loci of activity within the same Task Unit - and they do need to be independent. The second reason derives from the idea of a thread being only specifiable post hoc, which is assumed to be equally true for a mentally executed task. In this way, for any task, the two threads are likely to be different. Since the mentally executed task thread has implications for the planning of the task, and its coordination with other tasks, this mismatch may contribute towards inefficiency via an erroneous problem representation. This is just one example of a larger phenomenon - i.e. imperfect scheduling. Imperfect schedules can arise in two ways, firstly by the imperfect operation of the Solver. Previously, this was not thought to be the case; however, now it is modelled as having one or more heuristics at its disposal (with their inherent tradeoffs). The second way in which an imperfect solution can arise, is as a result of an inadequate specification of the problem. Failing to anticipate, or anticipating (thinking ahead) but making the wrong assumptions would be one way in which such an erroneous specification would arise.

Periodic, internally triggered, monitoring is accounted for by invoking a thread which periodically raises its priority. It might be further assumed that it would not
necessarily increment its priority sufficiently to preempt the Mental Workspace. This might be a way of perhaps modelling a person forgetting to monitor when engrossed in something else (i.e. another thread). Other internally and externally triggered interruptions would arise from a change in the problem representation and thus a change in the solution. This would include the case of doing something in passing. In this case, it is conceived that the current location is a factor in the problem representation and a rescheduling takes this into account.

6.4. Summary of the Model

In summary, it can be said that this has been a successful cycle in the development of the model. It was originally thought that the cooking job would not function to extend the scope of the model - i.e. to contribute to its horizontal development. It was clear that the initial description of the job in terms of JAs was unnecessarily limited. The horizontal development of the model has identified three BPs: Initialisation, Default Activities and Free Time. In this light, it is appropriate to consider modifying the set of JAs as follows:

- Interleaving of tasks: in addition to its previous expression, this also needs to consider the workload imposed on the person.

- Although the person may have experience with the individual tasks, they may have no previous experience with the set as a whole.

The observational study reported in the next chapter is particularly interesting with respect to the second point since it is concerned with an example of a multitasking job where the person has a great deal of previous experience with the particular task set.

The vertical development of the model in this cycle has also been significant. The summary of BPs (Section 5) shows the high degree of overlap with the preceding set which has contributed to the consolidation. Probably the most significant change in the model concerns its debugging in respect of task switching. Finally, the data have also allowed some aspects of the model to be unpacked.

6.5. Problems with the Model

The main problem with the model is still the arbitrary division of the set of ongoing behaviours into their associated tasks. This was also true of the model in its previous generation, where it was also expressed in terms of a level of description problem (i.e. with respect to the messages). The latter has been dealt with in this generation, but no solution for the former has been proposed. It is important to understand this problem in its true context, in that it is only a problem if it affects the application of the model. Since this is not imminent, it is not yet an urgent problem.
Chapter 5

5 Study 2: Computer Operators

Prologue

This chapter presents the second observational study, and thus the second full cycle in the evolution of the model. As well as data driven modifications and horizontal development extending its scope, the model in this generation undergoes an important change as a result of a revised conceptual basis.

Contents

1. Introduction
2. The (Abstract) Job Specification
3. The Study
   3.1. The Chosen Behaviour - Computer Operations
   3.2. Method
      3.2.1. Subjects
      3.2.2. Data collection
      3.2.3. Data analysis
4. Results
   4.1. Horizontal
      4.1.1. The Task set is shared with others.
      4.1.2. Previous experience with the task set
   4.2. Vertical - debugging
      4.2.1. Interruption
   4.3. Vertical - consolidation
      4.3.1. There is a Schedule, and tasks are coordinated.
      4.3.2. Monitoring
      4.3.3. Doing things in Passing (While I'm at it).
      4.3.4. The Tendency to Preserve the Existing Schedule.
   4.4. Vertical - unpacking
      4.4.1. Bias in favour of simplicity
      4.4.2. Validating the Problem Representation
5. Summary of Behavioural Phenomena

92
6. The Model
   6.1. Foundations of the new model
   6.2. Structure
   6.3. Mechanism
   6.4. Summary and Discussion

1. Introduction

This chapter is essentially very similar to both the previous chapters and the next one in that like these chapter it constitutes a complete cycle in the development of the model of multitasking.

Unlike the multitasking job observed in Chapter 4, which was, to a large degree, specified by the researcher, this chapter reports observations made of a group of individuals going about their everyday work. The extent of the researcher's control over the observed job is thus limited to choosing a particular period in the day. The job chosen is that of a computer operator; specifically that part of the job concerned with the daily backup of disks to other disks and tapes. This was taken to be a potential instance of multitasking since there were a number of disks to be backed up in parallel, with the added complication of too few drives, for example. It therefore seemed to offer an example of real time resource and task coordination - i.e. multitasking.

The abstract job specification which follows makes explicit the potential for the model to be developed horizontally. This is evident also in the observed phenomena, as are contributions to all three classes of vertical development.

Importantly, the model is revised not only as a result of these observations, but also in the light of a new conceptual understanding of tasks. The model is simplified in the process and re-expressed such that some of its previous structural elements become implicit.

2. The (Abstract) Job Specification

In the beginning, there were three JAs:

• Interleaving, for whatever reason, within a number of constraints.

• The person should have some experience with the individual tasks.

• A certain degree of unpredictability, so for example the possibility of interruption.

The previous chapter ended with a proposal for the following additions:
• An allowance in the requirement for interleaving for workload - the cooking task was not a particularly high workload task and it was thought that consequently, the BP of Free Time was observed.

• In addition to considering the degree of experience with the individual task, the previous experience with the actual set of tasks may be important. In the case of cooking, there was no previous experience with the actual meal to be cooked.

The job of computer operator is such that it is necessary to consider the following two additions:

• The operators *do* have experience with the job itself - i.e. the particular set of tasks under fairly standard conditions. This is the other extreme of second addition above.

• There are other operators present. The task set which they must consider is in practice a subset of the tasks which a number of operators are concerned with as a whole. There was no strict demarcation of responsibilities, but rather the various tasks were dealt with on a more opportunistic basis. It might be expected, then, that this will have an impact on any one person's ability to coordinate their own tasks.

3. The Study

3.1. The Chosen Job - Computer Operations

The operators who were observed, were employed by a computer centre, which provided general central computing facilities for both staff and students of a college within the University of London. The facilities encompassed both micro computers and a mainframe, with its attendant terminals. The operators were responsible for the running of the mainframe, with only limited other responsibilities.

A mainframe differs from the micro computers with which most people are familiar in several ways. Most obviously, it is much larger and typically more powerful. Rather than being dedicated to a single user at a time, it is shared by a large number at once, each one using a separate, remote, terminal in some other part of college. The suite itself, which consists of the computer and the central peripherals such as disk drives and monitors, was housed mainly in one room, with an additional room holding extra printers.

The mainframe was actually made up of 5 computers, one of which was termed the *hub* and formed the central control for the other machines - named the *rims*. Attached to these machines were numerous other devices, such as printers and plotters (these were located centrally and users had to come and collect anything they had printed), communications gateways (to the users and other computer sites in the university),
together with terminals by which the operators themselves controlled the computer(s). These terminals were both of the more modern VDU (Visual Display Unit) type, and also the older teleprinter type. The latter have no screen, but print their output onto a continuous roll of paper. There are at least two advantages of the latter for the operators: more than the last 24 lines is available for their inspection, and any activity is audible. One disadvantage is that teleprinters are slower than the screen based terminals.

Two teams of approximately three operators manned the suite in two shifts per day (it was unmanned late at night). Their jobs consisted of a general duty to keep the machines available for users, looking after the printers and associated output and so on, with a limited responsibility for answering queries from users of a micro facility in an adjoining room. As part of this were several specific jobs, one of which was the periodic backup which was observed here.

Backing-up in general involves copying the contents of one everyday disk to some other format which is taken away to a safe place, such that if some calamity befalls the everyday copy, the data can be recovered from the backup copy. The computer in the present study had a large number of large (6) and small (15) hard disks which needed to be backed up. This was achieved over a week, using a schedule specifying different disks each day. Some backup was to tapes, whilst some was to other disks. The hard disks were all of a removable type, which meant that some could be removed from a drive, and replaced by other disks. There were only a limited number of disk drives, which meant two things. Firstly it meant that there was a degree of competition for and juggling of these resources, and secondly that the computer had to be taken out of normal service and users excluded from using it while backup was taking place.

To minimise the inconvenience to the users, the periodic backup took place between 8.00 am and 9.30 am each day, and thus the operators were under pressure not to exceed the deadline for bringing the computer back on line.

User access was via a gateway called the DCX. The task action in the backup job was to switch this to exclude users and display a suitable message. Similarly, the last task would be to re-enable this link when everything was restored to normal. It was also the practice of most of the operators to halt and restart each of the computers at both the start and the end of the backup session.

During the backup period, one particular operator would deal with the backup itself, whilst any others would take care of less important tasks like opening the micro room for users and checking the paper in the printers. The backup operator would be kept busy for the hour or so that it took (they typically managed to beat the 9.30 deadline),
swapping disks, entering commands on terminals and logging his actions in the appropriate book, and at the same time monitoring the active VDUs and Teletypes for problems. Each backup task required both a source and a destination drive of the appropriate size (termed ‘big’ and ‘small’ in the protocols), a computer (a further constraint was that different drives were connected to different computers, but a computer could run more than one at once), and a terminal. The exact time to copy was unpredictable.

It might be expected from this description that the job of backing-up such a computer is fairly straightforward once a procedure has been learned. The job is sufficiently complicated with unforeseen problems and tasks that it is far from simply a matter of repeating a sequence of tasks from memory. This, coupled with the way it corresponds to the current definition of multitasking, makes it interesting for the current purposes.

3.2. Method

3.2.1. Subjects

The subjects were five computer operators, going about the job of backups as closely as possible to the way they would do it normally. There were two shifts of operators, who took it in turns to do the early shift, and therefore the backups. One of these shifts was smaller than the other, and was used first. For this reason, it was agreed that subsequent subjects would be left to do approximately the same set of tasks (i.e. those associated with the backups) and other shift members would confine themselves to the peripheral tasks, such as cleaning the printers.

The researcher visited the operators at work on several occasions prior to data collection in order to become familiar with the nature of the job. This had the added advantage of allowing the operators to become familiar with an observer being present whilst they worked.

All subjects gave concurrent verbal protocols whilst working. The subjects were all experienced at the job they had to do, with levels of expertise ranging from the shift leaders (several years) to newly trained operators (several months). All subjects were male, and aged between 22 and 35 years.

3.2.2. Data collection

The nature of the working environment was such that video recording was not feasible - the computer room was large and crowded with tall cabinets around which the operators were constantly moving. It was decided that audio recording would be
sufficient given its successful use in the first (cooking) study when coupled with
notes made by the researcher.

To minimise the intrusion, keeping the recording as low-key as possible so that the
other people with whom the subject had to interact could be assumed to be doing so
normally, a small micro-cassette recorder was used. This fitted neatly into the pocket
of the subject, who then wore a small tie-clip microphone. The latter could be
positioned close to the subject's mouth, thus overcoming the problems of the
excessively noisy air conditioning in the computer suite. The tapes in this recorder
were long enough to last for 45 minutes on each side, so they had to be turned over at
some point in the observation. The researcher was careful to do this at a moment
judged to unlikely to interfere with the ongoing tasks. A second cassette recorder
was used by the researcher to record any notes and observations as they occurred
during the job. Finally, the subjects were all de-briefed after each session to clear up
any remaining points that the researcher had not fully grasped.

The protocol instructions were given to each subject in writing before the start of the
session, and any questions were answered as necessary. The instructions themselves
were based on those used in the previous (cooking) study, which were deemed to
have been sufficiently successful, although it was necessary to make some small
modifications (see Appendix B).

3.2.3. Data analysis

All the protocols, and additionally the recorded notes made by the researcher were
first transcribed. These were then broken up into statements associated with
particular phenomena. These are presented using the framework of horizontal and
vertical development and understood in terms of the current generation of the model.
For a fuller description, see Chapter 4. A full example protocol is given in Appendix
B, along with the instructions given to the subjects.

4. Results

This section presents the data in the form of protocol extracts. The extracts are
grouped in terms of the BPs they support, which are grouped further according to the
Horizontal and Vertical classes demanded by the Method of Chapter 2. This chapter
uses the same conventions for the extracts as listed in Table 4.2. (Chapter 4).

In the small sample of operators observed, there were no large individual differences.
They had all worked together for at least a number of months and the more
experienced operators (S1 and S2) had been responsible for training the others. This
might reasonably account for the similarity of style. The subjects do vary a little more
in their level of skill - S4, in particular, is less experienced than the others and this is visible in some of the extracts below.

4.1. Horizontal

Horizontal development is concerned with those novel BPs which are thought to be largely attributable to the novel JAs of this particular instance.

4.1.1. The Task set is shared with others

In this task, it was clear that the particular set of tasks under consideration by the particular operator was a subset of the tasks under consideration by the whole group. The important point is that this subset is not fixed, and it is clear that they have to make allowance in this respect.

The problem could occur in both directions - that is the operator may on the one hand anticipating doing a particular task, but find that someone else has done it (this may mean that the operator has time to spare, but further, had they known about this in advance, their planned schedule may have been different). On the other hand, the operator may have planned a schedule which relies on someone else doing something - in which case they may need to be told to do this.

E.g.:

Subject 2 is obviously aware of the other people and takes them into consideration when planning what he himself should do:

S2179 Erm, right, I'm just trying to think now (what about?) What I've got to do and what has been done. I mean Bob has -

In the case of S4, the subject had stopped and was looking around thoughtfully. In the response to the researcher's prompt, it is clear that he has time on his hands and is looking for something to do. Crucially, he checks what the other members of the group have done, or are doing:

S4 (Right, so what have you got to do?) Right, so all the backups are going now, so there's other things to do, not to do with backups. Is John doing the printers? What I'm going to do now, is every morning we have to clean the tape drives

The following is almost the opposite of the above two cases in that here the subject finds that what he intended to do has already been done by someone else. Not only has this other operator done one of the tasks (i.e. stopped everything), but in so doing he has prevented the subject from doing other tasks (i.e. those which would have been done before the machines were stopped).
Now the first thing that I do is to check the rims just to check that everything is going and I check the hub. Dave has stopped everything already so - I haven't done retries Dave (??) - well we usually do retries but we won't bother with that either. We usually do the messages first as well, but I won't bother with that either.

4.1.2. Previous experience with the task set

In the previous study, the task set (i.e. the tasks to be interleaved) was specifically chosen to be realistic but novel for the subjects. In the current study, protocols were gathered from the individuals as they went about their usual work. As with many jobs, this one involved much that was the same from one occasion to another. Consequently, it was apparent that the subjects had established routines, and that these rarely changed from one time to another, thus the same schedule could be remembered and re-used. That they were nevertheless capable of planning or scheduling a set of tasks is evidenced by the fact that they still had to make adjustments to this schedule in the light of unforeseen events, and also that they were able to verbalise the reasoning behind the learned schedule.

(This learned schedule must have originated in someone’s mind, whether it is passed from trainer to trainee, or assimilated from scratch by each trainee is unclear.)

Examples of a pre-learned schedule:

S271 I've just remembered that I've forgotten to put this in as going back up on here. We've only had this on for the past couple of months and it's still not ground into the system that it's automatically there.

Note the last sentence - this seems to suggest that the subjects are aware that they develop routines for their job.

S1 The first procedure that we do is produce a copy of the disk log.

The way that this is phrased suggests that this is not something that he has decided to do, but rather that producing a copy of the disk log first is accepted practice - i.e. it is routine.

S3 Now the first thing that I do is to check the rims just to check that everything is going and I check the hub. Dave has stopped everything already so - I haven't done retries Dave - well we usually do retries but we won't bother with that either. We usually do the messages first as well, but I won't bother with that either.
This extract has been used elsewhere in this report to illustrate the need for being aware of the other people involved in the job. In the current context, the important feature is the way that the subject implies that what he does is routine (e.g. "usually").

That subjects are nevertheless still capable of thinking and planning for themselves is illustrated below:

| S3 | Now if I'd been on my own, I would have tried to sort it out, but like Dave could carry on and I could do the backups, so if I'd been on my own, I would have carried on with the backups, and kept going over there and hammering away at the DCX |

The above illustrates two things. Firstly that the subject is able to amend his routine - to cope with the fact that ultimately his actions will have to be situated in a possibly novel context, and secondly that the subject can plan theoretically - i.e. he can create an imagined solution from an imagined problem specification.

| S4 | (You can do it with 2 people?) Oh yes, you can, you can do it probably with one if you had to, but it would mean that, er (Is 2 a happy minimum?) Yes, because one would have to do everything, although the backups would take priority |

The researcher prompts the subject in response to an earlier comment about them being understaffed on that particular morning. The quote is intended to support the notion that this subject, like the one above, could react to an unexpected change in circumstances.

The operators are not blindly following a schedule which they have simply learned, rather there is evidence that they understand the reasoning which underlies it. The following quotes are intended as examples of this and thus as support for their potentially more flexible role.

| S262 | Yes, third out of four, so. (are you going to wait for the fourth one until one has finished?) Yes, I'll wait for those other two to finish and then I can run the other one on the hub and theoretically when that's finished that's when the backups are finished because I should have done the small ones and done all that other stuff as well. |

Although this subject seems to be interleaving his activities on a fairly routine basis, he is at least able to verbalise some of the constraints underlying this routine.
... even when there are two of us, the most important thing is to start the backups first thing in the morning, so once we have started the backups, we know they are going to take roughly about 20 minutes (That's your critical thing?) That's right - once you start the large ones, you can start the small ones, and you can leave the small ones running, then you can do the minor bits, clear the logs, clean the tape drives, the printers and PCs etc, open the terminal room at half eight, switching on the printer in there.

This extract is intended to show that the subject is sufficiently aware of the relevant scheduling constraints to be confident of constructing a schedule for a different number of people. A routine needs to be with respect to a set number of participants in this job.

Erm normally we have two people doing the backups, and somebody else does rim B.

Again, there is an interaction between the routine and the number of participants, and the subject is aware of it. Note that this statement implies, correctly, that the ongoing situation is a little abnormal, yet being coped with.

4.2. Vertical - debugging

The debugging component of vertical development is concerned with those BPs which should be addressable by the model, but which are incorrectly predicted.

4.2.1. Interruption

The current model certainly acknowledges interruption, but the nature of the job under observation here is such that it is possible to identify a potential refinement in the associated behavioural phenomena. An interruption, in this case, is the occurrence of some event in the world which demands that the operator do something.

The main observation is that it is possible to discern two stages in an interruption; hence this phenomenon is termed two stage interruption.

The first stage is at least the acknowledgement of the stimulus, and possibly also some further interrogation to discover more of the nature of the particular interruption. This is followed by, and crucially, potentially separated in time from, the actual performance of whatever task the interruption demanded. In the current model, the first stage could be understood as the entering of the necessary information into the problem representation, and thus the revision of the schedule, with the second stage being the result of this new schedule. This handling of
interruption means that the individual may retain more control over his sequence of actions.

In the following example, the operator is engaged in one task when he is interrupted by another operator who reports a problem with a plotter. The important feature is that he gets some details about this and tells the operator to 'hang on'. He finishes what he is doing, and then turns to the plotter.

S2223 so I'll have to online [Jane interrupts with a problem] Is that the same job? (Jane: no, its another one) right, hang on. So I'll online the backup disks as well, so they've come up, so then I just start that, and I don't need the 'INUSE' because it only onlined (The final big backup?) Yes, the final big backup - the two small ones have run successfully - oh I haven't checked this one on B - its come up completed successfully, right. Then, as I don't want B any more, I'll just stop it, and then in a minute I'll just take those disks off and put them back to normal. I'll just have a look at this plotter now, and see what Jane has done to it...

The next extract, below, is offered as a possible example of the first stage alone - just the acknowledgement of the stimulus, plus a certain amount of interrogation. At this point, the operator has just been distracted by the sound of one of the teleprinters and has looked round to see what it says.

S3 That's right. [hears teletype + looks] Its just displaying batch jobs - there are none running, just the ones that are queued.

The second new observation concerning interruption is to do with the recovery or resumption of the interrupted task. It is observed that this recovery can be either rough, in which case the interruption has had a disruptive effect on the ongoing task, or smooth, in which case the operator appears to be able to pick up where he left off as if nothing had happened.

In the next extract, the operator is interrupted by, and talks to another member of the shift, but appears to resume smoothly:

S1 OK, that's fine, so the next [interrupted by and talks to Peter] Peter is going to do the logs... OK, so I come back to here and find that

The alternative rough case, is illustrated in the following two extracts. In the first, the operator is speaking aloud a user message as he enters it into the computer. Then the telephone rings. Telephone interruptions such as this are not very suitable for a two stage handling, since the caller may not be in a position to hold until it suits the interrupted party. In this case the interruption is understood as two stages which are
run together. Consequently, the operator is away from his original task for a considerable period (4-5 mins).

S1 The message is "Please log out, running backups" [the phone rings] Can I answer the phone? (Yes, just do what ever you need to do) [Telephone conversation with the shift leader, explaining why he has not come in today] He's actually been knocked off his bicycle last night ... he's just bruised I think. Erm, where was I? Coming back to this.

The next, and final extract also suggests that recovery can be rough, and like the above, seems associated with a prolonged interruption. Very little can be said in this respect owing to the limited numbers of observations, but several hypotheses are discussed when the mechanics of the new model are presented later.

The operator is interrupted by Jane here, and initially gathers further information from her about the problem. This extends into reasoning about the possible causes and working out a solution (and explaining to the researcher) so that he is actually away from his original task for several minutes. Recovery is then rough, as the last two sentences indicate.

S2 Right, so I'll stick that back - now I still want that write enabled, so I'll leave that as it is. Jane, get some tea! Is that the only job in the queue? (??) And what happened? Well cancel it and change the priority to about 200. (What happened?) Well the plotters started as you can see and it says it is active, but that job - there's nothing happening on it, its stuck, so if she cancels it and changes the priority up a bit, because there's other jobs in the queue, and see if they start running (Are you doing plots before you start the machine again?) Yes, because they'll be - what you mean before we reload? Yes, because they'll be little jobs. When this last backup finishes, I'll avoid the plotter and the printers - so that you only interrupt them, but if there's little jobs, we should be alright. Right, what am I doing now? That's not going on properly - I'd either forgotten or hadn't pressed it on properly

4.3. Vertical - consolidation

This section is concerned with those BPs which are thought to be successfully addressed by the model and thus provide support for it.

4.3.1. There is a Schedule, and tasks are coordinated.

This notion is central to the model. For the purposes of consolidation in this generation, it is presented in terms of three associated behavioural phenomena. Firstly, and most simply, it is possible to find evidence in the verbal protocols that there is advance intention and organisation of activities. Note that the evidence
presented earlier in connection with pre-learned or routine schedules is also relevant here.

For example:

<table>
<thead>
<tr>
<th>S1</th>
<th>We might have to bring up rim D in minute to actually do the backup of rim D on the system, but er with rim C it hasn't got a spare drive, so we have to take it across and do it on the hub.</th>
</tr>
</thead>
<tbody>
<tr>
<td>S3</td>
<td>I can take USE3 off - USE3 backup we can do last, we leave that until right at the end</td>
</tr>
</tbody>
</table>

Secondly, the previous study gave an account which was understood in terms of scheduling activities in the shared context of each other - i.e. where conflicting resource demands and so forth are taken into consideration. There is evidence of this here also:

| S2159 | I don't want to do anything else on this machine so I should - I mean the proper way of doing it is to offline the disk and put them back, but I mean as I'm finished using the machine, all we're going to do after we've done the backups we reload all the machines anyway so there's nothing else I want to do on here, so I just halt it |

Finally, the last study also identified the phenomenon of Free Time, and suggested that its existence implied the coordination of tasks over time. Further, this can be seen as a last resort (given an assumption of diligence), after all other possibilities, including so called default activities have been exhausted. In the present set of observations, there are also several examples of this, some varying slightly within the theme.

Free time as a last resort, after checking all the other possibilities (i.e. validating the problem representation with respect to the world - see later) is exemplified in the following extract. Here, the subject has just mentally run through some of the things which might come up that morning:

| S1  | Right, so just check the notepad - everything is OK there - there's no complications last night. Erm so really we've just got to sit back and wait for that to finish. There's nothing on the spool either - so there's nothing to print out, users can't login |

An assessment of free time, without any such checking is given below:
Lastly, it is worth emphasising that Free Time can be relative. Consider the following excerpt from the protocol of S1. Here he is deciding to stand and do nothing, waiting for the computer to respond (which does take an appreciable length of time), but the statement implies that there are things he could do in the mean time, but maybe the time or effort involved is too great.

S1 I might as well hang about here and wait for the ready prompt to come up.

4.3.2. Monitoring

Monitoring behaviour was one of the phenomena introduced with the prototype model based on the original literature review, particularly from the work of Beishon. It was observed in the case of cooking, and is evident here also. Generally, it is concerned with parts of tasks which are proceeding independently of the operator, in a mainly feedforward manner. Feedforward control is acknowledged to be necessary for the anticipatory scheduling in multitasking (hence the original requirement for some experience with the tasks), but also the world in general is held to be unpredictable. It is only reasonable then that the feedforward control should be augmented by a little feedback control, i.e. monitoring. In terms of the model, monitoring is understood as a case of validating the problem representation against the world.

Originally, monitoring was a particular case of interruption, but since the latter has taken on new refinements since then, it is worth considering monitoring separately for the time being. It retains, however, the potential to be cued internally or externally (with respect to the operator).

Internally prompted monitoring, i.e. that which occurs for no apparent external reason, is exemplified as follows:

S1 I'm just going to have a look to see what's happening to it, get a bit of feedback - we're on logical disk 3 and there's 4 of them to actually be copied so once we're there, we know we're near the end.

S3 This isn't quite finished yet, so just see where its got up to - that should finish soon. How is B getting on? (??) yes, just to see where they've got up to, they're on the last disk, so it won't take long.
Before I run the tape test, I could just check the backups to see how they are going.

The chief occurrence of externally cued monitoring, in terms of the model at least, is that which occurs in passing. This is only one type of such behaviour, which is considered in the following section.

4.3.3. Doing things in Passing (While I’m at it)

Much behaviour can be accounted for in this opportunistic manner, including both monitoring ongoing processes, and the performance of actual tasks themselves.

For example, monitoring:

Ah, this is finished, good (the second big backup) yes, that’s right. (Were you expecting that to finish, or were you just checking as you came past?) They usually finish together, so I knew it would finish any moment - every time I go past, I just have a look.

Check the big ones again - nothing going there.

For example, stopping to do something:

Yes, while I’m here, I could switch the Benson on as well.

_A Benson_ is the name used to refer to the plotters.

Anyhow, I’ll finish bringing up rim B while I’m here.

4.3.4. The Tendency to Preserve the Existing Schedule

This tendency was first identified in the previous study, where it was proposed it could be understood in terms of the model in more than one way. It was suggested that it might be a specific heuristic in the solver which would be actively applied, or alternatively it might emerge ‘passively’ as a natural consequence of a partial planner.

The following excerpt is taken from a point towards the end of the backup session, when the subject is restarting all five computers. This is a complicated process, involving more than just pushing a single button, and he finds that one computer (C) will not come back up after being reset. He performs a minimal check of the fuse, but leaves further examination until later. It is assumed that he is preserving the ongoing schedule, or at least its critical path (note that this job has a critical path with the added constraint of a deadline), since he proceeds to bring up the remaining...
machines. In this way he limits the effect to beating the deadline with four of the machines (which is sufficient), instead of missing it with all five.

This adherence to the critical path and respect for its deadline, is also apparent in another way. In the following excerpt, we are again towards the end of the session, and all that remains on the critical path is to wait for a disk to finish copying and then reset the computers before letting the users back on. The operator has evidently decided to make use of some of this time and free resources to bring forward the printing of the queued items. (These are items from users the previous day.) This is normally something which would be done later, after the restart. Notice, however, that his intention is not to affect the critical path and he anticipates being able to interrupt the printing with little or no cost.

In the event, though, it turns out that the cost of aborting the printing is much higher than he originally expected. Nevertheless, the critical path and deadline prevail:

4.4. Vertical - unpacking

Unpacking is a part of the vertical development of the model which takes consolidation a step further. The BPs in this section are generally addressed by the
model already, but in addition offer some further detail such that elements of the model can be refined - i.e. unpacked.

4.4.1. Bias in favour of simplicity

Sometimes, time and priorities alone are not sufficient to generate an unambiguous schedule. It is possible to observe, in the protocols, a tendency towards solving such dilemmas on the basis of ease or simplicity. This is not really surprising.

For example:

S2 I'm doing it this way because then you leave it back in as near a state to normal running (??) yes, I mean, that's right, it simplifies ...

Leaving the system back in a state which is near to normal minimises the work which has to be done later, which simplifies the overall job.

S3 Yes, so I put it on this drive, which is write enabled from the time before (You could have used this drive?) Yes, but I would have had to open the door, un-write enable that one, and write enable this one, so... This is just easier

There are two possible paths, with the same overall effect, but he consciously chooses the easier one, the one with less work.

S5 I don't know what method S2 does, because I think that the method that S2 does is different from the method that I do - I mean the procedure is that once you have finished running backups you are supposed to reset the machines before bringing it up again, that way you are double checking that the disks are in the right drive, and the system automatically mounts them, but normally the disks that we take off, are one from drive 7 and one from drive 9 - USE2 and USE3 - those are the only two disks that we take off. Now if you can remember that after the backups are finished that all you have to do once the backups are finished is to type in a single command to mount the disks on those drives, that's all you need, therefore you don't have to reload the system, so it just depends on your way of thinking and your way of work. (And you're happier with that one command rather than resetting the whole lot?) Well it only takes a couple of seconds to do that, there's no point in doing that, in reloading the system again - there isn't a necessity. As an operator we shouldn't try to work things by the book so much, but rather use your initiative - you try to learn and not do things in a parrot fashion.

Again, there are two possible routes to the same overall effect, but knowledge of the system allows him to evaluate them against each other, so he chooses the one which is simplest for him, or rather involves less work.
Note that the last part of this extract (about not doing things parrot fashion) is not contrary to the idea of learned routines. The point that the subject is making is that there are procedures laid down by the management, but the routines of the operators need not be identical to these.

4.4.2. Validating the Problem Representation

This has already been mentioned above in several contexts (for example Free Time and Monitoring), but is repeated here in its own right. The model maintains that the person has an internal representation of the relevant aspects of the world, and that this world is changeable, so that it is necessary to actively keep the internal representation up to date.

The following excerpts are offered as examples of this behaviour.

| S2296 | (So what's the current status?) the current status is that we're just - oh, its finished now |

Here, the researcher asks the subject what the current status is. The subject starts to answer as if he is reading out of his current 'picture' (i.e. the Problem Representation), which is that he is waiting for something to finish before proceeding, but looks around to check that this view is in fact up to date. In the event it is not and he changes what he was about to say to reflect this.

Similarly, and more simply, the same subject at another point:

| S2179 | Erm, right, I'm just trying to think now (what about?) What I've got to do and what has been done. |

5. Summary of Behavioural Phenomena

The following is a summary of the particular behavioural phenomena emerging from this study. Some of these phenomena are novel to this study, whereas others are qualifications of phenomena identified in an earlier generation.

- 2 stage interruption (acknowledgement and action).

- Doing things in passing. It is proposed in the light of the observed behaviour that doing things in passing is a more general behavioural phenomenon, rather than being merely one expression of monitoring behaviour.

- Simplicity. For the first time, a bias in favour of simplicity and ease was identified. On reflection, such a bias is not thought to be surprising.
• Sharing a set of tasks. In a situation where the demarcation of task responsibilities is not fixed, it is observed that some coordination and cooperation is necessary between the individuals. Further, this cooperation has been observed.

• Routine Scheduling. It was observed that as a result of repeated experience with the job, much of the organisation of the operator's tasks had become guided by habit.

• Keeping up to date with the world. This is modelled as validating the Problem Representation, and has previously been implicated in the mechanism for several other behavioural phenomena. In this generation, it has been observed in its own right.

The following table (5.1) is included as a synopsis of both the previous combined set of BPs and those of this study.

<table>
<thead>
<tr>
<th>Previous, combined, BP set</th>
<th>BPs from this study</th>
</tr>
</thead>
<tbody>
<tr>
<td>Interleaving</td>
<td>✓</td>
</tr>
<tr>
<td>Planning ahead</td>
<td>✓</td>
</tr>
<tr>
<td>Limited capacity of the Mental Workspace</td>
<td></td>
</tr>
<tr>
<td>Limited capacity of the Performance System</td>
<td></td>
</tr>
<tr>
<td>Parallel Mental Activity</td>
<td></td>
</tr>
<tr>
<td>Interruption may be disruptive</td>
<td></td>
</tr>
<tr>
<td>Partial, incomplete, plans</td>
<td></td>
</tr>
<tr>
<td>Some tasks done opportunistically, in passing</td>
<td></td>
</tr>
<tr>
<td>Some information bound to plans</td>
<td></td>
</tr>
<tr>
<td>Free time</td>
<td>✓</td>
</tr>
<tr>
<td>Default activities in spare moments</td>
<td></td>
</tr>
<tr>
<td>Relative or absolute expression of the schedule</td>
<td></td>
</tr>
<tr>
<td>Initialisation heuristic</td>
<td></td>
</tr>
<tr>
<td>Local correction tendency (preserving the schedule)</td>
<td></td>
</tr>
<tr>
<td>Scheduling may be imperfect</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• 2 Stage interruption</td>
</tr>
<tr>
<td></td>
<td>• 2 Stage interruption</td>
</tr>
<tr>
<td></td>
<td>• Routine</td>
</tr>
<tr>
<td></td>
<td>• Validating PR</td>
</tr>
<tr>
<td></td>
<td>• Sharing set of tasks</td>
</tr>
<tr>
<td></td>
<td>• Simplicity bias</td>
</tr>
</tbody>
</table>

✓ BP present, blank cell denotes BP not observed.

Table 5.1. Summary of the BPs observed in this study, together with the complete set observed so far.
6. The Model

The generation of the model about to be described retains elements of its ancestors, but differs both structurally and more importantly in its conceptual basis. There are two main consequences of the changes: firstly the model becomes simpler, both structurally and mechanistically, and secondly, the revised conceptual basis allows for a clearer definition of what might constitute 'multitasking'.

6.1. Foundations of the new model

The evolution of the model so far has been understood mainly in terms of the analogies employed in respect of the two foundations - planning and execution. This is also broadly the case in the present generation.

The planning aspect of the model stays the same. The three structures (Problem Representation, Solver and Solution) are still present, and employ the same basic invalidation-validation scheme.

The execution aspect, however, does change. The new model will lose its explicit analogy with a multitasking operating system, retaining only certain features. In this way it starts to take on an independent meaning of its own as was suggested to be the case in the discussion of the development of models and theories in Chapter 2.

The new model keeps the notion of threads, but as will become clear, they are different in detail to the ones in the previous model. The main change is in the program-task analogy, which goes, along with its object metaphor (in the OOP sense). The two main reasons contributing to this decision are that:

• it is no longer thought necessary or useful to equate the person's Mental Workspace with the computer's processor.

• the structures of the left hand side of the model, as it was drawn (see Chapter 4, Fig 4.2), make very little contribution to an account of the set of behavioural phenomena so far. It is thought that this may be due in part to the method of observation and data analysis employed, and its suspected weakness with respect to many behavioural phenomena. This method, however, has proved to be largely satisfactory given the other constraints (e.g. ecological observation of behaviour), and will therefore continue to be used. It is not anticipated that its use in the next generation will differ greatly in this respect.

The description of the structure of the model which follows later is therefore limited to the three other structures previously derived from the foundation in planning.
Such a model might appear excessively simple. However, a model is more than just its structure (and mechanism), since it also makes many assumptions about the world and utilises particular concepts to describe that world. The concepts upon which the model has been based so far are the definitions of tasks and resources outlined mainly in Chapter 1, and also to some extent in the limited discussion of Operations Research in Chapter 3. It is now the intention to revise this underlying conception in the light of other theoretical developments in cognitive ergonomics (specifically, HCI). The revised conception draws on that originally proposed by Dowell & Long (1988), and developed further by them subsequently (Long & Dowell, 1989).

Dowell and Long conceive of domains in terms of objects, which may be physical or mental, concrete or abstract. An object would be described in terms of a number of attributes, each of which could assume a number of different states. An object’s attributes may exhibit a potential for a change of state. Such a potential is termed an affordance. The realisation of such an attribute state transformation, or affordance, is accomplished by a task. To give a simple example, in the job observed in the first study, one might conceive of an object fish, which would have a cooked attribute initially in a state signifying raw, but which could be transformed by the task of cooking to state cooked.

An important necessary extension to Dowell and Long (1988) is that in a given task, it may be necessary for objects other than those undergoing the attribute state change to be present. These objects would not themselves be changed in anyway and would therefore be analogous to catalysts in chemistry. It is proposed that the objects that have so far been termed resources can be defined this way - e.g. the saucepans and stove rings in the case of the cooking study, or the disk drives and monitors in the present job.

It is often the case that a particular object will undergo a sequence of transformations (An example might be a potato object and the sequence peel-wash-chop-cook-mash-season). It is proposed that such a sequence of affordances be termed a thread. Such sequences are very similar to those of van Dijk (1980), as discussed in Chapter 1, except that his were expressed in terms of goals etc. Pursuing this similarity, it is clear that if a thread is a sequence of affordances at one level of description, it must itself be synonymous with a task at a higher level. There are two related points in this respect however, which are that a transformation may cause an object to split or merge, and thus tasks/threads also, and that the choice of what will constitute an object is not entirely clear.

It is conceived that within a thread, individual affordances in the sequence need not be specified at the same level of description; for example, the earlier potato preparation
thread could be expressed as *prepare-cook-mash-season*, where *prepare* is covers *peeling*, *washing* and *chopping*. (A thickness metaphor was considered in this respect, but not taken further since it was not immediately clear whether thickness should indicate a high level or low level description of an affordance.)

The current model is restructured such that it avoids its previous dependence on a particular level of description (i.e. the Task Units), but the most important change is that the definition of a task is based around the (sequence) of transformations undergone by an object. This is crucially different from before, where a task was expressed in terms of things done by a person. The same was true of a thread in the previous generation, where it was thought of as a sequence of schemata.

In spite of this, the revised thread can still be considered as retaining some common characteristics:

1. A thread can still be thought of as extending over time.

2. Its ultimate specification must still be post hoc.

Probably the most important new feature though is that the new definition acknowledges that some transformations may be carried out other than by the person under consideration.

In this non-person centred definition of a task, transformations are conceived of as effected by resources, which have been defined above. It is assumed that once a particular object has been allocated the necessary resources for a particular attribute transformation, then this transformation proceeds. Thus there is no concept of an *actor* or *agent*. The act of allocation includes the commission of the particular resource. For example to boil an egg, one would conceive of the egg as an object, the state change as raw to hard-boiled, and the necessary resources as water, pan and heat source. The act of allocation would put these together in the usual manner. Allocation itself, then, is an attribute state change, and must itself require one or more resources to effect the change. This begs the question of what exactly the role of the person is considered to be. The person under scrutiny is thought of as having two roles, as a Controller, and as an Effector.

As a Controller, the person is responsible for allocating resources to affordances. As an Effector, the person thought of as two resources - corresponding to the Mental Workspace and the Performance System. All resources in the system are taken to be equal, and thus the control element of the person’s role is then modelled as a general allocation of resources. The actual allocation, or commission, of other resources in the system is one of the responsibilities of the effector resource(s) of the person.
All of the above implicitly refers to a single task, but it is now easy to see what multitasking might be. Multitasking is being concerned with realising the affordances associated with more than one object at once. 'Being concerned with' comprises both control and effect, which can be mixed to different degrees. Consider the following diagram (Fig. 5.1.), in which different roles are identified within a space.

![Diagram](image)

Fig. 5.1. Representation of the potential roles of the person in a (multitasking) job.

The four indicated extreme cases (A-D) are operationalised, as far as possible, in the following table (Table 5.2):

<table>
<thead>
<tr>
<th></th>
<th>Effect component*</th>
<th>Control component</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>High</td>
<td>Low</td>
<td>Possibly a busy Personal Assistant?</td>
</tr>
<tr>
<td>B</td>
<td>High</td>
<td>High</td>
<td>Cooking, Computer Operations.</td>
</tr>
<tr>
<td>C</td>
<td>Low</td>
<td>Low</td>
<td></td>
</tr>
<tr>
<td>D</td>
<td>Low</td>
<td>High</td>
<td>Air Traffic Control</td>
</tr>
</tbody>
</table>

* the amount of effect the person has over the transformation(s) relative to the effects of the other entities in the system.

Table 5.2. The role of the person in a Job concerning more than one object.

In the case of A and C in the table, it is difficult to imagine a job in which the person would have no autonomy (control component). Such a job would be highly interrupt driven. This aside, it would seem from this breakdown, that most jobs have at least some control component, since the person is most often at least responsible for managing themselves as two resources. Possibly only a very raw apprentice could be thought of as being controlled by someone else and having no autonomy.
To summarise, a conception of objects, resources and tasks has been proposed which allows the definition of multitasking to be decentralised. Previously, multitasking was defined in terms of the actions of a particular person. The person of interest now has two roles. Firstly there is a control role concerned with managing the transformations of states of attributes of objects, and then secondly there is an effect role in which the person is equal in status to all other resources in the system. Separating these two aspects allows a better understanding of the broad range of candidate jobs which fall within the scope of the model.

6.2. Structure

![Diagram of the Model](image)

Fig 5.2. The Structure of the Model

The above diagram (Fig. 5.2.) shows the three main structures in the model. These are largely as before, and maintain the same mechanistic relationship, although what was previously known as the Schedule has been labelled ‘Solution’. There are two reasons for this: firstly the notion of a schedule seems more suitable for a person-centred view of multitasking, which has now been dropped, and secondly the idea of a solution fits that of there being a problem (cf Problem Representation). The specific, detailed, differences are addressed below for each structure. An obvious addition is ‘The World’ on the right hand side, together with the two arrows leading to and from the World and the Model. In this way, the model appears to be a pipeline. To an extent, this is true, but it is intended that the level of expression is far removed from that of simple input and output; the arrows are only very general representations of perception and effect. The model is not one of simple event driven behaviour typical of so many pipeline models.
Taking the three main structures individually:

The Problem Representation is thought of as containing the threads - i.e. the set of affordances to be satisfied. In addition, there must also be some representation of the available resources. The threads contained in the Problem Representation are the anticipated sequence of affordances, which may be different from that which eventually happens (the ability only to specify a thread *post hoc* has been discussed already). Note that this relationship is different from the question of whether there is a break in the temporal sequence of transformations on an object. Such a break reflects the resources allocated to the affordances, and so is a product of the Solution. In this way, the model maintains a separation between what has to be done (in the Problem Representation) and when and how it will be done (in the Solution).

The Solution structure must then specify what resources are to be allocated to what affordance, and when. This can be looked at from two points of view. Firstly as the utilisation of a given resource over time, and secondly as the scheduled performance of the sequence of transformations associated with an object. It is reasonable to expect that such a duality might be observable in the way subjects express their intentions - which are taken to reflect the contents of the Solution. (Actually, subjects ought to be able to express the intended utilisation of any of the resources in the system. This is not entirely unreasonable, on reflection. For example, one might say "I will do X then Y then Z", which would be a statement of the intended use of the self as a resource, or equally, one might say "I'll use this saucepan for X then for Y").

Finally, the Solver. This structure remains unchanged. It is still drawn as a grey box to reflect both the original intention that it should be a ‘black box’ (which was only of interest in terms of its properties), and the later qualification of aspects of its internal workings.

6.3. Mechanism

The much simplified structure of the model has a similarly simple mechanism. This is the basic problem-solution scheme, in which the problem is imperfectly specified, and the solution thus the best estimate at a given moment. A change in the specification of the problem is taken to result in an automatic validation of the solution. This validation notion is extended to cover the relationship between the problem as it exists in the real world, and its internal representation. This is represented in the diagram (Fig. 5.2.) by the lower horizontal arrow. The strength of such a scheme is that many separate phenomena blur together, and can be understood in terms of the same mechanism. (That such an apparent strength may in fact be a weakness is discussed in brief in a section to follow.)
The state of the internal Problem Representation with respect to the problem in the world can be understood as the basis for many of the observed phenomena. For example, interruption, monitoring and sharing the set of tasks with other people. It is even observed that the operators explicitly validate their internal model of the world from time to time, especially when about to decide what to do - it would seem a good idea to update one’s representation of the problem before basing any decisions on it.

In the light of the specific changes to the conception of tasks and resources, interruption warrants a special discussion. Previously, the view of tasks was person centred, and so, consequently, was the view of interruption. The decentralisation of the model results in a corresponding decentralisation of interruption. In the new model, interruption has to be modelled in terms of each of the resources in the system, and thus it is the use of a particular resource which gets interrupted. This can be thought of in general as the use of the particular resource not being as anticipated. Previous observations of interruption are then interruptions of the Mental Workspace and/or Performance System resources. With respect to other resources, it is interesting to note the case reported above (Section 4.3.4), where the operator interrupts a printer.

General mechanistic issues aside, it is necessary to consider how the model addresses the novel behavioural phenomena observed in the present study. The observation of two stages in interruption has already been discussed in the results section, but at first, such an idea appears contrary to the foregoing discussion of decentralisation, although a prime example of the proposed validation scheme. However, when considered in detail, it does not appear too unreasonable. In such two stage interruption, it is considered that there is a potential separation of the acknowledgement of the need to divert a resource, and its subsequent actual diversion to another task. Consider an example of interrupting a printer in order to restart the computer. It is perfectly plausible to consider the first stage to be examination of the current state of the printer, and the second stage to be the termination of the current document. Separation would occur in the case where it was deemed worth waiting for the current document to finish.

It would be easy to misunderstand the intended nature of the Mental Workspace resource in the context of such an interruption for two reasons. Firstly, the person must usually become aware of the requirement to interrupt the use of a resource, and secondly, the person is also likely to have to implement the change of use. In answer to the first point, the model makes an assumption about the Mental Workspace resource, such that at a low level it is not necessarily exclusively focussed, but rather is aware of other things. In respect of the second point, in the model, the act of interrupting itself is a task, and can thus require resources.
The observation of simplicity or ease in deciding planning what to do is modelled simply in terms of a bias or heuristic in the Solver. The action of this structure is bypassed in the case of a learned routine. Observations of behaviour were made in this respect in this study, particularly at the start of the session. Such routine would be produced in the model by the initial specification of the Problem Representation from memory, and a corresponding filling of the Solution from memory. Such a solution would be valid, but the specification of the problem would necessarily be an assumption of normality about the world. Given the complexity of e.g. the operator's world and the nature of the backup process, it is not unreasonable that the greatest influence of routine should be at the start when there is the greatest likelihood of normality.

Doing things in passing is modelled simply in terms of the validation schemes. It is assumed that changing the specification of the Problem Representation is the result of both active (e.g. monitoring) and passive processes. A passive updating of the Problem Representation is merely being aware of the surroundings.

Sharing a set of tasks is a particular, special case of needing to maintain an accurate and up to date representation of the problem. Active validation in this respect, as well as the consequences of not doing so have been observed.

The 'output' of the model, represented as the upper, right pointing, horizontal arrow, is thus allocations of resources to affordances, which is more than the actual behaviour of the individual, even though the individual is likely to be involved in the actual physical allocation itself, as has been discussed.

6.4. Summary

This study has contributed to both the horizontal and vertical development of the model. In Section 2, two additions to the scope of the model were proposed, concerning the experience of the operators with the job, and the presence of other operators. The horizontal development of the model in this respect has proved successful, helping to identify such BPs as validation of the Problem Representation. In addition, this cycle in the development of the model has seen an important change in the underlying conception of tasks. This is reflected, along with other data based influences, in the simplified structure of the model. Many of the observed behavioural phenomena are modelled in terms of the same representation - validation scheme. Unfortunately, there is no way of knowing at this point whether such simplification is a good or bad thing. It may be that its ability to account for many apparently different phenomena results from it being unspecific. This will be partly put to the test in Chapter 7, when an attempt will be made to recruit the knowledge in the model for design. An important consequence of the revised conception of tasks
and the decentralisation of the model is that the definition of multitasking can now be applied to a greater range of jobs than before. The next and final study will take advantage of this.
Study 3: Railway Signalmen

Prologue

This chapter reports the third and final full development cycle. The observed job is that of a railway signalman. There are some additional, novel, behavioural phenomena reported, but the model is not changed. The model is summarised for the purposes of its testing in the following chapter.

Contents

1. Introduction
   1.1. The Chosen Job - Railway Signalman
      1.1.1. The Railway System
      1.1.2. The Station
      1.1.3. The Signalbox / Control Room
      1.1.4. The Station Signalman

2. The (Abstract) Job Specification

3. The Study

4. Results
   4.1. Horizontal
      4.1.1. “Hedging”
      4.1.2. Managing a variable workload
   4.2. Vertical - Debugging
   4.3. Vertical - Consolidation
      4.3.1. There is a Schedule, and resource allocations are coordinated.
      4.3.2. Interruption
      4.3.3. Preserving the schedule (making local changes)
      4.3.4. Simplicity
      4.3.5. Validating the Problem Representation
      4.3.6. Default activities
   4.4. Vertical - Unpacking
      4.4.1. Coordinating with others

5. Summary of Novel Behavioural Phenomena

6. The Model
   6.1. Mechanism
   6.2. Summary and Discussion
1. Introduction

The general layout of this chapter is identical to that of chapters 4 and 5, since it reports the third (and for the present purposes final) full development iteration. The multitasking job in this case is that of a railway signalman. This job is assumed to be unfamiliar to the reader, and is therefore described in detail in a section which follows. It will then be possible to present a partial, abstract description of the job, in terms of its Job Attributes. This will enable the relationship of this job to the scope of the current model to be understood.

After a brief description of the particulars of the study, the results are presented. The organisation of the observations is as before. Finally, a section is devoted to presenting the new generation of the model. This does not differ greatly from the previous generation.

1.1. The Chosen Job - Railway Signalman

The chosen job is that of a railway signalman. However, there is a great deal of variation in the particular responsibilities of different signalmen according to experience and location, ranging from a small signalbox on a suburban branch line, through a shunting yard to the busy terminus station. The present study was conducted in a signalbox of such a busy terminus station, manned by only the highest grade (some of whom had been signalmen for 30 years or more). Some aspects of the job remain constant, whereas others are specific to the particular area covered by the signalbox. The rest of this section considers first the general, constant, aspects of the railway system, and then the particular problems posed by the station in question. Finally, the pertinent details of the particular signalbox are presented.

1.1.1. The Railway System

Railways consist of trains, running on tracks between stations. The number of trains is usually large, whereas the track between any two locations is often severely limited. Consequently, it is necessary for trains to share such track. Safety considerations demand that there be some coordination to such sharing, and it falls to the signalman to implement this according to laid-down regulations.

Track is subdivided into sections, which can range from a fraction of a mile to several miles in length. Sections are guarded by trackside signals to the train driver, in much the same way that traffic lights control access to a road junction. The signals for the sections of track in a particular area are controlled remotely from a central signalbox. A signalman will thus be notified of the position of a particular train, and then set the signals for its intended course such that it either has a clear run or it is required to stop
and wait for a previous train to clear the section of track it needs to enter. It should be appreciated that this account is severely limited, but other relevant points will be made clear in the presentation of the results, as necessary.

In principle, a railway is run to a timetable. This takes into account the enforced sharing of sections of track such that in a perfect world, the signalman's actions would be entirely dictated. However, the world is far from being so amenable, and unreliability is present in most aspects of the system. This unreliability stems from such disparate causes as engine breakdowns and passengers leaving the doors open, and is usually manifest as lateness. This can be further complicated by the need to repair sections of track, and thus at the same time temporarily take them out of service. This means that the job of the signalman entails a great deal of decision making concerning what trains to put where, and in what order.

In making such decisions, the signalman must also take into consideration other factors. Probably the most important of these is the class of the train. Trains fall into 5 classes, and the class dictates the importance and thus the priority of the train. Class 1 trains are the large intercity trains. These are supposed to have priority over all others. Class 2 trains covers the local stopping trains. The remaining classes cover mail trains, goods trains, empty trains being shunted and so on. The signalman is aware of the class of the train from his special timetable, in which each train is identified by a four character code. The first character is number of the class. Subsequent characters identify the type of locomotive (diesel, electric, etc) and its destination. These are apparent in the protocols as e.g. “1M65” etc.

1.1.2. The Station

The particular region of responsibility of the signalmen studied here was a busy terminus station in the north of England. The track and platform layout is depicted in Figure 6.1.
Figure 6.1. Track and Platform plan of the terminus station controlled by the Signalman.

Orientation is the same as the mimic diagram viewed by the signalman.

All lines can be bi-directional except where indicated by arrows. "Up" refers to traffic going in the general direction of London (ie away from the main station). "Down" refers to traffic travelling away from London.
Much detail has been left off the diagram, including the position of the signals. Including such information would have made the diagram over complicated and unreadable, and it is thought that it is probably unnecessary for understanding the verbal protocol excerpts.

The main points to note are as follows:

The station consists of 13 platforms, indicated from 1 to 14 (there is no platform 12). For the most part, the station is a terminus, which means that trains have to depart in the direction opposite to arrival. Practically all the trains using this station were hauled from the front, and it was usually the case that a second engine would be backed onto a train’s carriages whilst it was in the platform, forming a new train ready for departure. Not only does this mean that this second engine has to be manoeuvred into position, but when the new train pulls off, the old locomotive is left ‘on the blocks’ (i.e. at the far end of the track in that bay) and may need to be removed before the next train into that platform arrives. This is only strictly necessary if that engine is required before the platform becomes free again. Some platforms are longer than others, and are thus needed for the longer intercity trains, or can hold more than one of the shorter local trains at a time. Again, the order of stacking of such trains is important.

In addition to the main station, there are two platforms (13 & 14) situated on through lines. There is another station less than a mile beyond these platforms (“Oxford Road”), which feeds all the traffic onto platform 13, and which is the first destination of all the traffic on platform 14. In the event of one of these lines being taken out of use, all traffic must then pass through this one mile bottleneck using the one remaining line. This is not an everyday occurrence. The more common problem associated with these platform becomes clearer when one considers the lines further out of the station.

There are four main lines into the body of the station: fast and slow in both up and down directions. ‘Up’ refers to track used in a primarily London direction, and thus away from this station. ‘Down’ refers to the opposite. These lines are in fact bi-directional (although not at once, and a signalman has to explicitly reverse some points to implement such a change). The two other lines into the station, the Up East and the Down East, deal with traffic to and from the East (often referred to as Ashbury’s by the signalmen). Any such traffic destined for the main station usually stays on the eastern side, using platforms 1 to 5. On the diagram, there are three arrows on portions of the Eastern lines. These indicate that these sections of track are used strictly in that direction only.
One of the most interesting complications for the station signalman concerns trains which wish to cross between the East and platforms 13 and 14 (and thus Liverpool etc). For such a train, the signalman must set up a route which crosses the front of the main station, and therefore prevents any other train from entering or leaving. As will become clear, it is a case of juggling the priorities of the trains with the expectation that they will move through a section fairly quickly.

Other, less major, clashes also occur. For example, whilst there is a train standing on platform 13, it is not possible to enter or leave platform 11 (which is typically used for mail trains); this would be an example of a physical block. The remaining clashes, of which there are far too many to list, result largely from platforms sharing a common entrance - see for example platforms 1 to 3. Only one of these can be accessed at a time.

Finally, just out of the front of the station, there is an engine siding. This is used by the signalman as a temporary holding for locomotives as they are brought out behind their original train (which now has a different engine on the other end), before they are backed onto their next working (also known as the return working, since most of the stock cycles round a particular route).

1.1.3. The Signalbox and Control Room

Most people are familiar with the appearance traditional signalbox which stands beside the track, and houses just the signalmen. Such boxes are often raised and have a large expanse of window, giving the signalmen a good view of what they control. In contrast to this, the signalbox in which the observations were made, was actually a part (albeit the largest part) of a more comprehensive regional control room. This control room was situated in the office block above the station, and had no outside view at all - all control was therefore remote. The control room covered all aspects of a large region of track, including at least two other small stations. The other functions which are relevant to understanding the observational data are the supervisor, who was in charge, and the station announcers. The latter sat right at the back, where they could see the whole of the mimic diagram and could thus see when trains needed to be announced into the stations. Figure 6.2. is a plan of the layout of the control room.
One whole wall was taken up by a large mimic diagram (approximately 12 metres in length). This showed the layout of all the track covered by the signal box, together with switches at signal positions.

Figure 6.2. Plan of the organisation of the station control room.

Wall mounted track diagram, showing state of signals and position of trains

Desktop track diagram with switches at signal positions

Vertically mounted keyboard + telephone screen and additional track display

Supervisor and other staff

Platform announcers
with small coloured lights indicating the position (and setting) of all the signals. The track itself was represented by strips of lights. These strips would be normally unlit, but would light up white to indicate that the signals had been set for that section. The position of the train was indicated by the track light strip turning red as it progressed along a (white) section. The white would extinguish behind the train indicating that once again the track was available for the signalman to set for something else.

The whole region was divided into four sections, each controlled by its own signalman. These were referred to, from left to right, as 'Longsight', 'Station', 'Oxford Road', and 'Bolton'. Each signalman sat at his own console which featured a track diagram with switches at each point there was a signal, as well as a computer screen and keyboard. Each signalman also had a telephone with which he was able to talk to other signalboxes, platform staff, and even train drivers using trackside telephones.

It was sometimes necessary to unset the signals for a section of track - if a train broke down, for example. This is sometimes referred to as 'pulling up', probably after the action used on the switches. The crucial thing to note about this is that there is a timeout period of approximately three minutes before the track is returned to its free state. This is a safety measure, but it means that the decision is often costly to the performance of the whole system and therefore not undertaken lightly.

Finally, the signalbox is manned around the clock on a shift system.

1.1.4. The Station Signalman

The observational data concerns the signalman with responsibility for the station section, whose area of responsibility corresponds to region in the diagram (Fig 6.1.). Like any signalman, his overall brief was to keep trains running safely and to time. A significant aspect of his job involved coordinating the passage of trains with the Longsight and Oxford Road men either side of him. Quite often, for example, a train would be cleared out of a platform in the main station, and immediately given all the necessary signals into the next signal box (i.e. past Longsight).

The amount of work varied a great deal according to the time of day, with the greatest workload in the mornings and early evenings when there could be as many as 60 trains an hour. At such peak times, the station signalman was unique in having an assistant. In the period observed in this study there were approximately 40 trains per hour through the station area.

The computer screen was not greatly used by the station signalman, although part of his (or his assistant's) job involved entering the four character codes for the newly
formed trains. These were then visible in bays on the main mimic board so that all
the signal men could identify individual streaks of red.

2. The (Abstract) Job Specification

This section takes the above specific details of the job of the station signalman and
abstracts some of its properties or attributes. In this case, this can be done in terms of
the existing Job Attributes from previous generations. There will thus be no new Job
Attributes emerging from this study. However, it will be necessary to qualify some
of the existing Job Attributes, and propose some slight variations. Such
modifications are taken to reflect an increased understanding of multitasking, and in
this way, the study is still able to make contributions to the horizontal development of
the model.

Previously, the foremost Job Attribute has been that there should be interleaving of
tasks, for whatever reason. Such an expression is centred on the person, and is
therefore now at odds with the decentralised view of multitasking developed in the
last generation. The equivalent attribute in the new conception concerns the control
and effect roles of the person. However, in order to arrive at an understanding of
this, it is first necessary to specify the objects, resources and transformations of the
world.

Given a conception which allows for abstract as well as concrete objects, many
different specifications are possible. Since there is no known way of deciding upon
one particular set of objects a priori, the most obvious, concrete, objects will be used.
It is proposed then, that the main objects of the system are the trains, and that the
transformations applied to these objects will be locational. The resources will then be
sections of track and platforms, etc. In this way a resource like a section of track
could be allocated to a train so that it could achieve a change in the state of its location
attribute. Similarly, a train would be thought of as being in possession a platform
resource whilst in the station. In terms of the model, the signalman as an effector
also constitutes a pair of resources.

Considering the demands for these particular resources, compared to others, in the
transformations of objects in the world, it is possible to establish the control and
effect roles of the person. On this basis, it is proposed that the job of the station
signalman is one of high control, but low effect, roles (i.e. the attributes decided
upon above are primarily concerned with the location of trains; the signalman as a
resource is not directly involved in transformation of such attributes), and thus differs
from the previously studied jobs, especially cookery. It could be thought that the
computer operators had a similar balance of roles. Both jobs are the same in that the
employment of the person as an effector is chiefly in the allocation of other resources
to tasks; for the signalman, this resource demand is taken to be less. Nevertheless, the signalman is subject to a high workload (note the need for second person in peak periods) and it is thought that the control role is correspondingly higher. The workload will be discussed later, as necessary.

Returning to the other Job Attributes of before, the following is a consideration of their applicability to the station signalman’s job:

• The world is changeable at short notice. In fact, it is suggested that the world controlled by the signalman is extremely unreliable, probably as a result of the sheer number of parameters (for example, points may seize, guards may fail to turn up, trains may break down). As if this were not enough of a problem, the high cost associated with revoking some decisions has already been outlined (i.e. the time lag in setting back a set of signals). These two factors together combine to add an extra novel detail to this particular Job Attribute.

• The signalman has previous experience with the individual tasks. This is true, even given the current conception of tasks. A task would be the transformation of location of a train - i.e. not actually driving it, but setting it the correct signals. All the signalmen studied had been doing this in one form or another for most, if not all their working lives, since only the most experienced were allowed to work in this particular box.

• The signalman has previous experience with the particular set of tasks, i.e. the job. This is true, since the movement of trains is largely dictated by a timetable which remains in force for many months at a time. This Job Attribute might not be applicable on the first few days when a new timetable is introduced, but the current timetable had been in operation for a considerable time before the observation were made. It is also worth remembering in this respect that the effect of experience will be tempered by the extreme unreliability of the system.

• It is necessary to coordinate with other people. In the case of the computer operators, the necessity for coordination arose because of the absence of demarcation of responsibility. On the whole, the particular responsibility of a signalman is clearly laid down, although in practice they will help each other out - for example the Longsight man may enter a train’s number into the computer. In the job of signalman, the cooperation arises from the need for different signalmen to be involved in successive stages of a train’s progress. This is manifest in the case of the station signalman mainly in his interaction with the Longsight man (since they exchange most of the traffic) as agreements on when and where to route particular trains.
3. The Study

As has already been described, the subjects were all signalmen with many years experience, of signalling in general (ranging from around eight to more than forty years), and at least a year in the station control room. In total, five subjects were observed, for an hour each. To provide some consistency, each subject was observed at the same time of day (just after lunch). It was standard practice on this particular shift to rotate around the different sections each day. The number of subjects and access to them was dictated in part by senior management who were not keen for any observations to be conducted during the busy peak periods, for example. Observing during the middle of the day also meant that there was only one person managing the station, which made interpretation easier.

The subjects were originally made familiar with the goals of the observation by a letter sent to the supervisor and passed around for their agreement. Before each session, the signalman was given further detailed instructions (see Appendix C) about giving a concurrent verbal protocol. These instructions were derived from those used in the previous two studies.

The protocols were recorded using the same small micro cassette recorder that was used in the previous study. This was left on the console in front of the signalman. The researcher sat to one side of and slightly behind the signalman and was able to make notes as well as prompt or seek the occasional clarification. Following verbatim transcription of each of the protocols, the analysis proceeded as in the previous two studies.

4. Results

This section is organised in the same way as in the previous two chapters. It commences by reporting the observed BPs thought to be concerned with the horizontal development of the model, and then proceeding to the vertical development. In this section, the excerpts are presented as before, with an abbreviation of the week day being used to identify the particular subject.

4.1. Horizontal

The observations in this section are those which are not adequately addressed by the model, but which are thought to result largely from the particular combination of JAs specific to this instance of multitasking (as distinct from other BPs not adequately addressed, which are dealt with under the heading of debugging).
4.1.1. "Hedging"

This section addresses an observed phenomenon which concerns the way the signalman tends to err on the side of caution in unreliable situations. This will be presented in terms of two aspects: i) the reliability of trains, and ii) the costs associated with being wrong.

i) reliability of trains

Coordinating many different events over time relies on feedforward control - anticipation, rather than feedback. Feedforward behaviour in turn relies on the world being predictable/reliable. As some of the following quotes illustrate, the signalmen judged that some trains (those in motion) were more predictable than others (those at a standstill), and took this into account when scheduling.

"F ... if you pulled off for the train and the train moved, it would make life a thousand times easier, but it doesn’t work like that. There are a lot of variables, there are a lot of things you can’t rely on - you don’t know at any moment whether a passenger is going to decide to jump in at the last minute, or whether someone has forgotten to put red star traffic in, or they’ve had an invalid turn up, or a mechanical fault of some sort - that’s why I say a moving train is a happy train."

"W7 It doesn’t do to pull off right through with a train - it doesn’t do to complete a move, let’s put it that way. ... I never tend to pull off for any move until the train starts moving ... you get let down too many times - and once you have committed yourself you see, its too late to go back (if you set a route up, can you undo it?) Oh, you can do, but it takes a couple of minutes to time a route out, once you have pulled the signals back you’re waiting 2 or 3 minutes for the route to drop out, and you can’t do that until you have informed somebody on the station that you are doing it - you can’t just pull it up willy nilly. (So your best policy is only to commit so far?) That’s what I do. (Different people do it in different ways?) Well, yes. I don’t believe in committing myself, I’ve been let down too many times. (You’ve just set this 1E62 off on 3) Well I’ve pulled off for him because he’s indicated out of the station, plus the fact that there is nothing to go that way, so if he is late, he is not going to stop anything."

"W134 And we’ll pull off number 2, and if he doesn’t go, he’s not going to stop anything"

This signalman is not blindly hedging in all cases, but only where it is warranted. However, later on, this same signalman does in fact get caught out:
Oh well there you are you see! This is what happens. I've gone and pulled off all the way up there, and they tells us now there's no driver - so that's er now blocking platform 13 (You cleared him all the way along to Longsight - have you undone that route now?) Yes, I've pulled the signals up on that route now. So we can't get off 10 and 11 now.

Th as soon as I see him move past 382 signal, which is the one he's stood at right now, I'll set the road for him to come right across

ii) the cost of reversing a decision

Putting back is one of the phrases used by the signalmen to refer to cancelling the allocation of a section of track to a train - i.e. setting the signals to clear. The problem, or cost, arises because of a safety feature built into the system which means that an action of cancelling a signal has a built-in lag of two or three minutes. This is presumably sufficient for any train to clear the section should it have entered. The timeout of two to three minutes is long enough to cause problems for the signalmen - who may see two or more new trains in such a time period. It should be pointed out that, in a very few circumstances, it is possible to set the signals back on a section of track with immediate effects - it depends on the position of the oncoming train in the preceding sections.

Occasionally the signalmen do make either mistakes or simple slips (and then feel annoyed with themselves):

Tu Oh Paul! I've pulled off the wrong way there, so I'll have to wait a few minutes while that dies out.

Fortunately, there are aspects of the station which will accommodate mistakes (i.e. mistaken assumptions of reliability) more than others. It is suggested that signalmen may use this to their advantage:

F [The signalman is discussing the advantages of crossing a train close to the mouth of the station versus further out]
the only trouble with that is, if there's any problem with the trains a mechanical fault or whatever, you've got to wait to timeout across the front of the station, and everything is at a standstill while you are waiting for that to happen whereas going towards 332 and the up slow, and crossing at Ardwick, the only thing you are stopping is any trains behind you - its just a matter of personal preference really.
4.1.2. Managing a variable workload

On the whole, it is reasonable to assume that the presence of the timetable considerably reduces the reliance on the scheduling capabilities of the signalman, and thus ought to be apparent in a reduced workload. In practice however, the signalman's workload is kept at a fairly high level as a result of maximising the throughput of the station. One of the consequences of this is that there is little or no slack available to accommodate the increased demands of abnormal situations.

This subject voices the important role of the written timetable:

<table>
<thead>
<tr>
<th>Th</th>
</tr>
</thead>
<tbody>
<tr>
<td>This is just a reference, for what you've got see (Do you use it a lot?) I use it a hell of a lot - I mean we'd be lost without this, we couldn't do this job without this.</td>
</tr>
</tbody>
</table>

The problems of workload and abnormal situations are exemplified in the following excerpts:

<table>
<thead>
<tr>
<th>W94</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes, it can go on its proper platform now (So that's taken your worry away?) It has, its reduced the workload considerably, because when you get a platform block like that on the junction it can be a pain, especially during the peak hours - these fellas will tell you. If you get one fail or something like that, you are in trouble. (These are your only 2 through lines, aren't they?) That's all. And you've got a train through there every 2 minutes in the next hour or so, and obviously they are going to build up</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>At the moment its OK because everything is going more or less - all the equipment is working. Even at this time of day you're fairly close - you saw with that London - we're fairly close to full capacity, you get a situation, in that case due to a mechanical fault, in which the platforms are occupied, you've got very little alternative basically. I don't quite know what the percentage rate is, but I know at Birmingham New St, the system was running at 130% capacity. We were actually running 30% more trains than the systems was ever designed to cope with - its probably more or less about the same here - we're probably running more trains ... its just a bit too much. I mean I like it, when these trains are going, but there are times when its - there are too many trains and not enough railway basically - there just isn't the slack in the system at peak periods to cope with any unforseen problems. You've got to rely as far as possibly on everything running to time, all the trains - all the train crew turning up, all the units being in perfect working order, and all your signals being in perfect working order - but it doesn't work like that (So if something goes wrong, its not an isolated thing, it backs all the way up) That's right it just snowballs and the job can just grind to a halt</td>
</tr>
</tbody>
</table>
There is a suggestion that subjects try to plan defensively to take account of the potential problems:

W262 You always keep the signals clear, through the junction, you know, 13 and 14, to keep them moving otherwise they build up.

The workload itself can affect the particular actions of the signalman. In the following excerpt, the signalman is explaining how he has a choice of how to complete a move. One option is simpler and less effort. The other is more involved, but perceived as more desirable. His choice is based on how much else he has to do at the time:

F I'm waiting to pull off into 14 until the previous train has actually gone - he then gets a full aspect rather than a subsidiary, call on signal, the call on signal tells the driver that the platform is occupied - if it is occupied, all well and good, but if it isn't, its probably better to wait until the train has gone, to allow him to come in on a full signal, so he knows the platform is clear and he hasn't got to come in at caution. Its just depends - if you're busy, you set the route up and let it get the cat's eye. ... I tend to wait - because the driver knows that he's coming in to an empty platform then

4.2. Vertical - Debugging

There are no observed behavioural phenomena which will be presented as contributing to the vertical development of the model in the way of debugging, in this generation. The possible implications for this will be discussed here, rather than later in a separate discussion section, for clarity, although the intention is to assume that there is no problem.

In order to discuss the potential implications, it is first necessary to briefly reconsider what would constitute debugging, in the present context. Debugging is simply when, for a given situation (i.e. input) the model would produce a behaviour (i.e. output) which is contrary to that which is observed. In a previous generation, observations falling under this heading have included, for example, the granularity of task switching and interruption. The contrast is with observed phenomena about which the model says nothing at all, which thus contribute to horizontal development, observed phenomena which it accounts for completely (Vertical - Consolidation) and incompletely (Vertical - Unpacking).

The most obvious reason for their being nothing to report under this facet of vertical development is that the model is correct and therefore does not need to be debugged. This is the best case, and the intention is to assume this situation here. However, this
is an assumption, and its worth considering it in the context of some of the other potential reasons.

Apart from the model's correctness, the next most obvious reason is that the data analysis might be incomplete. In other words, had the analysis been more comprehensive, some phenomenon might have been observed which would be contrary to the model. This is an inevitable criticism, and one reason why one might benefit from different forms of data. In the present case, the form of the data and the analysis were chosen because to favour non-invasiveness for example.

The other two potential reasons are more concerned with the status of the model, and its predictive capabilities and why the predictions of the model might fail to conflict with anything observed.

Firstly, the model might be too 'general'. This is not meant in the sense of Chapter 2, where generality was defined as the model's scope in terms of Job Attributes, etc. Here it is intended to convey the idea that the model might be too unspecific and thus require additional assumptions before it can generate a suitable prediction. This is similar to the second case, in which the model might be too powerful. Such a model would be one which was itself able to generate too many predictions which might conflict with each other.

It is suggested that none of the above can be conclusively rejected, and it is conceded that there is a danger of any or all of them being true. It is on this basis, and to be able to proceed, that it will be assumed that the model is correct. This, however, is not a serious assumption since the model will soon be evaluated as an engineering device and as a scientific device, and thus ultimately, the model's status can be determined in this way.

4.3. Vertical - Consolidation

The observations which can be interpreted adequately by the model are presented in this section, providing support for the model in its current form. Note that this represents the majority of the observations in this cycle.

4.3.1. There is a Schedule, and resource allocations are coordinated.

The signalman's scheduling behaviour is in the context of the railway timetable. It seems that this provides a useful general guide, even if the unreliability of the system means that the signalman's scheduling has to be more event based:
That’s the only thing really, most of the time I don’t worry about the train until he’s actually hit the panel - you’ve got to plan your station - that’s the only thing you’ve got to watch, is your platforming. You end up trapped otherwise.

The general usefulness of the timetable is illustrated when it no longer applies:

I’ll just have to play it by ear now, I won’t be able to plan anything up, its just going to be one through each way most of the time now - I’ve just got to play it by ear sort of thing, you know. You can’t plan your routes when its like this, because it depends when the trains come and how long they take when they’re in there

The remainder of this section addresses some of the other phenomena associated with scheduling.

a) Free Time

There are no suitable excerpts illustrating free time in the way that it could be found previously. It is not the case, however, that the signalmen concentrate solely on the job (this would probably be impossible): they often chat to each other, for example about football. They seldom take their minds away from the situation for long though. It is probably the high workload and constantly changing world which means that the signalman is kept constantly busy. The shift incorporates a fifth special needs man, to allow any one of the other signalmen to take a short break.

b) Thinking Ahead/Anticipation

There are many opportunities for a small action to have long term effects - for example putting a train on a different platform has an effect on other trains which were supposed to use that platform. It is therefore crucially important for the signalman to be aware of the longer term. The following excerpts are offered as evidence that the signalmen themselves are aware of this.

(You are planning both your platforms and your engine sidings so you don’t get stuck in the far end?) Yes. In that respect you’ve got to think ahead a little bit - and if a train is going to be late, you’ve got to find out how late its going to be before you commit yourself

Now what I’m doing there, is I’m going on what you call the bottom route onto number three because this - oh ... - what I was going to say is that because of going onto 4 we go bottom route then we can always go top route onto 4 as well. (So you’re just saving yourself some track for later?) Yes
c) Balancing resource competitions

It is important for the signalman to consider all competing demands for a resource if he is to stand a chance of successfully multitasking.

Tu 1H10 - now there is nothing sparring there, so he should be able to go straight onto platform 8.

Here he is assessing the competition for this section of track.

W 63 So this is where the messing about starts now - because we’ve got to go both directions through platform 14, but obviously they can’t all go ... Its fitting them in

Sometimes, there are no tasks competing for resources, so allocation becomes much more straightforward:

F (Longsight: Altrincham down slow Alan, and Paddington engine if you want it) - yes, we’ll have him - we’ve got no trains going anywhere, we’ve got no guards for them at the moment

d) Balancing priorities

The priority system is determined for the signalman, but he still has implement it:

M so I’m going to get him down in front of that freightliner, because he’s a passenger train and he’s an express

W 74 You try and keep em as near as possible to time on their proper path, but it doesn’t always work out that way - you’ve got to keep a route open for a class one more so than a class 2

Th I’ll put him on B platform for now, that allows me to still come out of 10 without stopping then, because that on 10 is a more important train

F where you have got a choice of routes, that’s the sort of criteria you use, what other moves is it going to stop you doing, is there anything that you want to do, where if you pull off for one route rather than another, that you won’t be able to do - its an assessment of priorities really, its the same with running a freight train through to er Trafford Park, you’ve got to try and assess whether you’ll knock a minute or two
Implementing the priority system is not straightforward, and for example, it can be complicated by the time of day:

Th: We have a policy here - people coming into town in a morning - there’s a hell of a lot of people coming in than people going out, so we have a policy of get the trains in, and don’t dote on the trains going out. (Opposite in the evenings?) And the opposite in the evenings.

4.3.2. Interruption

There are two aspects to interruption which need to be addressed: the idea of interruption with respect to resources other than the signalman, and the idea of two stage interruption.

Two stage interruption - i.e. suspension of the actual body of the interrupt after the initial stimulus, is certainly present:

In this excerpt, the Longsight man interrupts to handover a train.

M: ... just hang on, let us get this Buxton in. Now, what have you got?

In this case, two interruptions occur almost simultaneously. One is the handover of the engine for number 6, from the Longsight man (Mike). The other is the information about 2H56 from the Oxford Road man on his other side. One of these takes precedence, whilst the other is acknowledged and suspended, to be returned to later.

Tu: [Longsight - down the fast Paul, engine for number 6]. Just one moment Bill - (time passes with nothing said) Style line up slow Mike. Which one did you say Bill, 15.57 - 2H56?

It is also possible to identify cases of interruption with respect to other resources:

F: (Background: Put back on 5) Why, what’s up? (They’re working on the Buffet car). The train on platform 5 is getting carriage and waggon attention now. ... I’m stuck now. - I’m stuck with this one outside anyway - I’ve no chance. (That means that 1H10 can’t come in?) That’s right, yes.

In this case, the signalman has committed a resource to a particular task - i.e. he has set a section of track for the train on platform 5. He is informed that this train no longer requires it, so he has to put back. This really upsets his intentions.
4.3.3. Preserving the schedule (making local changes)

The signalman is somewhat limited in his freedom to make decisions with longer term consequences anyway. However, it is clear that he appreciates the advantages of localised changes to the schedule.

Th No, not really, but if I see a train that's late, I do me best to keep them going all the way like, you know. ... (Is that a general thing - if you know he's late, you help him make up time?) Yes, but you try not to delay another train to keep him going - my policy is that late trains stay late, early trains stay early - you try to keep the trains that are on time, on time, and the ones that are late, you try to push them as much as you can like, without delaying anything else.

M And you've got to keep them going - if you don't keep em going you're making it more difficult for yourself - and you're making the passengers less than happy because the more trains you have standing, and drivers come on the telephone, ringing you up, and you've got the extra mither* of that - you've got to keep on top of it - if you don't keep on top of it, you cause a lot of problems.

* Mither is a word used in Lancashire (at least) equivalent in meaning to 'bother', as in "don't bother me now, I'm busy".

4.3.4. Simplicity

The simplest, easiest, way of dealing with any train would be to give it its entire route when it needed it. The unreliability in the system means that this is not usually attainable:

M The idea on this job is to pull off and then you can forget all about it

Also, the tendency towards ease and simplicity is manifest as a way of choosing between alternatives:

W I prefer to keep it straightforward, where possible. (Longsight: Up the slow please) There's no point - I don't see the point in bending trains everywhere when its not necessary (So you brought it across) You see, some of them will use the alternative routes - when you're stuck you've no alternative have you? But when there's plenty of time, there's no need.

4.3.5. Validating the Problem Representation

The model maintains that this is a central, constantly ongoing process. The two excerpts in this section are intended as examples of special cases. The first one
would be an example of the signalman's problem representation being incorrect, his appreciation that this is not good, and then revalidating his view of the world.

Th  (Background: what's that one on 14?) It's the late running Warrington.
      (Background: Well what's that over there?) Eh? Oh shit! We'll ask the junction to identify it then [two berths show same number] "PHONE can you identify that train coming on 14 for us please - Kirby is it?" Kirby!

The next quote is simply an example of someone pointing out a change in state (i.e. the train on 13 is ready). It would seem that this occurs at an opportunistic moment for the signalman, and so is very similar to doing something in passing.

W 92  (Background: Ready on 13) Oh, that's handy - we're going on 13

4.3.6. Default activities

Previously, default activities have been tasks which are not actively scheduled, but which are done if and when a spare moment arises. The manifestation of default activities in the present job is provided by the priority system:

F  PHONE "alright -I'll shift him directly, or as directly as possibly" That's a relatively low priority move, going into the arrival and departure lines at Ashbury - its only a shunting engine for Longsight - Guide Bridge. (So you're not worried about that) No, I'm not worried about that - I mean we'll get him onto Longsight as and when we can - of all the moves on the panel at the moment, that is probably the lowest priority

4.4. Vertical - Unpacking

The data are such as to be able to suggest additional details for certain aspects of the model. These BPs are reported in this section.

4.4.1. Coordinating with others

In the case of the signalmen, there are two aspects to this. Firstly, there is the simple hand over of trains to and from the Longsight man and also (to a slightly lesser extent) the Oxford Road man:

Tu  Birmingham, Eric, up fast please

F  (Longsight: Altrincham down slow Alan, and Paddington engine if you want it) - yes, we'll have him
The second aspect of coordination concerns the sharing of less strictly demarcated aspects of the job, such as entering train information into the computer:

M  [Longsight: I’ve changed him on 5 for coming out Bill] - thank you very much

There is also the coordination with other people:

F  With dead end platforms like this, to make a move in that sort of situation, the driver has actually got to ask your permission, because there is nothing to stop you dropping a train on the other way, and obviously you don’t want head on collision

The lack of such coordination and sharing of knowledge sometimes has adverse consequences, e.g. had the signalman in the following quote known about this earlier, he might not have wasted his time getting to a state in which he could back up onto 10.

Tu  Fred, I’m just backing up on 10 with this van now. (Fred - they don’t need it now, they’ve put the other one right) Oh!

Finally, the cooperation of the signalmen extends to monitoring each other’s regions. To an extent this is not entirely surprising, given the interdependency, since it may provide for a degree of anticipation:

M  he noticed that - sometimes you’re watching that much that if he sees something coming to a standstill he’ll shout you

5. Summary of Novel Behavioural Phenomena

The novel or qualified behavioural phenomena arising from this study can be summarised as follows:

• Hedging. It was observed that in a situation which is very unreliable, and when the cost associated with failure is high, subjects tended to make the least risky commitment possible.

• Managing a High or Variable Workload. The signalmen have to work hard, and they, along with the rest of the system, are maintained at almost maximum capacity. There is some indication that they plan defensively, to avoid increasing their workload further. There is also an indication that they may make qualitatively different solutions to the same problem in high and low workload conditions.

• Coordinating with others. The previously documented coordination demanded by the sharing of poorly demarcated responsibilities was also observed here. In
addition, the nature of the railway system demands cooperation in the handing over of trains, from one section to the next. It was also observed that the signalmen seemed to share awareness of more than just their section, and cooperated to bring events to the attention of the relevant signalman where they thought it necessary.

As in the previous two chapters, the following table (Table 6.1.) summarises the BPs of this cycle and their relationship to the cumulative set of BPs observed so far.

<table>
<thead>
<tr>
<th>Previous, combined, BP set</th>
<th>BPs from this study</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Interleaving</td>
<td>✓</td>
</tr>
<tr>
<td>• Planning ahead</td>
<td>✓</td>
</tr>
<tr>
<td>• Limited capacity of the Mental Workspace</td>
<td></td>
</tr>
<tr>
<td>• Limited capacity of the Performance System</td>
<td></td>
</tr>
<tr>
<td>• Parallel Mental Activity</td>
<td>✓</td>
</tr>
<tr>
<td>• 2 Stage interruption</td>
<td>✓</td>
</tr>
<tr>
<td>• Partial, incomplete, plans</td>
<td></td>
</tr>
<tr>
<td>• Some tasks done <em>opportunistically</em>, in passing</td>
<td>✓</td>
</tr>
<tr>
<td>• Some information bound to plans</td>
<td>✓</td>
</tr>
<tr>
<td>• Free time</td>
<td>✓</td>
</tr>
<tr>
<td>• Default activities in spare moments</td>
<td></td>
</tr>
<tr>
<td>• Relative or absolute expression of the schedule</td>
<td></td>
</tr>
<tr>
<td>• Initialisation heuristic or learned routine</td>
<td></td>
</tr>
<tr>
<td>• Local correction tendency (preserving the schedule)</td>
<td>✓</td>
</tr>
<tr>
<td>• Scheduling may be imperfect</td>
<td></td>
</tr>
<tr>
<td>• Validating PR</td>
<td>✓</td>
</tr>
<tr>
<td>• Sharing set of tasks</td>
<td>✓</td>
</tr>
<tr>
<td>• Simplicity bias</td>
<td>✓</td>
</tr>
</tbody>
</table>

✓ BP present, blank cell denotes BP not observed.

Table 6.1. Summary of the BPs observed in this study, together with the complete set observed so far.

6. The Model

Accepting the absence of any requirement to debug the model, and given that the model in its current form is able to incorporate the necessary Horizontal development
and Vertical unpacking, there is no real need for the model to be changed. For this reason, this section will be relatively shorter than its equivalent in previous chapters. There will be no need to review the foundations of the model, nor to present its structure. These can be taken to be the same as the previous generation. It is necessary to address the mechanisms accounting for the novel behavioural phenomena summarised in the preceding section. Finally, the model will be summarised and a limited discussion presented, in preparation for its recruitment to design in the next chapter.

6.1. Mechanism

Hedging might usefully be accounted for in the model by assuming that it represents a bias or heuristic in the Solver. Such a tendency would be in line with the more general modelling of planning as a costly and therefore partial exercise. The same is true of the observation that the signalmen may try to plan defensively, in situations of high workload, in order to keep the problem within their control.

The observation that the signalman might choose the simpler or the more involved way of doing something, depending on his workload is less straightforward. Previously, a bias in favour of choosing the simplest or easiest course has been proposed. This must now be qualified such that it takes into account the workload of the person. However, even then it is inadequate, since there is no reason why laziness would not prevail all the time. It is necessary therefore to postulate something along the lines of diligence or professional pride which would act in opposition to laziness. The solver has then yet another factor which it must take into consideration. Another way of conceiving of the same effect can be found in Dowell & Long (1988), upon which much of the model is currently based. For them, performance has two aspects: Task Quality and Resource Costs. Workload would be equivalent to Resource Costs, and so this behaviour could be thought of as a tradeoff between Task Quality and Resource Costs.

Finally, it is necessary to incorporate into the model the observations of with others. Previously, the sharing of a common set of tasks was modelled simply as a special case of validating the Problem Representation. In the present case, this will also suffice given an additional assumption. The demarcation of responsibility here is more defined than was the case for the computer operators. However, the assumption will be that the signalman maintains an awareness or more than just his own region. This is not too unreasonable in the light of some of the comments, since an awareness provides the signalman with a useful means of anticipating some of the future events in his own section. It is this awareness which he maintains for his own purposes which enables cooperation in the form of, for example, bringing potential problems to the attention of a colleague.
The awareness of each other's problems is not necessarily extensible to a similar awareness of each other's solutions. This explains the need to coordinate the exchange of trains between signalmen.

6.2. Summary and Discussion

The model will be briefly summarized, to an extent which will support its recruitment in the next chapter. Following this, the status of the model will be reviewed, sufficient to support the use of the model in this forthcoming chapter.

The model represents the control of multiple tasks as a resource allocation problem over time. The person is conceived as maintaining an internal representation of the problem, which is solved to produce a solution. The actions of the person are then dictated by this solution. The internal representation of the problem is thought to be imperfect and incomplete relative to the problem in the real world. The solution at any moment is also thought to be incomplete. The mechanism of the model is centered around the notion of invalidation and subsequent (re)validation of a) the internal Problem Representation with respect to the Real World (by observation and enquiry), and b) the Solution with respect to the internal Problem Representation. The latter is achieved by the action of a Solver structure. The detailed mechanics of this structure are not of interest, but rather it is modelled as a black box about which certain biases and tendencies (possibly heuristics) are known. Thus it would be drawn as a grey box.

The definition of resources is based on a conception of objects and tasks derived from Dowell & Long 1988 / Long & Dowell 1989). The full potential of such a conception has not been exploited, but its limited recruitment has been useful in clarifying the role of the person in a complex system.

This generation has contributed to both the horizontal and vertical development of the model. The model is now the product of three successive development cycles, and for the present purposes, there is no intention to develop it further. At this point, then, the current state of the model can be assessed with respect to the desired state and subsequently tested.

The state of the model is its scope (its generality), in combination with its standing as both a scientific and an engineering device. At this point, the model has no standing as either a scientific device or an engineering device, rather it has been developed as a descriptive tool. In this respect, its scope should be considered to be reasonably good, given the large amount of consolidatory evidence presented in each generation. At the same time, there has been only a limited need to debug the model (although some of the problems with this have been pointed out already). It might be
mistakenly thought that the model would be stronger had there been more debugging, but this would be to mistake debugging (which is a post hoc exercise) for testing the predictions of the model in a scientific sense (which ought to be done in an a priori fashion).

Its scope has also been enlarged from that of the original prototype, taking on several additional Job Attributes in addition to the core of managing resource allocations over time. These include the presence of other controllers, and the previous experience with the particular job. The goal however, is to (broadly) develop a model with predictive capabilities in science and engineering, and in this respect, the model has the status of ‘not demonstrated’. The demonstration of its capabilities in this respect will be the focus of the next chapter (in which it is recruited to design) and the following, final chapter (in which some of its potentially testable predictions will be outlined).
Prologue

An experimental study aimed at demonstrating the recruitment of the model to design is described in this chapter. This goal is achieved, with a limited degree of success. Other points of interest to the model, arising from the unique opportunity of the study are also discussed.

Contents

1. Introduction
   1.1. Relationship of this chapter to the rest of the thesis
   1.2. The Nature of the Demonstration
   1.3. The Design Problem

2. Method
   2.1. The Multitasking Job to be used
      2.1.1. Job In Abstract
      2.1.2. Instantiation
      2.1.3. Implementation
      2.1.4. The programs
   2.2. Recruitment of the model
      2.2.1. A Summary of the Model
      2.2.2. The Particular Knowledge to be Recruited to solve the current Design Problem
      2.2.3. Expression in design
   2.3. What is efficiency in Multitasking & how to measure it?
   2.4. The Experimental Design
   2.5. The Experimental Procedure
   2.6. Summary of Measures
      2.6.1. Primary Measures
      2.6.2. Secondary Measures
      2.6.3. Tertiary Measures (for interest)
3. Results

3.1. Level of assertions

3.2. Considering the Manipulations individually
   3.2.1. The PR Condition
   3.2.2. The Sol Condition
   3.2.3. The Combined PR + Sol Condition

3.3. The comparative effects of the manipulations
   3.3.1. Primary Measures
   3.3.2. Secondary Measures

3.4. Tertiary Measures

4. Summary of Data Analysis and Conclusions
   4.1. PR Support
   4.2. Solution support

5. Discussion
   5.1. The demonstration of recruitment of knowledge to design
      5.1.1. Where it could have been better
   5.2. Tertiary goal - points of interest

1. Introduction

1.1. Relationship of this chapter to the rest of the thesis

In chapter 2, a developmental strategy was proposed. The purpose of this approach was to develop a model with a degree of generality (of the type specified), with a standing as both an engineering device and a scientific device. In the intervening chapters, the development of such a model of multitasking behaviour has been described. The status of this model should now be evaluated. Its standing as a scientific device will be discussed in the final chapter; this chapter will address its standing as an engineering device, i.e. something which has the ability to support the design process.

In outline, the main goal of this chapter is then a demonstration of the recruitment of the knowledge in the model. However, such a demonstration is also an implicit partial demonstration of the feasibility of the general developmental strategy (i.e. Chapter 2). A more complete demonstration of the latter would entail other considerations such as cost.

Whilst a partial demonstration of the feasibility of the general developmental strategy constitutes a secondary, minor goal, the study also presents an opportunity to learn about aspects of multitasking behaviour not currently explicitly addressed in the model. In this way, the study also addresses some aspects of the model's weaknesses, and although there is no intention to develop the model further in the current context on the basis of
this, any points of interest arising will be discussed from the point of view of developing the model further.

1.2. The Nature of the Demonstration

More precisely, the major goal of the chapter could be expressed as:

“A demonstration of the plausibility of the recruitment of the knowledge embodied in the model for the purposes of design.”

The two important words in the above which are worth considering further at this point are ‘demonstration’ and ‘plausibility’.

Demonstration: The contrast here is with proof, and is the weaker of the two. To establish or prove that it is plausible to recruit is unrealistic. A demonstration is taken to be a case of providing an example of the recruitment of the model with no attendant guarantees of the ability to do so in another context. It is of course believed that such subsequent recruitment will be possible.

Plausibility: There are many ways in which a good design might be achieved - the approach to be described in this chapter is proposed as one of them. Note that there is no offer of a guarantee that the application of the knowledge in the model results in a better design, nor in fact that if a better design is apparent, it resulted from the application of the knowledge. For this reason, the knowledge in the model will be traced through its derivation, application and implementation in a section to follow.

1.3. The Design Problem

The design problem which will be used, i.e. which will be (partially) solved by recruiting knowledge from the model, will be to improve efficiency in a multitasking job. The details of the specific job to be used will be discussed below, as will be the concept of efficiency.

2. Method

2.1. The Multitasking Job to be used

The multitasking job to be used as a basis for this study must satisfy two sets of requirements. It must be a good instance of multitasking with respect to the scope of the model, i.e. it should possess a set of Job Attributes which are addressed by the model, but which are not exactly the same as any of the instances upon which the model is based. In addition, it must be practical for the purposes of an experimental study involving design manipulations, given the available resources.
Specifying, in abstract, a multitasking job as a sub-set of known Job Attributes is not difficult. What is more difficult, is instantiating this in a form that a subject can be expected to relate to (in this case, with the minimum of extra training). Finally, for the purposes of the experiment, the instantiated job was implemented as a simulation rather than the real thing. Instantiating the (chosen) job for real would have been too costly in many ways.

2.1.1. Job In Abstract

The particular job created for this study is that of a carwash manager. The job is relatively simple in its attributes, and does not by any means involve all of the attributes which constitute the scope of the model. In this way, it might be argued that it only provides a limited test of the model’s generality. However, it was pointed out in the original discussion that such generality also encompassed the application of the model to novel permutations of job attributes within its scope. This is certainly the case in the job of carwash manager.

The actual Job Attributes can be stated as follows:

- The job requires the management of the allocation of resources to tasks over time. This represents the control role of the person. In addition, the carwash manager is given a non-trivial effector role in that they must act as a driver. This is arranged such that it is mutually exclusive with the control role.

- The job is not one with which the person has previous experience. Training is provided with respect to the individual tasks.

- There are no other people with which to coordinate.

- The job has no deadline, but does have a clear end point.

- The job is not completely specified in advance - i.e. it was configured such that some tasks would be added later as interruptions.

The main points of the relationship of this new job to those already observed is summarised in the following table (7.1.):
### Previous Job | Differing Attributes
--- | ---
Cooking | • Carwash manager has a lower effect, but higher control role.
• Job of carwash manager (i.e. queue of incoming cars) was not completely specified at the outset.
Computer Operators | • There are not other carwash managers to coordinate with.
• The carwash managers have no previous experience with the job.
Railway Signalmen | • There are no other carwash managers to coordinate with.
• The carwash job has a definite endpoint.

Table 7.1. The main JA differences between the carwash and previously observed jobs

#### 2.1.2. Instantiation

The carwash was conceived as, and presented to the subjects as, a manual service, rather than the probably more common automatic carwash. The carwash is described in full in the appendix in the form of the introductory instruction sheet given to subjects at the start of the experimental session. In short, the carwash was arranged as follows:

**A. The Carwash**

The carwash itself was presented as consisting of three bays, in which cars could be processed. Once a car had been chosen and driven into a bay, it could not be driven out again until it had been fully processed. In addition to the bays, there was a resource store, containing a fixed number of resources of different kinds, a queue of incoming cars, and a large number of imaginary **minions** who would see to whatever was given the necessary resources. This enables the simplistic view of Chapter 5 in which it is thought sufficient to merely allocate the resources to a task.

**B. Cars**

The cars which came into the carwash were all identified by using different real car names (e.g. Escort, Jaguar, Mini, etc). No car name ever appeared twice. The cars were classified on two dimensions; size and level of service required (see later). The sizes were Small, Medium and Large, and cars were given a size which corresponded approximately to their size in the real world (so, for example, the Mini and the Nova
were both Small, the Honda and the Escort were both Medium, and the BMW and the Granada were both Large).

C. Processes and Levels of Service

There were three levels of service offered by the carwash, Standard, Extra and Deluxe. The processes associated with each of these is indicated in the following table (7.2):

Table 7.2. Table of processes in the different levels of service.

<table>
<thead>
<tr>
<th>Service</th>
<th>Wash</th>
<th>Rinse</th>
<th>Ashtrays</th>
<th>Vacuum</th>
<th>Polish</th>
</tr>
</thead>
<tbody>
<tr>
<td>Standard</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Extra</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Deluxe</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>

Details of each of the above processes, in terms of times and resources required, are given in the next table (7.3):

Table 7.3. Times and resource requirements of the various processes.

<table>
<thead>
<tr>
<th>Process</th>
<th>Time</th>
<th>Needs</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>S</td>
<td>M</td>
<td>L</td>
</tr>
<tr>
<td>Wash</td>
<td>20</td>
<td>30</td>
<td>40</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rinse</td>
<td>20</td>
<td>30</td>
<td>40</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ashtrays</td>
<td>10</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>Vacuum</td>
<td>20</td>
<td>30</td>
<td>40</td>
</tr>
<tr>
<td>Polish</td>
<td>20</td>
<td>30</td>
<td>40</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Driving IN</td>
<td>10</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>Driving OUT</td>
<td>10</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes:

1. The times in the table are in ticks (see later).

2. Some processes depend on others having been done beforehand. This was arranged in a manner likely to be thought realistic - so for example it is relatively easy to
understand and remember that Rinsing can’t take place until a Washing has finished, but that Emptying the Ashtrays can take place at any time.

3. Some processes vary in the amount of time they take with the size of the car. Again, this was arranged to be as realistic and easy to remember as possible - so Emptying the Ashtrays takes the same amount of time irrespective of the size of the car, whereas Washing or Rinsing vary. The relationship between time and size was kept deliberately simple. Given a process which varies in time with the size of the car, it takes twice as long to do a large car as a small one, and a medium car falls exactly in between.

Both of the above tables were available to all subjects throughout the experiment (see Appendix D).

D. Resources

The above table (Table 7.3) also shows the resources needed by a process. Care was taken to choose a set of resources, and make the availability of those resources such that the task was realistic but not too complicated. As described earlier, the carwash had a resource store in which all the available resources could be seen, and into which resources were automatically returned when finished with.

E. The Control Role vs The Effector Role of the Carwash Manager

In the model, the multitasking person is modelled as having two roles - as a controller and as an effector. The role of the carwash manager as a controller involved deciding which cars to drive into the available bays, and where to allocate the available resources. Once the command was issued, the implementation of, for example, Washing the Mini with a Bucket and a Sponge happened automatically (it was explained to the subjects that this was taken care of by imaginary minions). The exception to this is the process of Driving cars in and out of the bays. This constitutes the carwash manager’s effector role - the explanation given to subjects is that a recruitment oversight meant that none of the minions is able to drive. Importantly, it means that there are periods of time when the manager must suspend the controlling role.

F. The Carwash Manager’s Goal

The goal of the manager (i.e. the subject) is to process all the cars in the incoming queue in the shortest possible time.

2.1.3. Implementation

The carwash was implemented as a simulation on a computer. In this respect it is a simulation since many of the features of a real carwash were left out (cars, soap, water,
etc). However, the important features or properties of allocating resources to tasks over time were represented. To minimise both the programming effort and subject training, the subject was not required to operate the computer. Rather, they issued commands which were then entered by the experimenter. To maintain consistency, the interaction between the subject and the experimenter was constrained according to the following rules:

1. The subject can ask experimenter for clarification, but not to supply information which is already available in some form.

2. The subject should issue commands to the experimenter, and thus the system, based on the following command language:

   DRIVE <car> INTO <bay>

   DRIVE <car> OUT

   USE <resource> TO <process> <car>

   TAKE AWAY <resource> FROM <process> <car>

In the case of one of the design manipulations, the following two commands were also available:

   QUEUE <resource> TO <process> <car> AFTER <process> <car>

   QUEUE <resource> TO <process> <car> NEXT

3. Only instructions expressed using the allowed command set should be entered. Since the focus of the investigation was on the strategic decisions issued, rather than the subjects' ability to use command languages, this constraint was relaxed such that minor variations were acceptable, so long as the expressed command was still in line with the constraints underlying the derivation of the commands in the command set. These constraints were that the command set:

   - should be as small (and therefore general) set as possible
   - should be as near to commonly spoken language as possible
   - should constrain the user to express one thing at a time (i.e. one allocation or one deallocation etc)
   - should provide all the basic information (i.e. exactly the information required).
Figure 7.1 is a reproduction of the basic carwash screen with the different parts labelled. Appendix D contains a photograph of the actual screen.

The programs were written in a 4th generation data management language (Dataflex, by Data Access Corporation). This language is designed to facilitate the handling of information, allowing the basic carwash system to be developed from scratch in a very short time. One of the trade-offs against this is the flexibility in designing the interface - screen paints, for example, are restricted to characters only.

The programs were run on BBC microcomputers, acting as dumb terminals attached to a Minstrel 4 multiuser computer (kindly loaned by HM Systems PLC), running the TurboDOS operating system. A particular drawback of such a system is the...
unavailability of a real-time clock with accuracy down to seconds. For this reason, the programs were written in such a way as to have an internal tick, which would vary slightly with respect to real time if the subject fired off a large number of commands in a short time, but which would provide complete internal consistency for the rate of acceptance of commands and the relative time taken by different processes. This internal tick count is the basis for the results (to follow, Section 3).

2.2. Recruitment of the model

2.2.1. A Summary of the Model

In the model, there are objects, which have attributes, which in turn have states. The transformation of a state of an attribute of an object is accomplished by a task. Resources are objects in the world which are required in a task, but which themselves are not subject to transformation (are not transformed).

The person whose multitasking behaviour is being modelled is conceived as having two roles; one as a controller, and another as an effector. In the case of the latter, the person is modelled as a pair of resources, which are the Mental Workspace and the Performance System. The person as a controller is represented diagrammatically as a set of three boxes, called the Problem Representation, the Solver, and the Solution.

The problem in multitasking is one of allocating resources over time to competing tasks. The Problem Representation internal to the controller is the current problem as it is seen. This is modelled as being incomplete for various reasons (see elsewhere). Its incompleteness has consequences for the suitability of the decisions based upon it.

The solution in multitasking is, as might be expected, the plan of what to do when. The solution internal to the controller is also thought to be incomplete.

2.2.2. The Particular Knowledge to be Recruited to solve the current Design Problem

Two areas of the model have been chosen, and as a consequence, there will be two corresponding solutions to the design problem. The two areas correspond broadly to the Problem Representation and the Solution respectively. These will be discussed in turn.

A. The Problem Representation

In the model, the internal Problem Representation is always an imperfect representation of the problem as it exists in the Real World. Reasons for this include the constantly changing nature of the problem in the Real World, and the cost associated with keeping up with this in sufficient detail.
Examples of the observed consequences of this would include:

i) Failing to attend to tasks early enough (e.g. it is possible to imagine a strategy which would have been much more efficient if a particular task had been started at an earlier opportunity).

ii) Allocating a resource to the wrong task (from the point of view of the overall efficiency of the strategy). This could be expressed more generally as not being objective when allocating resources.

The goal would be to improve the relationship between the problem as it exists in the Real World, and the problem as it is represented internally by the controller.

B. The Solution

In the case of the Solution - i.e. the current plan of what should be done when, and with what - there are two ways in which efficiency might be expected to suffer.

i) Remembering the plan. The controller must remember what they intended to do, both for the purposes of doing it, and also for the purposes of knowing what else to plan for.

ii) Remembering to implement the plan. The controller must implement the plan, but may be distracted and thus forget or be unable to do so.

The goal in this case would be to find some way of aiding the remembering of a plan, and also of aiding its implementation, such that any plan conceived by the controller, however intrinsically inefficient, could not be made worse.

2.2.3. Expression in design

It should be noted that the knowledge in the model has been recruited in two ways. Firstly, it has been recruited to specify design goals. This however, does not itself directly specify the form of the design manipulation. The design process has been characterised by Whitefield (1990) as having two processes (and being cyclical). The first is a creative, generative step, which is followed by a second, evaluative step. It is in this later process that the second recruitment of the knowledge in the model can occur.

It was thought important for the current purpose that any proposed design manipulation should not change the nature of the job performed by the carwash manager. Therefore, an important constraint in designing something which should increase efficiency in multitasking, adopted here, is that any manipulation should be concerned with the presentation of information already present in some other form, rather than presentation of extra information or prompting with the solution. The principal motivation for this
somewhat arbitrary limitation is that it makes it easier to compare performance with and without the manipulation present.

The final design involved giving the user additional screens (this was the easiest way of implementing it on the available computer system), one each for the two design manipulations.

In the case of supporting the Problem Representation (subsequently referred to as the PR manipulation), the screen (see Fig 7.2.) lists the total set of resources and next to each one showed the car-process combinations still competing for that resource. Note that this screen did not show the current use of a resource, nor the count of resources currently available - it was designed as an adjunct to the basic screen from which this information was available. This screen represents what still had to be done - when it was blank, there was nothing for the subject to do (subjects were in fact encouraged to think of it in this way). Only tasks associated with cars already in bays were shown.

<table>
<thead>
<tr>
<th>Bucket: 2</th>
<th>Wash Escort</th>
<th>Wash Daimler</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cloth: 1</td>
<td>Polish Daimler</td>
<td></td>
</tr>
<tr>
<td>Dustbin: 1</td>
<td>Ashtrays Escort</td>
<td>Ashtrays Daimler</td>
</tr>
<tr>
<td>Hose pipe: 2</td>
<td>Rinse Escort</td>
<td>Rinse Daimler</td>
</tr>
<tr>
<td>Sponge: 2</td>
<td>Wash Escort</td>
<td>Wash Daimler</td>
</tr>
<tr>
<td>Vacuum Cleaner: 1</td>
<td>Vacuum Escort</td>
<td>Vacuum Daimler</td>
</tr>
<tr>
<td>Wax: 1</td>
<td>Polish Daimler</td>
<td></td>
</tr>
<tr>
<td>You: 1</td>
<td>Drive Out Mini</td>
<td>Drive Out Escort</td>
</tr>
</tbody>
</table>

List of all resources in the system: Quantity of this resource in the system: List of tasks which require this resource

Fig 7.2. Reproduction of the screen used to support the Problem Representation.

The other manipulation (supporting the Solution, subsequently referred to as the Sol manipulation) involves providing an additional mechanism whereby subjects could input their intended commands ahead of time. Such commands are by definition ones which could not be implemented immediately and thus corresponded to allocations of resources currently in use. An extra command was added (see above) which allowed the subject to
specify a process to be done by the first instance of a particular resource to come free - in other words they could Queue resource allocations. The additional screen showed any currently queued allocations for each resource (thus supporting the remembering of the current solution), and implementation of that allocation was automatic (thus obviating the need to remember to do it). This screen is reproduced in Fig 7.3.

<table>
<thead>
<tr>
<th>Resource Name</th>
<th>Next use</th>
<th>Sunsequent use</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Task</td>
<td>Car</td>
</tr>
<tr>
<td>Bucket</td>
<td>Task</td>
<td>Car</td>
</tr>
<tr>
<td>Cloth</td>
<td>1</td>
<td>Polish</td>
</tr>
<tr>
<td>Dustbin</td>
<td>1</td>
<td>Daimler</td>
</tr>
<tr>
<td>Hose pipe</td>
<td>2</td>
<td>Polish</td>
</tr>
<tr>
<td>Sponge</td>
<td>2</td>
<td>Daimler</td>
</tr>
<tr>
<td>Vacuum Cleaner</td>
<td>1</td>
<td>Vacuum</td>
</tr>
<tr>
<td>Wax</td>
<td>1</td>
<td>Escort</td>
</tr>
<tr>
<td>You</td>
<td>1</td>
<td></td>
</tr>
</tbody>
</table>

List of all the resources in the system
Number of resource of this type
Queue of future allocations of this resource

Fig 7.3. Reproduction of the screen used to support the Solution.

2.3. What is efficiency in Multitasking & how to measure it?

The efficiency of performance in a given situation must relate to the goal being pursued in that situation. In the current case, the goal of the carwash manager has already been stated - to complete the processing of the set of cars in the shortest possible time.

Alternative goals might have included maximising the utilization of particular resources, or processing individual cars in as short a time as possible.

Efficiency in the current case could be measured simply as time taken. However, there are certain complications. The total time (i.e. real time) to complete a job would not be a very accurate measure since, as has been pointed out above, the machine on which the simulation was to be run was not designed to support this. Instead, it is possible to utilize the fact that the system was written around an internal 'tick' and use this tick count as a measure of the total strategy time for the completion of a particular job. The term 'strategy time' is used to refer to the number of ticks taken to implement the plan adopted by the subject. This is to distinguish it from the real job time, which would vary according to the load on the computer.

Notice that this only yields a relative measure of efficiency. In theory, such a measure could be related to a theoretical strategy time by using a computer to solve the resource
allocation problem. Such programs exist in the form of Project Management tools, but none of sufficient power and flexibility were available (which, incidentally says something about the complexity of resource allocation problems which people have to solve when multitasking).

Finally, there ought also to be some way of quantifying the effort expended by the person doing the multitasking to achieve the above strategy time. Measuring the effort required is not straightforward. It could be done in terms of the number of commands issued - but the design of the particular job used in this study means that there is unlikely to be much variation in this (there is a finite number of things to do, and therefore a fairly fixed number of decisions to make).

Instead, it was decided to utilise the concept of free time as a measure of effort. It will be remembered that in the model, free time is that time when the person, both as a controller and an effector, is not required. The assumption is that the more free time which is observable, the less effort the person is expending.

To summarize:

Efficiency in multitasking will be taken to be the completion of the prescribed job in as short a time as possible with the least effort, as measured in terms of free time.

2.4. The Experimental Design

The above discussion has specified that there are to be two manipulations - PR & Sol. The hypothesis under test is that the presence of either design manipulation alone should have an observable effect on the efficiency with which the multitasking job is accomplished. The effect of the presence of both together is unclear a priori (one would expect that they should be additive, but there are other considerations which would suggest that this might not be visible in the data).

The experimental design must be such that the effects of such manipulations can be seen both separately and in combination. Any effects are likely to be masked to a degree by individual differences and learning effects - these must therefore be controlled for.

These requirements can be satisfied by using 4 experimental conditions and 4 trials per condition. The following table (7.4) indicates the presence of each of the 3 screens across trials and groups.
Table 7.4. The experimental design, showing screens used in the different conditions on trials 1 to 4.

Points to note about the above:

- The Base condition, in which all 4 trials are identical (see below) is included to reveal any learning effects. The results section will discuss each of the other 3 conditions with respect to the Base (control) condition. This is therefore a between subjects, rather than within subjects design.

- All conditions are identical on trials 1 and 4. The purpose of the similarity on trial 1 is to provide a basis for controlling for individual differences for trials 2 and 3. Trial 4 returns to the base condition in case there is any after effect of a manipulation, or overall learning.

- Two experimental trials are used (trials 2 and 3) rather than one. This is intended to allow for any effect of learning within a manipulation.

To further enhance comparability, the job of the carwash manager was kept as constant as possible across all trials and all conditions. Given that the system was interactive and responded in a realistic way to the commands of the subject, it is not possible to ensure that the parameters of the job were identical for all subjects in all conditions.

A job consisted of processing 6 cars. 4 cars were present in the incoming car queue at the start of a trial, with the other arriving at 150 and 200 ticks. The sizes and levels of service associated with each of these cars are shown in Table 7.5.:
Table 7.5. The configuration of a job in terms of cars.

The fact that the subject was doing the same job four times was disguised by changing the names of all the cars on each trial - for example, Trial 1 involved ESCORT, HONDA, DAIMLER, UNO, ROBIN and SIERRA, whilst Trial 2 used NOVA, ORION, GRANADA, PORSCHE, VOLVO and YUGO. All subjects were asked at the end of the experimental session whether they had realised that they were in fact doing the same job four times. The subterfuge was successful.

The other constant in the system was the number of available resources. The resources in the system were:

BUCKET, CLOTH, DUSTBIN, HOSE PIPE, SPONGE, VACUUM CLEANER, WAX, YOU.

There was one each of all resources except BUCKET, SPONGE and HOSEPIPE (i.e. two separate instances of each of WASHING and RINSING could be going on at once).

4.4.1. Subjects.

A total of 35 subjects were observed. Three of these could not be used because their data had been incompletely logged, leaving a total of 32, or eight in each group. Subjects were assigned to conditions according to a predetermined semi-random pattern (which allowed for preliminary analysis after 16 and 24 subjects). The subjects used were all adults, of both sexes, aged between 19 and 55 years. Recruitment was through notices placed around the college and all subjects were at least at an undergraduate level of education. The only stipulation in recruitment was that the first language should be English (this was because of the heavy use of English language text in both the instructions and the screens).
2.5. The Experimental Procedure

The general scheme for an experimental session is laid out below:

1. Subject reads the general instructions (see Appendix D)

2. Subject reads the detailed instructions (see Appendix D)

3. Experimenter answers any questions and describes the various parts of the base screen.

4. Training trial. This uses just 3 cars, which are only used for this purpose. The experimenter starts the trial off, demonstrating the use and effects of the available commands, and then the subject is permitted to finish off the set of cars.

5. Trial 1.

6. (For PR, Sol and PR+Sol groups only). Written explanation and demonstration of the extra screen(s). The additional features of the system were introduced to the subjects in the form of an extra instruction sheet (see Appendix D) and a brief demonstration by the experimenter.

7. Trials 2, 3 & 4.

8. Debriefing, including CFQ questionnaire.

An experimental session took approximately one and a half hours (each trial lasted in the region of 15 minutes). Subjects were paid the sum of £5 for participation.

Figure 7.4 is a plan view of the experimental setup, showing the positioning of the experimenter’s and subject’s screens - such that only the experimenter could see all four.
2.6. Summary of Measures

Three categories of potential measures have been identified according to their relationship with the hypothesis under test. The two members of the first, primary, category have been described above. Other measures to be taken during the course of the experiment are now also summarized.

2.6.1. Primary Measures

These are the measures which are directly related to the hypothesis being tested.

A. Total Strategy Time

After a small initial pilot study, it was decided to consider the strategy time to start when the first command is given (i.e. the first car is driven into a bay), and to end when the last car was driven out. The reason for this is based on using the internal tick clock and the fact that it tends to exaggerate the initial idle period.
B. Amount of Free Time

Free time would be when there is nothing for the person to do. Calculating precisely this would require more effort than the experiment warranted. Instead, it was decided to define free time, for the purposes of measuring it, as having less than or equal to a criterion number of things to do. This could be determined quite easily without modifying the experimental programs, since this is essentially what is displayed anyway by the PR screen. The criterion was established at three items, since it was observed that it was quite common for there to be nothing to do except wait to drive out 3 cars when they became ready.

Rather than use this raw figure, it was decided to express this as a proportion of the individual's total strategy time on that trial.

2.6.2. Secondary Measures

The secondary measures are those which, whilst not directly related to the hypothesis, are included for their potential to provide extra information useful for understanding behaviour.

A. Self Rated Efficiency & Difficulty

After each trial, subjects were invited to rate how efficient they thought they had been on the last set of cars, on a scale from 0 to 10, and also how difficult they had found it, similarly on a scale of 0 to 10. Examples of these rating slips are given in Appendix D.

B. Comments

Also on these slips, there was room for the subjects to make any comments or observations they wanted to. Whether or not to write anything was left to the subject. It should be emphasised that no verbal protocols were collected in this study.

C. Direction of Gaze (DoG)

If one accepts the assumption that time spent looking at an information source is correlated with the interest in the information provided there, and given the manipulations in the current experiment involve additional information sources in different spatial locations, it seems reasonable to record some measure of where the person is looking. It was decided to use a simple form of such a measure and record the proportion of time spent looking in various directions. This was accomplished using a video camera positioned in front of, and level with, the subject's head. It was easy to determine from such a recording at which screen (or desk, etc) the subject was looking.
There are of course reasons why one might not want to accept the above assumption, but it will be accepted for the current purposes - especially since DoG has only been given the status of a secondary measure.

D. Debriefing

At the end of the experiment, all subjects were asked generally how they had found it and whether any manipulations they had experienced had been useful. In addition, those who had not had the Sol condition, had it described to them and were asked whether they thought such a facility would be useful. The same was not done for the PR condition since it was found to be difficult to describe it to them without an excessive amount of background information.

E. Experimenter Observations

The experimenter, although primarily engaged in entering the subject’s commands into the computer nevertheless had plenty of time to note down any interesting observations. In the diagram of the experimental setup shown earlier, it can be seen that the subject faces slightly away from the experimenter, meaning that whilst the subject cannot see the experimenter’s screen, the experimenter has a full view of the subject’s screens and is thus completely aware of the current state of the carwash. In addition, directly facing the experimenter, albeit slightly obscured by his terminal, is a video monitor showing the subject’s face.

2.6.3. Tertiary Measures (for interest)

There is only one tertiary measure, and it is classified as such since it is not at all contributory to testing the stated hypotheses of the investigation. This measure is concerned with individual differences and is included for interest. It is largely taking advantage of the design of the study. (The experimental design, since it utilises the same first trial for all subjects, is able to provide data on individual differences at no extra cost.)

The Cognitive Failures Questionnaire (CFQ) was originally proposed by Broadbent et al (1982), and has subsequently been used by many others (e.g. Reason 1988). It is a set of 25 questions, each answered on a scale of zero to four, concerned with the frequency of cognitive failures. The questionnaire is reproduced in Appendix D, and the reader interested in its content is directed there.

It has been proposed as a measure of ability to do complex tasks, since it correlates well with performance on these, but not at all with performance on traditional simple laboratory tasks. It was thought that multitasking as defined here probably constituted an
example of such complex behaviour, and if successful, the CFQ could serve in the future as an a priori indication of multitasking ability. Subjects were given the questionnaire to fill in as the final part of the experimental session.

3. Results

3.1. Level of assertions

The study was designed to allow the gathering of quantitative data. These data have been analysed and the hypotheses of interest subjected to statistical test. In many cases, it was not possible to reject the null hypothesis at the 0.05 level typically accepted in Experimental Psychology (Cognitive Ergonomics had no convention of its own and this seemed the most appropriate to use).

Multitasking behaviour is very complex. Even in a relatively simple and controlled job such as that of a carwash manager used here, there is a great deal of scope for different strategies and tactics. The variability in performance is apparent in the data. Fig 7.5 shows, for all subjects, the initial idle period (until the first car is driven in), and also the unadjusted time of driving the last car out (see Section 2.6.1.A). It can be seen from this that there is a wide range of times taken. The graph also serves another purpose. The data points are arranged according to the four conditions, and it is apparent that, broadly, the pattern of times does not differ between groups. It is thus reasonable in the rest of this chapter to compare performance across conditions. It is interesting to note that two of the three longest times for driving the last car out correspond to unusually long initial idle periods, giving support to the decision to utilise the difference between the two. The dotted line on the graph is the average of this adjusted strategy time, and it can be seen that there is very little difference between groups.
Fig. 7.5. Graph of the unadjusted time data for all groups.

Even given these adjustments, there is still variability, and it is believed that this is the main reason for the inability to support statistically many of the trends which, it will be argued, appear to be present. The level of assertion associated with the following results is therefore necessarily low. However, there are nevertheless visible trends in many of the primary measures, and it is the intention to reinforce these by considering them in combination with the secondary measures, some of which are quantitative, others of which are qualitative.

The statistical tests performed were non-parametric, largely because the fewer assumptions underlying these tests were more compatible with the above manipulations. Unless stated otherwise, tests for a difference in two populations were Mann-Whitney U tests ($n_1 = n_2 = 8$, $U$ must be $\leq 15$ to reject $H_0$ with $p < 0.05$).

Finally, to partially overcome the problem of individual differences, it was decided to adjust an individual’s performance measures to be relative to their performance on the first trial (it will be remembered from the experimental design that this is identical for all subjects in all groups). In the following figures, which show an average of this adjusted
measure, within a condition, first trial performance is therefore always represented as 1. Performance on subsequent trials is relative to this, so if, for example, the average relative improvement in time (i.e. the time taken was less) on trial 2 was 25%, then this would be represented as 0.75.

3.2. Considering the Manipulations individually

The main question to be addressed in this section concerns whether the presence of either manipulation, or both together, had any noticeable effect. The relative effects of these manipulations will be considered later.

Each of the three experimental conditions will be considered in turn, relative to the base, control condition, starting with the PR condition, followed by the Sol condition and the combined PR + Sol condition. Each condition is presented in terms of the primary data (Strategy Time and Free Time), then the secondary data (Perceived Efficiency and Difficulty, and Direction of Gaze), and then also any other comments or observations made by the experimenter. However, to simplify this data presentation, when considering Direction of Gaze, only the information pertaining to the screens of interest will be given. This, of course, can only be appreciated in the context of the direction of gaze for the remaining time. This information is presented once, now, rather than repeated for each case below.

The video recordings of subjects were analysed for gaze in five directions. These are the MAIN screen, the PR screen, the SOL screen, the DESK, and OTHER (which includes staring the the ceiling, walls and floor etc). Fig. 7.6 shows the average proportion of time spent looking in these directions for the four conditions on trial 1. It can be seen that the greatest proportion of time is spent looking in the direction of the main screen (average 88%), with smaller proportions accounted for by the desk and ‘other’ (9% and 3% respectively). Two points in particular need to be made:

- In the conditions where the PR and Sol screens are present, any time spent looking in the direction of these screens is accounted for by a general decrease in time spent gazing in the other directions.

- The pattern of time spent looking in the three basic directions remains the same on trial 4, with no perceivable differences between conditions. Thus there is no indication of any after effect on this trial - for example looking in the direction of a screen no longer present. There is a slight shift in balance across all groups so that slightly less time is spent looking in the direction of the main screen (average 86%) and at the desk (7%) and more time spent looking elsewhere (6%). These differences are only slight.
Fig. 7.6. Histogram showing the average proportion of time spent gazing in five directions, for all four conditions, trial 1.
3.2.1. The PR Condition

A. Primary Measures

i) Strategy Time

<table>
<thead>
<tr>
<th>Condition</th>
<th>Trial 2</th>
<th>Trial 3</th>
<th>Trial 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Base</td>
<td>0.95 (0.09)</td>
<td>0.92 (0.05)</td>
<td>0.93 (0.09)</td>
</tr>
<tr>
<td>PR</td>
<td>0.92 (0.07)</td>
<td>0.90 (0.09)</td>
<td>0.94 (0.09)</td>
</tr>
</tbody>
</table>

U (* = significant) 32 30 31

Table 7.6. Strategy time data for Base and PR conditions.

![Strategy Time Graph](image)

Fig. 7.7. Graph showing the average relative strategy time over four trials for the Base and PR conditions.

Although there are no statistically supported differences in the above, there is a slight visible trend in the direction which would be expected if the PR manipulation were having its predicted effect on efficiency measured in this way, i.e. performance on trials 2 & 3 is relatively better for the PR group, compared with the Base, control, group. The apparently identical data for trial 4 adds weight to the notion that there may be such an effect, albeit a small one.
ii) Free Time

<table>
<thead>
<tr>
<th>Condition</th>
<th>Average (S.D.)</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Trial 2</td>
<td>Trial 3</td>
<td>Trial 4</td>
</tr>
<tr>
<td>Base</td>
<td>1.10 (1.09)</td>
<td>1.10 (1.19)</td>
<td>1.79 (1.89)</td>
</tr>
<tr>
<td>PR</td>
<td>2.78 (1.28)</td>
<td>3.23 (2.45)</td>
<td>2.17 (2.48)</td>
</tr>
<tr>
<td>U(* = significant)</td>
<td>6*</td>
<td>11*</td>
<td>29</td>
</tr>
</tbody>
</table>

Table 7.7. Free time data for Base and PR conditions.

Fig. 7.8. Graph showing average relative free time (as a proportion of strategy time), over four trials for the Base and PR conditions.

Here, the data is statistically more supportive. Performance, on this measure, is better when the manipulation is present, shown as a higher proportion of free time, returning to base level on trial 4 when it is removed.
B. Secondary Measures

i) Perceived Efficiency

<table>
<thead>
<tr>
<th>Condition</th>
<th>Trial 2</th>
<th>Trial 3</th>
<th>Trial 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Base</td>
<td>1.15 (0.21)</td>
<td>1.22 (0.34)</td>
<td>1.32 (0.42)</td>
</tr>
<tr>
<td>PR</td>
<td>1.02 (0.05)</td>
<td>0.94 (0.43)</td>
<td>0.98 (0.35)</td>
</tr>
<tr>
<td>U (* = significant)</td>
<td>21</td>
<td>15*</td>
<td>13*</td>
</tr>
</tbody>
</table>

Table 7.8. Perceived efficiency data for Base and PR conditions.

![Self Rated Efficiency Graph](image)

Fig. 7.9. Graph showing average relative perceived efficiency, over four trials for the Base and PR conditions.

From the graph, it appears that for the control group, perceived efficiency increases steadily over the four trials. This would appear not to be so for the experimental group (i.e. with the PR manipulation). In the case of trials 3 & 4, the differences between the two groups are statistically significant.
ii) **Perceived Difficulty**

<table>
<thead>
<tr>
<th>Condition</th>
<th>Average (S.D.)</th>
<th>Trial 2</th>
<th>Trial 3</th>
<th>Trial 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Base</td>
<td>0.93 (0.22)</td>
<td>0.87 (0.31)</td>
<td>0.74 (0.36)</td>
<td></td>
</tr>
<tr>
<td>PR</td>
<td>0.99 (0.64)</td>
<td>1.10 (1.01)</td>
<td>1.01 (0.55)</td>
<td></td>
</tr>
<tr>
<td>U (* = significant)</td>
<td>20</td>
<td>27</td>
<td>23</td>
<td></td>
</tr>
</tbody>
</table>

Table 7.9. Perceived difficulty data for Base and PR conditions.

![Self Rated Difficulty Graph](image)

**Fig. 7.10.** Graph showing average relative perceived difficulty, over 4 trials for the Base and PR conditions.

The average perceived difficulty for the control group appears to decrease steadily over the four trials, whereas that for the PR (experimental) group remains constant, except for trial 3 where it would seem on average the subjects found it *more* difficult. This is not supported by the applied statistical test.
iii) Direction of Gaze

<table>
<thead>
<tr>
<th></th>
<th>T2</th>
<th>T3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average</td>
<td>12</td>
<td>6</td>
</tr>
<tr>
<td>sd</td>
<td>3.1</td>
<td>3.1</td>
</tr>
</tbody>
</table>

Table 7.10. Direction of gaze data for PR condition on trials 2 & 3.

![Graph showing average percentage of time spent looking in the direction of the PR screen in the PR condition.]

Fig. 7.11. Average percentage of time spent looking in the direction of the PR screen in the PR condition.

On average, subjects spent 12% of the time looking in the direction of the PR screen on trial 2, and only half this on trial 3. All subjects spent some time gazing in this direction. No statistical test is applied here since there would appear to be no requirement. This data is included as confirmation that the PR screen was at least looked at.

iv) Comments, Debriefing and Observations

Comments, both those elicited on the post-trial forms and other voluntary statements made in passing and noted by the experimenter, were approximately equally for and against the PR manipulation. Those in favour included those who clearly appreciated it (to differing degrees):

[the format of the code on the left is SUBJECT/TRIAL, where trial is 1-4, and D for debriefing]

13/T2  "It was easier having the left hand screen 'cos you could see at a glance what needed doing"

15/D   "The left hand screen was a slight help, a confidence boost to see things wiped - knowing that there is nothing left to do"
Also, there were those who admitted they missed it in the final trial:

20/T4  “Missed the information on the left hand screen”

25/D  “Was reluctant to use the left hand screen. Had a mental picture, but NOW think it would have been useful”

Comments from subjects who did not find the PR screen helpful included for example:

30/T2  “The left hand screen was not particularly useful. Because of the small number of cars being dealt with, it was possible to keep in memory the lists of tasks still to be done”

12/D  “The left hand screen was no use - was able to use the resource store etc to the same effect”

The (unsupported) observation of the experimenter was that subjects could often perform more efficiently if the PR screen is both present and used. However, some subjects expressed a preference for the single screen trials, finding multiple screens generally unhelpful.

The most common errors made by subjects were concerned with attempting to allocate more resources than were present. This may be a problem which is overcome through learning, but the PR screen could have been enhanced to emphasise the ‘current resource availability’ aspect of the problem. Had this been done, it would be reasonable to expect that more user comments at least would support the notion that the PR screen was useful.
3.2.2. The Sol Condition

A. Primary Measures

i) Strategy Time

<table>
<thead>
<tr>
<th>Condition</th>
<th>Trial 2</th>
<th>Trial 3</th>
<th>Trial 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Base</td>
<td>0.95 (0.09)</td>
<td>0.92 (0.05)</td>
<td>0.93 (0.09)</td>
</tr>
<tr>
<td>Sol</td>
<td>0.90 (0.08)</td>
<td>0.87 (0.11)</td>
<td>0.89 (0.10)</td>
</tr>
</tbody>
</table>

*U (\* = significant) 22 18 24*

Table 7.11. Strategy time data for Base and Sol conditions.

![Strategy Time Graph](image)

Fig. 7.12. Graph showing the average relative strategy time over 4 trials for the Base and Sol conditions.

In a similar way to the equivalent data for the PR condition, there does appear to be a trend in the direction expected if the manipulation were having its effect (i.e. the time, as a proportion of the time taken on the first trial, is lower on trials 2 and 3). However, efficiency measured in this way does not appear to return to a level similar to that of the control group on trial 4, which is not as one might expect. Any differences between these data series are not supported by the statistical tests applied, and so will have to be interpreted in the context of the other measures.
ii) Free Time

<table>
<thead>
<tr>
<th>Condition</th>
<th>Trial 2</th>
<th>Trial 3</th>
<th>Trial 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Base</td>
<td>1.10 (1.09)</td>
<td>1.10 (1.19)</td>
<td>1.79 (1.89)</td>
</tr>
<tr>
<td>Sol</td>
<td>3.01 (3.02)</td>
<td>3.75 (3.45)</td>
<td>0.76 (0.64)</td>
</tr>
</tbody>
</table>

U(* = significant) 13* 12* 21

Table 7.12. Free time data for Base and Sol conditions.

Fig. 7.13. Graph showing average relative free time (as a proportion of strategy time), over four trials for the Base and Sol conditions.

The presence of the Sol manipulation on trials 2 & 3 has an effect in the desired direction, according to this measure. This difference is also statistically supported. In the case of trial 4, the graph indicates that efficiency on this measure drops below that of the base condition, possibly indicating a negative effect of removing the manipulation. Such a difference is not statistically supported however.
B. Secondary Measures

i) Perceived Efficiency

<table>
<thead>
<tr>
<th>Condition</th>
<th>Trial 2</th>
<th>Trial 3</th>
<th>Trial 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Base</td>
<td>1.15 (0.21)</td>
<td>1.22 (0.34)</td>
<td>1.32 (0.42)</td>
</tr>
<tr>
<td>Sol</td>
<td>1.05 (0.25)</td>
<td>1.07 (0.33)</td>
<td>1.02 (0.25)</td>
</tr>
<tr>
<td>U (* = significant)</td>
<td>26</td>
<td>21</td>
<td>13*</td>
</tr>
</tbody>
</table>

Table 7.13. Perceived efficiency data for Base and Sol conditions.

![Self Rated Efficiency](image)

Fig. 7.14. Graph showing average relative perceived efficiency, over 4 trials for the Base and Sol conditions.

The experimental group does not show an increase in perceived efficiency over the four trials, even on trial 4, while the control group shows a steady increase. The difference between the two only becomes statistically significant on the final trial.
ii) Perceived Difficulty

<table>
<thead>
<tr>
<th>Condition</th>
<th>Average (S.D.)</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Trial 2</td>
<td>Trial 3</td>
<td>Trial 4</td>
</tr>
<tr>
<td>Base</td>
<td>0.93 (0.22)</td>
<td>0.87 (0.31)</td>
<td>0.74 (0.36)</td>
</tr>
<tr>
<td>Sol</td>
<td>0.90 (0.22)</td>
<td>0.70 (0.24)</td>
<td>0.87 (0.38)</td>
</tr>
<tr>
<td>U (∗ = significant)</td>
<td>23</td>
<td>20</td>
<td>27</td>
</tr>
</tbody>
</table>

Table 7.14. Perceived difficulty data for Base and Sol conditions.

Fig. 7.15. Graph showing average relative perceived difficulty, over 4 trials for the Base and Sol conditions.

There would appear to be a trend in the above such that subjects perceive the job to be less difficult on trials 2 & 3 when the Sol manipulation is present. There is an anomaly in the case of trial 4 where this apparent easing is wiped out when the manipulation is no longer present - to the extent of being worse than the base condition. There differences are not statistically supported and thus the final interpretation of any of these points will have to be in the context of the other measures.
iii) Direction of Gaze

<table>
<thead>
<tr>
<th></th>
<th>T2</th>
<th>T3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average</td>
<td>5.4</td>
<td>5.4</td>
</tr>
<tr>
<td>sd</td>
<td>3.0</td>
<td>3.2</td>
</tr>
</tbody>
</table>

Table 7.15. Direction of gaze data for Sol condition.

![Graph showing average percentage of time spent looking in the direction of the Sol screen in the Sol condition.]

Fig. 7.16. Average percentage of time spent looking in the direction of the Sol screen in the Sol condition.

There would appear to be no difference in the average percentage of time spent gazing in the direction of the Sol screen when present.

iv) Comments, Debriefing and Observations

The comments and observations were generally in favour of this manipulation. Many subjects (irrespective of condition) spontaneously tried to issue commands on trial 1 to achieve the effect of queuing. Only a single subject (No. 19), when asked afterwards whether such a facility might have been useful, thought that it would not have been (this subject was in a condition where it was not present). Examples of comments in favour of the Sol manipulation include the following (* indicates the subject was in a condition where they did not experience the Sol screen):

9/T2  “Queuing facility is more natural, i.e. better”

3/T2  “I liked having the queuing facility because I could set things up in advance and so did not have to be vigilant in spotting when a resource became free”

As in the case of the PR manipulation, it usefulness might only have become apparent after it was withdrawn:
6/T4 “Missed the opportunity to queue, despite the fact that I under used the queuing whilst it was available”

27/T4 “Once you get used to using a queue, having it taken away makes mistakes happen more often”

It is also clear from the comments that there is a dissociation between the facility to queue and the screen based reminder of this. The balance of comments probably falls against the need for the screen, given the queuing facility. For example, comments in favour of the the screen included:

*20/D “Queuing would probably be useful, although [unprompted:] would probably need a reminder of what I had queued”

27/D “The right hand screen is useful, but could possibly manage without it”

and those against:

*23/D “could probably remember 2-3 items”

7/D “because it was quite slow, didn’t use the screen”

13/D “the right hand screen itself was not useful. Queuing facility was very useful”

35/D “would probably have preferred something on just one screen”

There are no other observations in addition to that of subjects spontaneously trying to queue allocations.
3.2.3. The Combined PR + Sol Condition

A. Primary Measures

i) Strategy Time

<table>
<thead>
<tr>
<th>Condition</th>
<th>Average (S.D.)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Trial 2</td>
</tr>
<tr>
<td>Base</td>
<td>0.95 (0.09)</td>
</tr>
<tr>
<td>PR + Sol</td>
<td>0.88 (0.14)</td>
</tr>
<tr>
<td>U (* = significant)</td>
<td>21</td>
</tr>
</tbody>
</table>

Table 7.16. Strategy time data for the combined PR + Sol and Base conditions.

![Strategy Time Graph](image)

Fig. 7.17. Graph showing the average relative strategy time over 4 trials for the Base and PR + Sol conditions.

The graph for this measure is very similar to that for the Sol only condition - i.e. there appears to be a trend towards faster performance on both the experimental trials (2 & 3), with a carry over benefit on to the 4th trial. None of these differences is statistically supported.
ii) Free Time

<table>
<thead>
<tr>
<th>Condition</th>
<th>Average (S.D.)</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Trial 2</td>
<td>Trial 3</td>
<td>Trial 4</td>
<td></td>
</tr>
<tr>
<td>Base</td>
<td>1.10 (1.09)</td>
<td>1.10 (1.19)</td>
<td>1.79 (1.89)</td>
<td></td>
</tr>
<tr>
<td>PR + Sol</td>
<td>4.30 (4.19)</td>
<td>3.75 (2.62)</td>
<td>1.90 (2.33)</td>
<td></td>
</tr>
<tr>
<td>U(* = significant)</td>
<td>17</td>
<td>10*</td>
<td>31</td>
<td></td>
</tr>
</tbody>
</table>

Table 7.17. Free time data for the combined PR + Sol and Base conditions.

Fig. 7.18. Graph showing average relative free time (as a proportion of strategy time), over four trials for the Base and PR + Sol conditions.

The effect of the experimental manipulation (both extra screens together) displayed in the above graph is as would be expected - i.e. there is a benefit when the manipulation is present (trials 2 & 3), which disappears when removed (trial 4), rendering performance equivalent to that in the control condition. Note that only the trial 3 difference is statistically significant. Although the difference on trial 2 appears larger, it fails to be supported statistically at the chosen level, using the chosen test. Closer examination reveals one anomalous data point in each group. Also if the equivalent parametric test (i.e. t-test) had been chosen, this would have appeared to be significant.
B. Secondary Measures

i) Perceived Efficiency

<table>
<thead>
<tr>
<th>Condition</th>
<th>Average (S.D.)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Trial 2</td>
</tr>
<tr>
<td>Base</td>
<td>1.15 (0.21)</td>
</tr>
<tr>
<td>PR + Sol</td>
<td>0.98 (0.31)</td>
</tr>
</tbody>
</table>

U (* = significant) 19 30 19

Table 7.18. Perceived efficiency data for the combined PR + Sol and Base conditions.

Fig. 7.19. Graph showing average relative perceived efficiency, over 4 trials for the Base and PR + Sol conditions.

Statistically, there is no significant difference in the 2 conditions on trials 2, 3 & 4. Given the absence of even a suggestion of a trend, it is probably wisest to accept that there is probably no difference in these two series.
ii) Perceived Difficulty

<table>
<thead>
<tr>
<th>Condition</th>
<th>Average (S.D.)</th>
<th>Trial 2</th>
<th>Trial 3</th>
<th>Trial 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Base</td>
<td>0.93 (0.22)</td>
<td>0.87 (0.31)</td>
<td>0.74 (0.36)</td>
<td></td>
</tr>
<tr>
<td>PR + Sol</td>
<td>0.97 (0.53)</td>
<td>1.01 (0.66)</td>
<td>0.78 (0.61)</td>
<td></td>
</tr>
<tr>
<td>U (* = significant)</td>
<td>25</td>
<td>31</td>
<td>23</td>
<td></td>
</tr>
</tbody>
</table>

Table 7.19. Perceived difficulty data for the combined PR + Sol and Base conditions.

Self Rated Difficulty

![Graph showing average relative perceived difficulty, over 4 trials for the Base and PR + Sol conditions.](image)

Fig. 7.20. Graph showing average relative perceived difficulty, over 4 trials for the Base and PR + Sol conditions.

There is very little difference, either apparent or significant in these 2 data series. The difference on trial 3 might be taken to suggest that subjects do not find the job getting easier whilst the manipulation is present.
iii) Direction of Gaze

<table>
<thead>
<tr>
<th></th>
<th>PR screen</th>
<th></th>
<th>Sol screen</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Trial 2</td>
<td>Trial 3</td>
<td>Trial 2</td>
</tr>
<tr>
<td>Average</td>
<td>6.40</td>
<td>4.90</td>
<td>2.60</td>
</tr>
<tr>
<td>sd</td>
<td>2.40</td>
<td>1.90</td>
<td>1.60</td>
</tr>
</tbody>
</table>

Table 7.20. Direction of gaze data for the PR and Sol screens in the combined PR + Sol condition. Trials 2 and 3.

Fig. 7.21. Average percentage of time spent looking in the direction of the PR and Sol screens in the PR + Sol condition.

Subjects would appear to look at the additional screens, although not very much. There is little difference between trials 2 & 3. They look at the PR screen more than the Sol screen.

iv) Comments and Observations

Any comments and observations concerned with the effectiveness of each of the two manipulations per se have been presented above. The purpose of this section is to present any further points which arise specifically out of the combination. The only issue here is that having three screens at once was thought to be undesirable. For example:

33/D "Preferred to focus on just one screen"

Although for some subjects, this could be weighed against the advantages, e.g.:

13/T2 "Having all three screens going at once involved more time to think, but the gains were greater"
3.3. The comparative effects of the manipulations

The aim of this section is to reveal the relative effects of the different manipulations. The data presented here have all been presented above, separately, but are shown here again in combination.

3.3.1. Primary Measures

A. Strategy Time

Fig. 7.22. Graph showing the average relative strategy time over four trials for all four conditions.

The visible effects in the Sol and the PR + Sol conditions are very similar, and slightly greater in apparent extent than the PR conditions. Note also the general learning effect visible in the data for the base condition.
B. Free Time

Fig. 7.23. Graph showing average relative free time (as a proportion of strategy time), over four trials for all conditions.

The effects of all three experimental manipulations appear very similar, with the exception of the two points already discussed above (Trial 2 for the combined PR + Sol condition, and trial 4 for the Sol condition). Note also that any learning effect for the control group only becomes apparent towards the end. How such a curve would progress on subsequent trials is not known.
3.3.2. Secondary Measures

A. Perceived Efficiency

![Graph showing average relative perceived efficiency, over four trials for all conditions.](image)

Fig. 7.24. Graph showing average relative perceived efficiency, over four trials for all conditions.

With the exception of the combined PR + Sol group on trial 3, the members of the three experimental groups would appear similar on average in not perceiving themselves becoming increasingly efficient with successive attempts at the job.
B. Perceived Difficulty

Fig. 7.25. Graph showing average relative perceived difficulty, over four trials for all conditions.

It is hard to make any comment at all about the above graph. What could be said might include the following:

1) None of the experimental conditions appears to have the consistent decline in perceived difficulty of the control condition.

2) The Sol condition is much better than either of the other manipulations in that it does appear to enhance the decrease in difficulty when present.

Direction of Gaze

Comparing the amount of time spent looking at the two extra screens, it was found that a greater proportion of time was spent looking any one screen when it was the only one present than when both were present. When both were present together, both were looked at relatively less overall. Consistently though, the PR screen received a greater proportion of gaze than the Sol screen.

3.4. Tertiary Measures

There were no significant correlations between the CFQ scores and any of the primary or secondary quantitative measures.
4. Summary of Data Analysis and Conclusions

This section considers the two design manipulations independent of their incorporation in the experimental design.

4.1. PR Support

In the PR only condition, the visible trends in the two primary measures are in a direction consistent with the predicted effect of the manipulation. One of these is statistically significant. The perceived difficulty and efficiency data are not consistent with the prediction. The comments and observations are generally divided evenly, although the experimenter's observations were that the PR screen should still be useful and probably was for some subjects.

Taking all these data sources, including the direction of gaze, into account, the conclusion is that the PR manipulation did not fail outright in its aims, but was only weakly successful. Possibly its usefulness depends to a large extent on the particular individual and thus an overall effect is being masked.

4.2. Solution support

The case for the effectiveness of the Sol manipulation is stronger. Of the quantified measures, only the perceived efficiency does not show the desired effect, to some degree. Coupled with this is the largely supportive set of of comments and observations.

The general conclusion is that the Sol manipulation was at least more successful than the PR manipulation, which was itself not unsuccessful. However, there are two riders to this:

i) The dissociation of the queuing facility from the screen (which directly parallels the two reasons upon which this manipulation is based), means that in this case, the screen itself was largely unnecessary. The relatively lower proportion of time spent looking in the direction of this screen, compared with the PR screen lends weight to this.

ii) There would appear to be an effect of removing the screen on trial 4 in the case of the Sol manipulation, such that performance worsens. This is supported by the Free Time, Efficiency, Difficulty and Comments data, although the visible trend in the strategy time data is in the other direction.

5. Discussion

In the light of the above conclusions, this section addresses the question of whether the study can be considered to have achieved its main goal (the demonstration of the recruitment of knowledge in the model to design). In addition, the discussion will also
address the subsidiary questions of how well it does this, and how it could have been better. There will be no discussion of the secondary goal of demonstrating the feasibility of the whole approach, as this is considered to be implicit in the achievement (or otherwise) of the primary goal. Finally, some of the points of interest emerging from the study with respect to weaknesses in the model will be discussed.

5.1. The demonstration of recruitment of knowledge to design

The above conclusion, that one manipulation was only weakly successful, and the other more strongly successful, allows the further conclusion that in a broad way, the plausibility of the recruitment of the knowledge in the model to design has been demonstrated. However, this demonstration cannot be said to be as clear cut as might have been hoped for.

To understand where the weaknesses in the demonstration lie, it is useful to consider the extreme cases. At one extreme, it is possible that the conclusion might have been that both manipulations failed totally. In this case, it would not be possible to separate the effects of the knowledge itself from the recruitment of that knowledge - so it might be the case that the knowledge contained within the model is good, but that its recruitment, i.e. its expression in some artefact is very poor in the particular instance.

The opposite extreme is the case where both manipulations turned out to be very effective. In this case, it would be reasonable to conclude that both the knowledge in the model and its incorporation in the designed artefact were good. (There is always the additional assumption that the manipulation was not inert and that any effect was not due to something else).

The current case lies somewhere between these two extremes, and the following is an attempt to outline possible areas of weakness in both the relevant knowledge in the model and the method of recruitment of that knowledge.

A. Weaknesses in the particular knowledge recruited from the model

The first potential weakness concerns the PR manipulation. In this manipulation, the aim was to optimise the internal problem representation with respect to the real problem in the world, with a view to enhancing resource allocation decision making. The most common types of error made by subjects, at least in the early trials, related to not appreciating the current availability of resources in the system. The model treats this as just another aspect of the problem definition, and such knowledge could easily have been recruited to emphasise the status of the resources on the PR screen. The speculation is that if it had been, then subjects would have found that the PR screen reduced this type of error and thus made them more efficient.
The second area of weakness in the knowledge in the model concerns the chosen measures. Firstly, some of the measures were not very sensitive, particularly the Strategy Time measure. The most likely cause of this is the vast scope for individual differences, both in ability in general, and in chosen strategy in particular, which complex multitasking as studied here provides. It might have been more satisfactory if the model had provided some means of quantifying and allowing for individual strategic and tactical decisions.

Secondly, on the subject of measures, the chosen measures sometimes disagree. So, for example, the free time measure may say that a manipulation is making subjects more efficient, but ratings of their own perceived efficiency fail to reflect this. Admittedly, the chosen measures of perceived efficiency and difficulty were crude, but this is an aspect not addressed by the model and as such is a weakness.

Finally, it is worth considering briefly whether some of the weaknesses in the measures were due to ceiling or floor effects. Overall, this is unlikely given the fastest and slowest times achieved, however, within a particular subject it is conceivable that there is only a certain degree of improvement which can be achieved in the short term with manipulations such as these, and that greater benefits would only appear through longer term training and experience.

B. Weaknesses in the recruitment of the knowledge - i.e. its instantiation in a designed artefact

There are two areas in which the particular instantiation of the knowledge might interact negatively with its intended impact on performance. The first of these concerns the speed of the system and size of the job. It will be remembered that more than one subject commented that had the job been more demanding, a particular screen would have been more useful (this point is also relevant to the previous section).

The other area is that of the use of multiple screens. Obviously there is a limit to the amount of information which can be crammed onto a single screen and for ease of implementation, it was decided to solve the problem by using additional screens. In retrospect, it might have been better if some other way could have been found which kept the number of screens to one, as preferred by numerous subjects. This would have made the collection of Direction of Gaze information difficult though.

5.2. Tertiary goal - points of interest

This study also provides an insight into some aspects of multitasking behaviour not otherwise accessible. This section will briefly touch on two of these points of interest, in terms of the model.
Firstly, it is clear that there are learning effects - the counter claim, that there are no learning effects was never seriously entertained, but this study has provided some interesting potential insights.

These learning effects can be understood in terms of the use and development of heuristics. It will be remembered that the model incorporates heuristics in its Solver, to deal with the occasions when a sufficiently complete Problem Representation is not available, for whatever reason. In this light, two things appear in the comments. The first is that subjects must be relying on heuristics, or at least making informationally incomplete decisions anyway, since there are admissions that they only manage to attend to all the different dimensions of the problem with time:

1/T2  "Noticing more factors, got the ‘services offered’ table in use, more sophisticated plans, but also worse mistakes”

1/T3  "Getting better - fewer mistakes ... learning how many resources available”

4/T3  “Do not feel I am using size information”

32/T3  "I am not doing very well at managing this business. This is due to a failure on my part to take into complete account the respective times required for specific actions”

32/T4  “The experience was frustrating ... I failed to work out a satisfactory formula”

18/T3  “not been paying attention to the size of a car”

Secondly, one can see evidence of that subjects are attempting to derive heuristics:

16/D  “I used a pattern which seemed to work the same each time”

12/T3  “Less difficult, still making mistakes though - I have not concluded whether ’tis better to put larger cars in, or more small ones”

14/D  “put a large car in to start with, then later realised that this created problems with the polishing”

Secondly, there were also suggestions that subjects were reluctant to change a plan once it had been established (usually termed preserving the schedule in the previous chapters). For example, there were very few instances when subjects wanted to move resources from one ongoing task to another, even when the experimenter made it quite clear that this is something which they could do.
This chapter has provided a demonstration of the recruitment of the model developed in the preceding four chapters. It has been proposed that this recruitment has been successful, to an extent. The next, and final, chapter is a summary and an assessment of both the model and the development method followed.
8 Summary and Assessment

Prologue

This chapter concludes the thesis, summarising and assessing the work done.

Contents

1. Introduction
2. Summary
   2.1. The Job Attributes
   2.2. The Behavioural Phenomena
   2.3. The Model
3. The Achievement of the Original Goals
   3.1. The Model
      3.1.1. As a Scientific Device
      3.1.2. As an Engineering Device
   3.2. The Method
4. General Assessment
5. In Conclusion

1. Introduction

This is the final chapter in the thesis and is therefore concerned with both summarising and assessing the work reported in earlier chapters. In addition, there will be a commentary in the form of some related points from the literature.

The first section presents a summary of the complete set of Job Attributes, the complete set of Behavioural Phenomena, and the Model itself. Later sections are then concerned with assessment in two different ways. Firstly, there is an assessment of the model, as a product, with respect to the goals originally set out at the beginning of the thesis. This is followed by an assessment of the method both as it relates to the achievement of these goals and otherwise. Secondly, there is a more general assessment of the model and the contributions made by the thesis to Cognitive Ergonomics.
2. Summary

This section summarizes the Job Attributes, the Behavioural Phenomena and the Model. It is necessary to summarize the Job Attributes because they represent the scope of the model and the definition of the conceptual variable of multitasking. It was stated in the discussion of models in Chapter 2 that a model ought to be accompanied by such a statement of its intended scope.

The Model is summarised here since its most recent full description was in Chapter 5. The Behavioural Phenomena are restated here for completeness.

2.1. The Job Attributes

The expression of the JAs which has developed in the course of the thesis is really one of six classes of properties, each of which represents a number of states. This scheme will be maintained here also.

The first, and most important, JA concerns the role of the person. This replaced the requirement for there to be interleaving. The additional JAs could be thought of as secondary to this one, providing for the differences in the various instances of multitasking. To understand this JA, it is necessary to understand the conception of tasks upon which it is based. This conception is also crucial to understanding the model, which follows in a section below, and is therefore presented here, once.

In the conception, the world is taken to consist of objects. Objects have attributes, which in turn may be in a number of different states. An example of an object might be a floppy disk for a computer, which could have attributes specifying whether it had been formatted and whether it was locked, amongst others. Attributes of objects can exhibit affordances for change of state. Changes in the state of an attribute are accomplished by a task. In the example, the task of formatting a disk would change the state of the formatted attribute from unformatted to formatted. Probably the most important part of the conception concerns that which is thought to accomplish this state change. Resources are objects which are required in a transformation, but are not themselves transformed. In the example, a disk drive would be a resource required for formatting a disk (amongst other things). The person in such a world is conceived of as having two roles. Firstly they are an effector - that is to say they are modelled as a resource. Secondly, they have a control role, responsible for managing the allocation of all the other resources to all the necessary tasks.

The first JA takes, as its starting point, a need for attribute state transformations of more than one object to be going on in parallel. This is then further qualified in terms of the role of the person under consideration, such that they may have differing
degrees of control and/or effect roles in these transformations. It is proposed that for a job to be considered as multitasking for the purposes of the model, there should be at least a high control role.

The other JAs specifying the scope of the model are then as follows (in no particular order):

- Cooperation requirement. The presence of other people with whom it is necessary to coordinate one’s activities has implications for multitasking behaviour. This could range from being the only person present, through coordination concerned with different people being responsible for successive stages of tasks, to the extreme of sharing the pool of pending tasks with a number of others in a non-demarcated way.

- Reliability of the world. Efficient multitasking requires some advance planning. This JA concerns the degree to which such planning is possible. The behaviours of objects in the world could vary in their reliability.

- Familiarity of the person with the individual tasks. It was proposed early on that, because of the requirement to plan ahead, the person should be familiar with the tasks to be interleaved. This is arbitrary, but the model at present does not address behaviour arising from inexperience. This obviously has similarities with the previous JA, since experience is often required to be able to predict the behaviour of objects.

- Familiarity of the person with the particular task set. Although it may be the case that the person has experience with the individual tasks, they may not have experienced the particular combination before. This could have implications for their anticipation behaviour.

- Cost associated with error. In a world which is predictable to differing degrees by a particular individual, the cost associated with erroneous decisions is thought to have important consequences for multitasking behaviour.

The JAs constitute the terms of the definition of the scope of the model and so provide a framework by which different instances of multitasking behaviour can be identified, compared, and related to the model.

2.2. The Behavioural Phenomena

The following is the complete set of Behavioural Phenomena associated with multitasking, as proposed in the thesis. Each is accompanied by a short explanation, and they are in no particular order. It should be noted that it is envisaged that only a
subset of these phenomena will be apparent in any given instance of multitasking (see Chapter 2).

- Interleaving. It is observed that people are capable of controlling several tasks in parallel.

- Planning ahead. There is evidence to suggest that people are able to plan and coordinate the interleaving of tasks.

- Limited capacity of the Mental Workspace. The person is modelled as two resources (see next section). It is inferred from behaviour that the person is limited to being able to think about one task at once.

- Limited capacity of the Performance System. As above, it is similarly maintained that, mainly, the person can only be physically concerned with a single task at a time.

- Parallel Mental and Physical Activity. It is observed that a person can be physically concerned with one task and at the same time mentally concerned with another.

- Two stage interruption. Interruption is sometimes handled in two stages. The first of these is concerned with finding out more about the new task, and is thought to be minimally disruptive. This information enables the person to potentially delay the second stage of the interruption - i.e. actually doing what is required - until some more suitable time.

- Partial, incomplete, plans. Although it is observed that people do plan their activities, they do not construct elaborate, complete plans, or at least they do not voice them. Instead, the plans tend to be at a higher level and varying in detail.

- Some tasks done in passing. This phenomenon is thought to be closely related to that of only forming partial plans (above). It refers to the observation that while there is some planning in advance of what to do, other behaviour is not anticipated and arises opportunistically according to the circumstances in which a person finds him or herself.

- Some information bound to plans. It is thought that it is necessary to mentally simulate the performance of some tasks in order to remember important details. Mentally executing a task in advance and thus making these points explicit has implications for planning the coordination of tasks.

- Free time. It is observed that people can be (correctly) aware of when there is nothing for them to do. This is important since it implies an awareness of all the tasks for which they are required and when they will be done.
• Default activities in spare moments. In addition to those tasks which are planned for in advance and those which are done in passing, some tasks can be of sufficiently low importance that they will not be scheduled at a particular time, but rather slotted in around the other tasks.

• Relative or absolute expression of the schedule. When voicing their (partial) plans, it was observed that people would do so in either an absolute or a relative way. An individual might use both expressions at different times.

• Initialisation heuristic. In the case where the person had no previous experience of the set of tasks (i.e. the job), some people were observed to use a heuristic which enabled them to start doing something as soon as possible whilst they thought through a plan or schedule, which then took over and dictated what should be done when.

• Learned routine. In the cases where there was previous experience with the job, there was evidence of a learned routine which was used to dictate the order of tasks, thus removing some of the burden of planning afresh each time.

• Local correction tendency (preserving the schedule). When things did not go according to expectation, it was observed that there was a tendency for corrective measures to be applied locally, rather than forcing a global rescheduling of all the tasks.

• Scheduling may be imperfect. People make mistakes (i.e. intentional, but inappropriate, actions; the contrast would be errors, which would be unintentional slips.) and may construct plans which are sub-optimal for time, for example.

• Keeping informed. This is expressed in the model's terms as Validating the Problem Representation. It is observed that people make an effort to be aware of the ongoing state of the world in which they are operating. This is especially important in a rapidly changing, unpredictable world.

• Sharing a set of tasks. This behavioural phenomenon directly corresponds to the same JA. It refers to the observation that people are able to take the actions of others into consideration when planning their own behaviour and that of resources under their control.

• Simplicity bias. There is a tendency to choose the simpler, less demanding route. It is suggested that this phenomenon is likely to be apparent in the management of high workloads.
• Hedging. In unreliable systems, it is observed that the person may opt for the safer prediction which is associated with a lower cost should it go wrong.

2.3. The Model

The model is more than just the structure and mechanism to be presented below, since it must also include the conception of tasks and resources described earlier. It is, however, also less complete than it might be; there is no intention to specify exactly a mechanism for every one of the above BPs.

It was stated earlier that the role of the person is twofold. They are both an effector, and a controller. As an effector, the person consists of two resources - the Mental Workspace, and the Performance System. As a controller, they are responsible for allocating resources to tasks over time, and it is mainly the control role which is addressed in the structure and mechanism described below.

Structurally, the model is very simple. It is conceived in terms of three elements, a representation of the Problem, a Solver and the Solution to the problem. The problem, in the case of multitasking, is the set of tasks which need to be done together with the resources available. The solution to this problem is an allocation scheme for each resource over time. The Solver, which produces a solution from a problem, was originally assumed to be a black box, i.e. a structure whose mechanism was not of interest, since it was assumed that the problem was soluble and that any shortcomings would be due to an imperfectly specified problem. This is still largely the case, except that it has been possible or necessary to specify some properties of this solver in a little more detail. This has been done in the form of heuristics or biases - for example, the startup heuristic, or the tendency to make localised corrections. For this reason, the Solver is now thought of as a grey box.

The gross mechanism of the model is also very simple. The solution, at any time, is thought to be the best that can be constructed given the specification of the problem at that time. A change in the specification of the problem - for example, as a result of being interrupted - is considered to be sufficient for the solver to act, re-validating the (now) invalid solution. This notion of validation is extended to cover the relationship between the internal Problem Representation, and the state of the external world, i.e. the real problem. Thus in monitoring something, the person is constantly validating their internal representation of the state of the world - the Problem Representation.

It was suggested in Chapter 2 that in addition to its scope, a model should be accompanied by an explicit declaration of what its intended purpose is thought to be, and similarly, the intended user of the model within this purpose. In the present case, the model is intended to have dual standing as both a scientific and an engineering
device. From this it follows that it should be thought of as having two purposes. Firstly, it is intended to function as a means for reasoning about and predicting behaviour such that it furthers knowledge of this behaviour. Secondly, it has a purpose which is to aid in the design of artefacts. These purposes will be expanded on in a section which follows below. The user of the model must be stated as a Cognitive Ergonomics Researcher in both cases. The implications of this will be addressed in the following assessment section.

3. The Achievement of the Original Goals

The goal of this thesis was to produce a general model of multitasking behaviour, in the context of Cognitive Ergonomics. This goal was expanded in Chapter 2, to specify what it means to develop a model in this context, and what generality means. Thus the initial goals were in terms of a product. To achieve this, a process was proposed (the method of Chapter 2). Once the state of the product, with respect to the goal state, is clear, then the weaknesses and deficiencies in the process can be considered.

3.1. The Model

The broad goal of developing a model in the context of Cognitive Ergonomics was expanded in Chapter 2 to become the goal of developing a model which could serve as both a scientific and an engineering device. Within each of these goals, it is possible to consider several aspects of the model - such as its scope, purpose, intended user and status. This section takes the two sub-goals in turn.

3.1.1. As a Scientific Device

The scope of the model has already been stated in terms of JAs. This will not be elaborated on, except to say that it has obvious implications when using the model for its intended purpose. The purpose of a model as a scientific device is to form a basis for reasoning about and predicting behaviour. In this way it can contribute to furthering the understanding of this behaviour. It was suggested that for a model to have a status as a device suitable for such a purpose it should be possible to generate testable predictions. It was suggested that it is unnecessary actually to test any of these predictions; exemplifying their existence should suffice. Testing them would contribute to whether the model was strictly true, which is not an issue here.

The following are offered as examples of predictions made from the model.

i) The Solver structure is assumed to be able to produce a full, optimal, Solution if given an adequately specified Problem Representation and sufficient time. Sub-optimal plans are thought to arise through ill specification and lack of time. The
prediction is that if a person were given sufficient time, and all the necessary information, then they could construct an optimal plan. This is testable - there exists computer software designed to tackle the solution of such problems and their output could be compared to that of a group of subjects. Such a test could obviously not be conducted in real time.

ii) The model has no limits to the number of tasks which a person could be expected to control in parallel. In principle this could be tested by giving subjects larger and larger numbers of tasks to do and charting their efficiency with the intention of being able to identify a point of breakdown. In practice the model is weak when it comes to specifying what exactly constitutes a task (rather than a collection of tasks). This would have consequences for the clear establishment of an effect and the ability of others to reproduce it.

iii) The model represents the world in terms of resources, and the person of interest as two such resources. The typical expression of intention is in terms of what the individual intends to do in the future. It should follow that an individual ought to be able to state the future use, over time, of any of the resources under their control. This is an important prediction for which there seems no ready means of disproving. It might be easier to establish supportive evidence - for example using concurrent verbal protocols and directed prompting.

The model can at least generate predictions, and some of these would appear to be testable. However, it is clear that there are weaknesses in the model (discussed below). The conclusion would then be that the model has a limited status as a scientific device.

Finally, it is necessary to consider for whom the model is intended. For a scientific device, it is entirely reasonable for the intended user to be a Cognitive Ergonomics researcher.

3.1.2. As an Engineering Device

The purpose of an engineering device is to aid the design of an artefact. The study reported in Chapter 7 demonstrated that the model was able to serve such a purpose, although not strongly. Possible reasons for lack of strength were discussed which included the details of the study as well as the recruitment of the knowledge in the model. It was the case that the job used in the study fell within the scope of the model, which therefore ought to be applicable.

The model’s purpose has only been shown with respect to one user - a Cognitive Ergonomics Researcher. This is clearly not the most desirable end user of a product intended to be of use in design. It would be more desirable if the model could be
used by, for example, an industrial designer. The model's status in this respect is unknown, but it would be possible to test it in the future.

3.2. The Method

The process which gave rise to the product with the above weaknesses is the Method proposed in Chapter 2. This saw a model as a device for producing a set of behaviours in a scenario. The intention was for the model to produce the same set of behaviours as the entity being modelled (in this case a person), in a given scenario. To do this, scenarios or jobs were conceived of as specifiable in terms of attributes, with jobs having different attributes in common. The set of attributes addressed by a model would constitute its scope.

The method proposed that a model could be developed Horizontally, i.e. have its scope extended, and Vertically, i.e. increase its detail and robustness. Finally, a reasoned choice of research paradigm was made to support this scheme.

The main weakness identified in the previous section concerns the model's ability to generate fully testable predictions, and its recruitment to design which was both poorly demonstrated and by other than an optimal user.

The inability to test some of the predictions of the model probably lies in the conception upon which it is based and its inability to deal adequately with the question of what constitutes a task. In the thesis, this problem is dealt with by arbitrarily adopting a particular view. This is not really attributable to the method, but more to the knowledge used with it. The method of Chapter 2 does not pretend to proceduralise the development of a model, but is rather a guiding heuristic which must be used along with other knowledge. It is this other knowledge which is probably lacking in this case.

It has already been stated in the thesis (in Chapter 2) that a more comprehensive framework for understanding design would probably be necessary to improve significantly the development, and hence the recruitment, of a model whose purpose was to aid design. In Chapter 7, the model was recruited in an evaluative manner, such that a design was first created and then assessed according to the model. More useful might be a model which was able to prescribe elements of a design for a particular situation. It is not clear, at present, how such a model could be developed in Cognitive Ergonomics.

At present, the model is at least usable by the particular Cognitive Ergonomics researcher responsible for its development, and probably it is reasonable to assume, other such researchers. However, for it to have a useful impact on the design of artefacts in the real world, it needs to be usable by others with less specialised
knowledge. It was acknowledged in Chapter 2 that to develop a model for more than one user at a time is likely to be very costly, and so the intended user would have to be chosen at the outset. It would not be sufficient merely to state such an intended user. It would probably be necessary to have a view of both their existing knowledge, and how they went about design. In the current scheme, this then ought to have an influence on the vertical development of the model (on the choice of analogies to recruit, for example). There is probably also an interaction between the intended user of the model and whether the model is to be used in a prescriptive or an evaluative manner. It is likely that a prescriptive model could be constructed which required less background knowledge in its user.

The method also has some strengths. It has proved able to generate and integrate a large number of behavioural phenomena. Importantly also, the model is not a development of an existing model of multitasking behaviour, but rather was created afresh. To the extent then that the weaknesses are tolerable or rectifiable, the method could be recommended for other conceptual variables which are being modelled for the first time.

4. General Assessment

The final model which has been developed throughout this thesis is different in many ways to that originally proposed as a prototype. The purpose of this section is to consider some of the features, properties and assumptions of this model in the light of other work available in the literature. Some of this could have been incorporated into the initial model, undoubtedly resulting in a different final model, had it been identified as relevant in the early stages.

In Chapter 3, the prototype model was established on the basis of three Job Attributes. The choice of these attributes has obviously had important consequences for the way in which the model has developed. The second JA in that original set stated that the people observed should have some experience with at least the individual tasks to be interleaved. This had the consequence that the model is based largely on observations of expert subjects. A more detailed consideration of what it means to be expert reveals interesting links between the proposed model and other research.

The concept of expertise is usually associated with that of skill. However, in psychology, there is no clear cut agreement on what constitutes skill or expertise. For example, Leplat (1989a) defines a skill as "the possibility acquired by an individual of executing a class of tasks at a raised level of efficiency". This is a very general statement, with which it is easy to agree, but which still leaves many
questions to be asked. The conception of cognitive behaviour offered by Rasmussen (1983) will be used as a basis for understanding some of the aspects of expertise.

Rasmussen (1983) proposes that cognitive behaviour can be characterised in terms of three levels. These he terms the Skill based, Rule based and Knowledge based levels. The lowest of these levels, the skill based level, is associated with highly automatic, smooth, integrated patterns of behaviour, occurring without conscious control - i.e. without the intervention of a ‘Mental Workspace’. The middle, rule-based level, would be associated with behaviour controlled by the operation of stored rules or procedures. The highest level, associated with behaviour observed in unfamiliar, novel, situations is termed the knowledge based level. Behaviour at this level typically involves reasoning, problem solving and planning. This sort of mental behaviour is in turn that which psychologists often interpret in terms involving the operation of Mental Models.

It is important to emphasise that these three levels are not strictly demarcated and that observed behaviour is likely to be interpretable as a mixture of behaviours at the different levels.

For the present purposes, it serves to take the two extremes - skill based and knowledge based behaviour and explore two particular areas of the literature. These will be the properties of mental models and knowledge based behaviour, and the development of routines or automatic skills.

The first thing which is apparent from the literature is that there are qualitative differences in knowledge based behaviour depending on the amount of experience of the individual concerned, although the role of experience will be discussed in more detail below as it applies to the development of skills and routines.

Larkin (1983) was concerned with the differences in the way novices and experts solve physics (applied mathematics) problems. She proposes that the experts have the advantage as a result of being able to construct a mental representation (i.e. a model) which contains fictitious entities such as force and momentum. The novices are confined to a simple representation of the physical entities present. There is some similarity between this difference and the distinction between models-1 and models-2 discussed in Chapter 2.

DiSessa’s (1983) account of novice-expert differences in problem solving behaviour utilises a concept of cuing priority associated with each of the available phenomena. An attribute’s cuing priority is some measure of how strongly it suggests itself as useful to the solution. Her suggestion would be that the cuing priority of the relevant parameters of the problem is more likely to be higher than other irrelevant factors in
the expert compared with the novice. Errors arising in this way are analogous to the 'first come, best preferred' bias of Reason (1987).

The common theme between Larkin and diSessa, expressed in terms of the current model, would appear to be that novices could be expected to be less efficient in solving the multitasking problem as a result of paying attention to inappropriate aspects of the Problem Representation. Note that the theme running through the development of the model has been to assume that the Solver could be a black box and capable of solving the problem if it could be sufficiently specified. In this way, deficiencies in Solutions were modelled as resulting from inadequately specified problem representations. The current points could be integrated into the model by asserting that the world-sampling strategies of experts differed from novices. Alternatively, it might be appropriate to reconsider whether the Solver can remain as a black (or grey) box with such optimal capabilities. Dörner (1990) documents the relationship between questions asked of the environment by the subjects and the decisions subsequently made by them. He finds a difference in the pattern of questioning and decision making between novices and experts such that experts are more likely to ask several questions and then make several decisions whereas for novices, the two are more mixed together. In the expression of the model, this would be a case of experts tending to validate their Problem Representation first, and then acting.

This aside, there are other suggestions that expert's mental models are far from perfect. Rouse et al (1986), for example, suggest that even after instruction there is a tendency to retain the naive model (or presumably aspects of it). It is easy to understand how this might lead to difficulties in subsequent novel situations. Conversely, Williams et al (1983) propose that subjects may in fact utilise more than one model, appealing to a separate model specifically in novel situations to resolve ambiguities.

Utilising multiple mental models, between which there are almost certainly clashes, would be one reason for observing inconsistent behaviour. There has been no attempt to document such in this thesis, since one would probably have to confine the sampling of behaviour to a smaller number of subjects over a more prolonged period, to avoid confounding any effects with individual differences. Inconsistency, however, could also arise in other ways, according to Brehner (1987). His conclusions are based on studies of decision making using statistical regression analysis of symptoms and judgements. Briefly, he states that in such problem solving behaviour, people use only few of the available symptoms and integrate the information additively. (Dörner 1987, discussed below makes a very similar point, saying that subjects tended to think in causal series rather than perhaps the more
appropriate causal nets). Brehmer goes further, and says that “Inconsistency seems to be a general feature of cognitive systems faced with tasks that contain uncertainty” (p90). Increasing the complexity of the system would be expected to decrease its predictability and thus the consistency of the control behaviour associated with it. Currently, the model developed in this thesis has no clear position on the limits of its functioning - probably as a result of the type of behaviours observed. However, one potential development that could be explored would be to put a capacity limit on the problem representation. This is not in fact a great leap since it is already implicitly associated with Working Memory, which in turn is classically modelled as having a restricted capacity (e.g. Baddeley, 1986).

It is perhaps easier to integrate into the model some of the following more general documented properties of 'knowledge' based behaviour as, for example, heuristics in the solver (although, as stated above, there is probably a limit to which one can productively pursue such a mechanism).

It will be remembered that one of the heuristics proposed as a result of observation was that of simplicity - people could be expected to take a course of action which resulted in less effort. Norman (1983) proposes something very similar, although in slightly different terms. He suggests that people are 'superstitious', in that they maintain unneeded behaviour patterns (i.e. actually doing more) because they cost little in physical effort but save on mental effort. It would seem, then, that it is the unobservable mental effort expended by the subject which is really worth studying. Norman acknowledges that one of the reasons for such behaviour might be that with a little extra effort in some situations, the same behaviour can apply more generally, across situations, thus resulting in a simplified model, which is presumed to be easier to remember and reason with. Pursuing the theme of mental effort, there is other evidence in the literature which support the notion that people are unwilling to think. This is obviously important for modelling problem solving (in the general sense, which would include controlling a novel, complex situation).

Dörner (1987, 1990) seems particularly concerned with identifying such biases in individuals. His data concern the behaviour of both expert and novice subjects in controlling a computer based simulation of the economy of a small town or country. The following are amongst his conclusions:

Thematic Vagabonding. This is the name given to the observation that when faced with a difficulty, subjects tend to leave the problem and try to tackle something they perceive to be easier, i.e. this is an escape behaviour.

Encystment. This is the label he chooses for the observation that subjects may 'enclose' themselves in a particular (sub-) problem, although not usually a difficult one (in which case it would be hard to reconcile with Thematic Vagabonding).
Delegation. Subjects delegate if possible. It is interesting to consider what might have been the case in the behaviour of the Computer Operators of Chapter 5, where there was at least sharing of tasks and the potential for delegation, had the job been more difficult. Possibly the overall cooperation might have suffered.

There are specific situations for which he reports qualitative differences in behaviour - particularly in emergency situations. Here he suggests a reduction of the intellectual level, (i.e. using Rasmussen's scheme, presumably moving away from knowledge based towards skill based behaviour), resulting in stereotyped, reflexive behaviour. Objectively, there are both advantages and disadvantages to this. Armies spend much effort in drilling routines into soldiers so that they react predictably in emergencies, so if the stereotyped behaviours are appropriate satisfactory behaviour may emerge. However, some emergencies will be novel situations for which stereotyped reactions might be inappropriate.

What Dörner (1990) terms 'repair shop principles' can also account for more general mistakes that people make. Behaviour of this type is concerned with just tackling the obvious, apparent problems, without a view to the global or longer term consequences. In the behaviour documented in this thesis, similar behaviour is also apparent as, for example, failing to take all the competing demands for a resource into consideration when allocating it (see the cooking study reported in Chapter 4 for evidence of this). In the model, this could be explained in terms of an inadequately specified Problem Representation. However, it could also be an effect of the localised planning bias in the Solver suggested as beneficial to the person in some situations. It is not unreasonable that such a heuristic bias could generate mistakes in some circumstances. Possibly the most interesting point he makes is that subjects show insufficient consideration of processes in time. The ability to do this is considered central to the ability to manage concurrent tasks in the prevailing model. However, there are observations in the three studies documented in this thesis which would support his statement (and consequently, multitasking is often sub-optimal).

Finally, in this discussion of knowledge based or mental model derived behaviour, Dörner (1987) makes two points concerning individual differences. It will be remembered that there was an attempt in the experimental study of Chapter 7 to identify individual differences using the Cognitive Failures Questionnaire, which failed. Dörner reports that in his complex planning and decision making tasks, there were distinct differences between strong and weak subjects. Some subjects (the strong ones) were able to deal successfully with complex structures without prior training. Others, however, have difficulty even after some training. It is certainly the present experimenter's belief that it should have been possible to identify some individual differences along these lines in the data from the Car Wash simulation.
Such findings, particularly if they were pronounced, would have serious implications for many cognitive modelling enterprises since the overriding assumption is one of commonality across individuals. The second, and lesser point, is that he suggests that there is a significant relation between self assessment and performance for poor performers. If this were a general result, and was specifically not the case for superior performers, then this might be expected to have resulted in a masking of any effect between self assessment and performance in the Car Wash data.

Leaving knowledge based behaviour and turning to the other extreme of the Rasmussen (1983) framework - the development of skilled, automatic behaviour, the most obvious question to answer first is why routinization might be useful. Planning from scratch in a given situation is taken to be time-consuming and therefore unattractive (see the above discussions on biases in problem solving). However, a great deal of the time, the person is faced with situations which are familiar, which have been encountered before, and which are to a great extent, predictable. It is generally accepted that the person uses this predictability to their advantage and develops "routines" which can be executed with the minimum of conscious control.

The mechanism by which this is thought to come about has been the focus of much effort in psychology, and it is worth considering two major views. The first is the ACT model proposed by Anderson (Neves & Anderson 1981, Anderson 1982). This model assumes two stages in skill acquisition - an initial declarative stage in which the person establishes a database of facts about the world, in the form of a propositional network, and a subsequent procedural stage in which the control of behaviour is managed by a production system (IF - THEN rules). The transition from the first to the second stage they term knowledge compilation, which itself divides into composition and proceduralisation. Although the expression in ACT is novel, the stages themselves have a longer history and are likened, for example, to those of Fitts (1964). They assert that the major qualitative features of skill development - speedup, elimination of verbal rehearsal and decrease in flexibility can all be explained by compilation. An alternative mechanism for the qualitative changes in the development of a skill would be that of Shiffrin and Schneider (1977). In this scheme, the differences in skilled and unskilled performance are attributed to a capacity limitation - skill acquisition is the transformation of a task from being controlled and thus capacity bound, to being automatic and capacity free. It is not reasonable to go into the full implications of the two (and other) schemes here; the interested reader is directed to e.g. Cheng (1985) or Shiffrin and Dumais (1981).

What is clear, is that there would appear to be a general consensus endorsed by Experimental Psychology that skill is not an all or nothing phenomenon. Rather, skilled performance can be seen at many levels of detail and acquisition seems to
consist of changes from smaller to larger units of performance (see e.g. Duncan 1975). In this way, there can be seen a hierarchical structure in cognition. In the ACT model, this could be thought of as successive compositions of groups of productions into single productions. The reader will have noticed a strong similarity between this and the macro-rules of van Dijk (1980) described in Chapter 1. In that case, the discussion was concerned with what constituted a task and where one could cut across the hierarchy (stopping rules). However, it is now possible to see that the degree of experience, and thus skill, of the individual concerned should play an important role in such hierarchical task analysis. It will be recalled that the concept of hierarchy in routines has already been alluded to in the model (see discussion of thickness and length of threads in Section 6.1 of Chapter 5).

An interesting aside is whether it is appropriate to consider skill acquisition as a continuous or discontinuous (i.e. stepped) process. Fitts (1964) suggests that it is primarily a continuous process, although the fine grain structure of the performance may involve discrete operations (and thus gradual shifts may be apparent in learning). This would have implications for our observations and hierarchical analysis of performance, particularly the search for a discrete stopping criterion. There are other ways of representing skill acquisition which utilise a hierarchic structure (for example, Gagné 1977), but this seems to be the most useful for the present purposes.

The ACT scheme of Anderson does provide a convenient mechanism for understanding some of the properties of skill or routine development seen in the literature. For example, Moray et al (1986) point out the great variety in the sequences of actions used to achieve the same goal. Such differences across subjects would mean that different routines might be compiled and thus create different skill-task hierarchies, resulting in problems for the Psychologist attempting to analyse the skill. Differences within subjects are more directly a problem since consistency is viewed as central to the establishment of productions and thus skills. Brehmer (1987) makes a similar point. He states that to learn from experience, people would have to have hypotheses relevant to the task. However, these hypotheses must be statistical, due to the uncertainty present in many complex situations. The problem is that people don’t employ such statistical hypotheses (see the earlier discussion of the biases and weaknesses of Knowledge based reasoning behaviour), and therefore they don’t benefit from experience as perhaps they might.

Leplat (1989a) makes the point that a high level of skill is often associated with inflexibility and ‘set’ for a particular solution. This is often termed Einstellung and can be understood in terms of ACT (Anderson, 1982).

Leaving the issues of the two modes of cognitive behaviour, it is possible consider two other relevant points of contact between the developed model and the literature.
Chapter 8

Taking the nature of the Solver, it is assumed in the model that this structure contains the ability to construct a schedule. Sanderson (1989) has compared the ability of humans to construct schedules with formal rules for the same (as would be embodied in computer software to perform the task). The finding is that both have weaknesses and Sanderson’s proposal is to consider a combination which plays on the strengths. It was a conscious choice in the development of the model to assume this ability and account for sub-optimal behaviour in other ways. Possibly this was optimistic, especially given the number of additional heuristics and properties proposed for the Solver. It also needs to be remembered though, that there is more to multitasking than just the abstract scheduling investigated by Sanderson.

Finally, it is apparent that there is a developing body of research on what Harris (1984) terms Prospective Memory. It would seem important, in the context of planning future intentions, to consider the ability of an individual to remember to do something in the future. Harris presents what is essentially a review of some of the issues, rather than definitive statement of the properties of this category of memory. Two points which do emerge are that he suggests that prospective memory in the short and long terms are likely to be different, and that in the short term at least, it is easy to become absorbed in some other task very quickly and thus forget the intention.

Turning to the general contribution of method for developing a model with certain properties, it has weaknesses, but is probably still worth consideration. It is tempting to put both of the major weaknesses down to things which might reasonably be expected within a framework for Cognitive Ergonomics. This field is currently still young and developing and the necessary conceptual advances may be available one day.

Even given these limitations, it is suggested that the model has potential to be developed further in the future. For example, it could be enhanced to have an additional purpose to support a computer simulation of multitasking behaviour. The model has been developed based on just one type of data - derived from concurrent verbal protocols. It might be instructive to continue the development method with alternative data collection and analysis techniques.

5. In Conclusion

This thesis has derived and utilised a method for developing a model of multitasking behaviour. In the course of this development a large number of behavioural phenomena have been identified. The model is not perfect, but explanations of some of its weaknesses have been proposed. The model has been assessed both for itself and with respect to Cognitive Ergonomics more generally.
Bibliography


Bainbridge, L. 1979 Verbal reports as evidence of the processes operator's knowledge.


Brehmer, B. 1980 In one word: not from experience. *Acta Psychologica* 45 223-241


Hodgkinson, G.P. & Crawshaw, C.M. 1985 Hierarchical task analysis for ergonomics research. Applied Ergonomics 16(4) 289-299


Nisbett, R.E. & Wilson, T.D. 1977 Telling more than we can know: Verbal reports on mental processes. *Psychological Review* 84(3) 231-259


Rouse, W.B. & Morris, N.M. 1986 On looking into the black box: Prospects and limits in the search for mental models. *Psychological Bulletin* 100(3) 349-363
Sanderson, P.M. 1989 The human planning and scheduling role in advanced
manufacturing systems: an emerging human factors domain. Human
Factors 31(6) 635-666.

Sebillotte, S. 1988 Hierarchical planning as a method for yaks analysis: the example of
office task analysis. Behaviour and Information Technology 73 275-293

Shepherd, A. & Kontogiannis, T. 1987 Development of a specification for an automated
instruction system for training process control skills. Paper Presented at
First European Meeting on Cognitive Science Approaches to Process

Shepherd, A. 1975 Hierarchical task analysis and training decisions. Programmed
Learning and Educational Technology 22(2) 162-176

Shepherd, A. 1976 An improved tabular format for task analysis. Journal of
Occupational Psychology 49(2) 93-104

and their acquisition (ed J.R. Anderson). Lawrence Erlbaum Associates:
Hillsdate New Jersey.

Smith, E.R. & Miller, F.D. 1978 Limits on perception of cognitive processes: A reply to
Nisbett and Wilson Psychological Review 85(4) 355-362


Stammers, R.B., Carey, M.S. & Astley, J.A. 1990 Task analysis. In Evaluation of
human work - A practical ergonomics methodology (ed J.R. Wilson &

Suchman, L.A. 1987 Plans and situated actions - the problem of human machine
communication. Cambridge University Press.

Van Dijk, T.A. 1980 Macrostructures. An interdisciplinary study of global structures in
discourse, interaction, and cognition. Hillsdale, New Jersey: Lawrence
Erlbaum Associates.

Wallis, D. 1980 Comments on the criteria for judging models of man. In Models of man

Jones). The British Psychological Society.

218


Appendix A

This appendix contains additional information relevant to the study reported in Chapter 4. The first item is a full protocol, from subject 7. This has not been edited in any way except to underline the sort of statements which were taken to be of interest in the analysis. Line numbering is provided so that the reader can identify the position of excerpts in the main chapter. The use of brackets is as described in Chapter 4 (Table 4.2.), with the addition of “(??)” which represents a comment made by the observer which is unintelligible in the transcript.

1. Protocol of Subject 7.

(Just tell me any thoughts that you have as you read it) I've never cooked trout fried in butter, with almonds before. Oh, it tells you how to do it. (Just general instructions, a guide). Well, obviously, reading about the trout, the last thing to prepare is the browning of the trout I should think. Roast potatoes. OK, well the first thing that I've got to do is turn on the oven, for the roast potatoes for later on, and the next thing that I've got to do is get the potatoes ready so that they can be par-boiled, and I'm going to need some hot water for the potatoes so I may as well put that on next, erm, and I can get the carrots ready at the same time as I'm getting the potatoes ready, and I don't need to do anything to the fish for the moment. So that's what I'm going to do. Just check this again, and I'd better check that I have put the oven up to the suggested temperature which is wrong at the moment. (What temperature had you picked before?) At first I put it up to about 350. (And what made you pick that?) Erm, I put it to about 300-350 because er I was assuming that that was quite a hot temperature to heat the oven to, erm, but I think its probably better that I do put it up to about 400-450 now, so that I don't have to wait later on to get it to that hotter temperature. OK, now I'm going to boil some water, and I'll use the middle pan, because I won't want that many potatoes really. For roast potatoes I'll probably want about 3 potatoes, so I'll use this pan here - the middle sized pan, and there isn't a kettle so I can't quickly boil some water so I'll have to do it the hard way. And there's no difference in the ring size so I'll put it on this one and turn this up to the maximum, and I'm not sure whether or not you are supposed to put salt into water when you par-boil them (If there are no specific instructions, then you can improvise) That's right, its just deciding what I want to do - I think in fact I'll rub salt into them - a little bit of salt just before I put them into the oven. OK, so I'm going to get the potatoes ready now - I assume that there is no difference in these 2 bags except that these have already been opened, last night. But I won't be greedy - I'll just make do with these (Points out scrubber) Fine, good, OK. (??) Yes it will be thankyou, its fine, and because they are newish potatoes, I don't think that I'm going to need a knife, so I can just scrub them clean, which is very handy. And I
haven't looked at the clock yet, but I think that I must have put the water on at about 25 past 6, erm, and the potatoes should take about 20 minutes to cook, but I may have said, I can't remember, how long to allow, but I can test them with a fork in about 10 minutes after the water has boiled, and after I have put them into the saucepan. I can check then. I don't know, well I've never really thought about how long potatoes take before really - I just do it by - you know roughly but you don't necessarily know that that sized potato is going to take 12 minutes in hot boiling water, you just sort of stick them in and test them with a fork every now and then after about 10 minutes I suppose. But the thing with roast potatoes is that you have to make sure that they are done all the way through. Oh, and the water is beginning to boil, so I'm going to turn it down a bit, but I'm not going to turn it down too much. And looking at it, this won't be enough water, so I'm going to add some more, because I want to make sure that the potatoes are covered so that they are covered all the way through, so now I'm going to turn it back up again. But I've gutted trout before. Well when I was on a fishing holiday in Ireland, all we had was trout, and we had to cook it just about every way you can, but it was mostly roasted, but it was never roasted with flaked almonds. Well, I don't know, it might add to the taste quite a lot, and this is my last potato, so I'm now beginning to think about my carrots and how I want to do those, and whether or not I'll leave the skins on, and I suspect that I probably will, and I've got to cut a little bit out of this potato because it's a black bit and its bad, and I can hear the water boiling again, which is OK, because these are just about ready, once I've cut these bits out - I'll cut it into the bin, which will save me picking them up and chucking them in later. You didn't get very good potatoes here! Ha ha. (research grant cuts, you know). And I've just spotted a green bit that I haven't spotted before so that's coming out as well. And I'm left with potatoes the size of marbles. Right, I can put my put my potatoes in now, and once I've done it I'm going to check the time, and I still haven't quite covered them, but I think they'll be alright if I turn them over in a minute. So its just before 25 to 7, so I'll look at them again - if I haven't got too much else to do, I'll probably look at them again in a minute. Oh I'll leave that on full actually, so I'll put the lid partly on so that they'll cook a bit quicker, but not too much because they might boil over and I don't want to have to keep looking at it, so now I'm going to get the carrots ready - I'm not going to do anything complex with the carrots, I'm just going to take off the ends and erm cut out the bad bits if there are any - I'll scrape them a little bit I think. And then I'll just put them into a saucepan of cold water - ready for later on, or I might even use some of the potato water. I'm not sure - I might use the potato water later on. But the carrots are going to take, I would think, about 7 to 10 minutes, and I like them crispyish, so I probably don't need to do that until that last stages of the trout. Or when I'm frying the trout, so that it all comes together at the right time. I'm really looking forward to this meal. And my mind has gone back now to my cycle ride here this evening - so I'm not really concentrating on
scraping the carrot at all. erm and I'm going to put the carrots on the plate for a minute I think. Well no I'm going to put them in cold water so that they don't - so that they stay freshish. But I think that I'll put them in whole rather than chopping them, because it will be less fiddly. And I'll just put them to one side and wash the knife, in case I need it in a minute, which I probably will. Right, I'm just going to check now what I'm going to do with the trout now I think. And, erm, I've got to wash it in cold water and dry it with the kitchen towels provided, and then I've got to make a mixture of seasoned flour and herbs. And the herbs are there, right. So, I'm going to check again on the roast potatoes and how long they are going to take once they're in the oven, which is 10 to 15 minutes. So, if the fish needs to be fried either side for 6 minutes then that's a total of 12 minutes, so in theory I should put the roast potatoes in and the start frying the trout. I guess that will be alright for the potatoes. And when I'm frying the trout I'll have to keep basting the potatoes because I like them quite brown and crispy. Erm, and I'm just checking again what I need to do with the trout and the almonds, and on this browning bit, and you should allow a minute it says, to brown in with the almonds. So, that's OK, and I'm thinking now about my plate when its ready, and whether or not I should warm it up. But as there isn't a grill or anything like that, I'll probably just put it into some hot water before I serve up everything onto it. So, just checking again on putting the roast potatoes into the oven, and I've got to make sure that the oil is hot when I put the potatoes in, and I can do that in a couple of minutes. I think I'll wash the fish first and get that ready, erm, right! I'm going to wash the fish now, and while the fish is drying again in that I can make the seasoning for it. And I'm going to check the time again for the potatoes and I think I'm going to test the potatoes and I might turn it down a bit. And they're doing reasonably well. And that water's gone green! Ha, and I'm worried - I've never seen green water before (??) No, its definitely greency yellow. Right, I'm going to turn these right down now, so that they don't over cook, because I don't want them too soft by the time that they go in the oven. I'm going to wash the fish, and I'm going to get the kitchen towels ready for when I take it out of the cold water. And I'll get another one ready just in case, and I'm reading these instructions again, checking on the seasoning - its a cold, slimey fish, so I'm going to take the fish to the sink because I don't want to drop it. Well its not that cold, and it smells alright. And I want to see how much you spent on it. Oh! You spent a lot. And now I'm going to wash the fish, but its quite clean already, so I don't need to do too much to the fish, and its not that slippery, so I don't need to hold on too tightly, which is a relief, and its very clean. It still looks very fresh because its got a good sheen to the skin. So that's that, really, so I just need to dry it a bit - I'll shake it a bit first, so that the paper towels absorb most of the moisture, and so that I don't need to get some more paper towels out to dry it again. Right, I'm going to leave that there for a minute and now I'm going to make the seasoning, but first I'm going to wash my hands. And then I'm going to dry my hands so that I don't get flour all over them. OK, so I need a plate for this seasoning because - because er a bowl might be better. I want to make sure that - I think that I can
probably use a plate actually. Yes. [reads seasoning instructions] So I need a tablespoon and a teaspoon. That's not a tablespoon is it? I'm wondering if this is the best way to do it now, I think I probably should have used a bowl. (??) OK. Well, I'll add a teaspoon and a bit. I'll put this back out of the way again, and hope I won't use it, and I've mixed this well up, and I'm worried about the potatoes again in case they're overdoing. So I'm just going to spread this on a plate and then check the potatoes - cos I might need to turn them down a little bit, because I'm not ready to do anything with them yet. And, they're getting quite well done, so I'm going to stop playing with the fish and heat some oil up in the baking tray. This must be butter - I think I need oil. Right, I'm going to put a bit of oil into the baking tin, and then I'm going to put that in the oven and heat it up a bit. No problem. Well I'm just going to judge this actually. Its a huge baking tin for three potatoes (Notice that the last person didn't use enough oil, and burnt it, so use plenty!) OK, well I'll add a bit more then. Right, well I'm going to put this in the oven for a minute, and the oven is very hot. OK, well its just before quarter to, so I'll play with the fish a bit more, and I think in fact I'll take the potatoes off and I'm not going to use that water because it looks a bit odd, to boil the carrots in. And I'll just leave these here for a minute, and play with the fish again. (Were you anticipating using that water?) Yes I was earlier, I thought I might use it, but looking at the strange green colour, I don't think I will. Right, the fish is ready now I think, unless I have to, no, I just have to mix it into this. And I don't think a plate was the best way of doing it, because I'm going to get flour onto the table, and the floor. Everywhere! And now I'm going to put flour into the middle, and it says coat it very well, so that's what I'm going to do. And I'm thinking again about the oil in the baking tray and if its heated up, and after I've done this I'll check it, because I think it'll probably be ready. And thinking again about the potatoes because I think they'll probably take as you said 10 to 15 minutes, so I've got to move sort of sharpish after I've done this, because I've got to start frying this, because if it takes 6 minutes each side, then, er approximately, well I've got to make sure that the potatoes are alright, so, well, the fish looks very pretty now. (Is there any one thing that you are more worried about than anything else?) Erm, I'm worried about my potatoes not being done enough, because I quite like them crispy. (Are you going to allow for that?) Erm, I'm going to give them a few minutes more than you suggest, I think, and I'm also going to make sure that I baste them quite a bit, but the only thing about basting them is that you open the oven a lot, so you loose some heat, so I'm thinking a little bit about that, but I haven't thought how many times I'm going to baste them yet. If I put them in for 15 minutes, I'll probably look at them once every 5 minutes, something like that. So now the fish is ready, I think, I'm going to get the baking tray out. I'm going to wash my hands first, again, but they'll get flouiry again in a minute when I pick the fish up. Erm, it reminds me of cooking Sunday dinners at home when I was little - the thought of basting potatoes. So I can use this tea towel to take the baking tray out. Ooh, and I think I've burnt the fat - no I haven't. Its just very hot in here, and its a smelly oven. It won't shut, and my eyes are watering, and I
can't see anything. And now I'm going to make sure that these are - well I'm not going to
need these spoons any more, so I'll chuck them in the sink for the time being, and I don't
need this kitchen roll any more, but I'll use that in a minute to wipe up the flour. And
that's not too hot to put on there, so I'm just going to make sure that these are all well and
truly basted, or covered in oil even. (Is the oil warm?) The oil is very hot now, yes, and
so is the potato - in fact its rather a hot corner here. Well at least I don't have to wash up a
fork, and I can, yes, its the way that I've always done it though. Right, well its a huge tin
for three little potatoes. But there isn't another, so that one will do. My hands are really
oily now. And now I'm going to put the potatoes in the oven, and I'm going to check the
temperature again of the oven, and I'm going to put them on a middle shelf - I should have
checked that before, so I'll need to get another dish-cloth (The top shelf is probably best in
that oven) Yes? Well I'm not sure if there's enough room. Ah, OK, I'll put them on the
top shelf then, you're right, so that'll do. (Is it shut?) Yes. And the oven is at 450, and it
is just after 10 to 7, so I'll try and look at them every 5 minutes now, I think. And I don't
think I'm going to cook the trout for a few minutes, so looking at the carrots, they're in
cold water, so they'll need about 10 minutes, so I'll put the carrots on, lets think, at about 5
to, 4 minutes to. No, mm, I'll put them on in about 7 minutes or so. I won't need this
saucenpan again, so, in fact I could do a little bit of washing up while I'm waiting. Where's
the plug? (Under the dish cloth) Under the dish cloth, oh yes, you're right. I'll see what
else I can wash up as well, erm, I shouldn't need the knife again actually, but I might do. I
can clean this up a bit. And, I'm not thinking about anything at the moment, except that the
oil will take - I'll have to melt the oil and the butter in a minute. I think I'd quite like to use
butter to fry the trout in, I think. Because I think it will taste nicer. That looks very clean,
except that its got a green tinge now. So I'm worrying a little bit about the potatoes. And
whether or not they'll taste green. And there's the lid of the saucepan to do as well. I can
dry these up now - well I might leave them to air dry actually, cos, there's not much point
in drying them if I don't have to, and I can hear the potatoes cooking, so I'm going to look
at the fish again now, and what I have to do with it. So, erm, I've got to shallow fry it in
butter or a mixture of oil and butter. Right, well I'm going to use butter, or maybe I'll add
a bit of oil as well. (The butter is in the tin foil) Yes, OK. Right, I'm going to look at the
clock again now, and I'm going to look at the potatoes in a minute or so, to baste them,
because they'll be drying out quite a lot, because the oven's very hot. And I'm going to
move all this out of the way because I don't want it, but I do want the flaked almonds, so I
think I'm going to open these and look at them. Well, and I'm going to eat one. Its very
good, and I'm going to put some on a plate, or maybe I'm not, yes I am going to put some
on a plate, to judge how many I'll want, and I want a lot, and I'm really looking forward to
this meal. So that's the almonds done. OK, now I'm looking at that wooden spoon and
wondering if I should have used it already ha, but I don't think so. I think I might well put
the carrots on now, as they're in cold water, but I'm going to have to slice them first. So
I'm just going to take them out individually and cut them into the pan, so that I don't have
to use that chopping board. Sometimes I might use a chopping board, because I think that
you get a more even carrot. (Good grief) But I'm not too fussed really. And the next
thing I do after I put these on at a slow heat, after I've checked the potatoes is to get the
frying pan out and to start melting the butter, and I'm going to need a fish slice or
something like that, of which there's one under there, so that's no problem. Right, so
that's the carrots and I'm going to put them on here now. And I'm going to turn it on
slightly, to a low heat, and now I'm going to look at the potatoes, and I'm going to take a
spoon with me. Knocked back by the heat again. They seem to be doing very nicely, and
I didn't rub any salt into them, I forgot, I said I would, and I forgot, so I'll just have to
make sure that I put a lot of salt on them when I eat them. OK, and its 2 minutes to 7 and
I'll put the trout on now, and because its 2 minutes to 7, I'm going to turn the carrots up
slightly (Are you aiming for any particular time?) Erm, well, if the trout is going to take
about say 12 minutes, plus putting that in, I mean that's going to be something like 15
minutes by the time I've got the plate ready and everything, so I'm aiming for about quarter
past 7 now I think. So we'll see what happens, but it might actually be before that
depending on when the trout is ready. I'm going to turn this on, and I'm going to put it on
a very low heat to melt the butter. And I'm not quite sure how much butter to use, but I
think I'll use - well I can always add to it later on, so I think I'll start off with a small
amount. And I'll use the same knife again, I think. Its just got a bit of carrot on it but
that's OK. Ah, there's tin foil there, so in fact I could have put that over the potatoes and
then - I could still do it but I don't think I will (Would that make it any faster?) Well its just
that some people do do that with their roasts, they put foil over it because. Yes, in theory it
does make it a bit faster, and also then you take it off for the last 5 minutes, to make them
really crisp on the outside, but I do like them quite crispy, so I'm not that fussed that I
didn't, but it's generally more efficient with heat to use it. And I noticed it earlier on, but
when I saw it earlier on I thought I was going to have to roast the trout in it. But I knew I
was having fish, so I thought that I might have to roast the fish. So that's an oversight,
really, but then I wouldn't, if I was normally doing roast potatoes I probably wouldn't use
foil anyway. (??) I might have done, but I guess that you tend to do the way that you, you
do things now the way you did things when you were growing up, and I mean I don't
think my mother ever used foil for roasts. I'm going to turn the carrots right up now, well.
I think I'm going to need more butter than that. And I'm going to eat another almond. I've
got to make sure that the butter doesn't get too hot and burn, so I'm going to keep an eye
on it. (Is that something that you would be aware of anyway?) Well, it just caught my eye
[in the instructions] but I mean thinking about it, its something that I would do anyway,
and the carrots are begging to boil. And it's about a minute past 7, so I probably ought to
turn the carrots down again, having just turned them up. (Are they boiling?) They're not
boiling yet, no, but I want to slow them down now, because I haven't started cooking the
fish yet, and I don't think that's going to be enough butter, so I'm going to add a bit more.
I really don't know how much butter you need for trout because I haven't fried it before.

Appendix A
(It's not a very oily fish) Yes that's true, so it'll absorb quite a lot. OK, and I'm probably going to need the fish slice for this and I'm going to need a fork as well I should think. I'm going to look at my washing up again, in case there's any washing up I can do, but there isn't really, so I'm just going to stand here and look for a plate to serve it to myself on, and I'll use one of these plate I suppose, and I think I'm just going to sort of heat it up in hot water, later on, in a minute. So I'll put it there, and I'm not going to get a knife and fork ready yet, but where do I sit to eat it - where I like? I don't need the vegetable oil again, so I can put this back on there. Keeping an eye on the butter meanwhile. OK, so I don't need any of this, but I might need the salt though. I don't want any of that again, and I don't really need the wooden spoon again, and the butter is well and truly melted, and the carrots are sort of steaming, but they're not boiling yet, so I reckon I can put the fish in. And I'm going to turn it up a little bit, to just under 2, because I can't remember what the instructions said about erm how hot it has to be for the trout. (It doesn't) That's why I don't know - well I'm going to put it up to 2 and a half them and see what happens. And I'm going to look at the potatoes again, because its now about 5 past, with the intention of basting them again. (Do you think the potatoes are early, late or what?) I think the potatoes are on time, but everything else is late - well no, the trout is late. My main preoccupation in life is potatoes - ooh the potatoes look wonderful, but they need to do on the other side a bit. The potatoes look excellent, and I'm going to turn them down I think at this point, because I don't want them to do too much more, so I'll turn them down to about 350 and turn my attention to the fish. And my carrots are well and truly hard, so I haven't got too many worries about them doing too quickly. And I was beginning to worry a bit about that, and I'm not going to use this again, so I'll chuck it. I could in fact use the same plate again I suppose, I'll put that back. And the fish is beginning to fry, so I'm going to look at that again, and I'm going to turn it up to 3 now (Why did you choose to use the same plate again?) Well there's not much point in - well I suppose that one was dry, well why use 2 when you can use one. And I'm not going to put a lid on the carrots because I don't want them to speed up in cooking. The fish has been cooking now for about 2 minutes or so, so there's not an awful lot more I can do at the moment (??) Yes it sounds like its cooking, and its bubbling gently, and it smells nice, and the carrots are beginning to boil, so they are going to be ready way too early, so I'll turn them down (How long do you think carrots take?) Well, erm, well as I said, I like them reasonably hard, so I reckon that they only take about 5 minutes, and that means that I only probably will have just turned this over to cook on the other side, so I might in fact take them off the heat for a minute, I think. That was me being too eager - I'll eat another flaked almond while I'm waiting. (That's why you put them on the plate!) Plan ahead - you never know when you might get peckish. It's funny, when I'm cooking, and things, and I'm sure everybody is like it, and you've got to get things ready for the same time, I'm always aware of what the ideal is, and what you are doing in practice, and for one reason or another, either because, I don't know, like putting carrots on too early, you start feeling
quite stupid for doing things too soon. I mean I'd feel much worse I think, if I was cooking for a bunch of people, rather than cooking for myself, because then, time is more of the essence, er, just because you're keeping other people waiting other than yourself.

(Why do you think that you did put them on too early?) I think what happened was that I should have put it on a much lower heat, to heat up more slowly - why did I put them on too early - I don't know why I put them on too early. Its a major faux pas. Well, I don't know how to tell if this is done or not [the fish], so I'm going to poke it, and its quite tender, and its been on now for about 5 minutes, and it smells good, so I think I'll turn it over, taking its tail with it. I've got one or two problems here, eh! And it looks lovely, and it looks ever so nice, and I'm wondering if I'll need any more butter. I think I might, so I'm going to add some. And it's about 10 past 7, I think, well 11 minutes now, so the fish should be ready I would think in about 5 minutes or so, and I'm going to have to check the carrots again. And they're, the knife is going in reasonably easily, so, I'm going to try one - its very hot. Well they're not done enough yet, but they're only going to need another minute or two, so, I could just heat them up a bit - well they're going to cook like that actually. And I'm going to turn the trout down a bit, so it cooks a bit more gently than it is, but it seems to be cooking very well. And I'm just going to peek inside. And I'm going to have another quick look at the potatoes, and see if I need to baste them any more. They look very nice, they're coming on very well, and I'm going to stick a knife in I think. Oh yes! I'm going to put them on the other side a bit though, away from the oil, to try and dry them out a bit. And I'm going to get rid of this as well, this saucepan, I'm going to put it over here as well, with that. Just make sure that the fish isn't burning (Did you put some more butter in?) I did put a little bit more butter in, yes. You'll see it on your video later. And I'm going to look at what I'm supposed to be doing with the almonds again. Erm, (??) So you take it out of the pan and put it on a plate. Right. And I'm going to turn this down a bit, because it sounds like its doing very well. Right, so I don't really need to do too much more, apart from - well I can dry those up now I suppose, to make sure that they're ready. I would have thought really though that I could put the almonds in at the same time as the fish, I mean for the last stages - except I suppose that you can control - I suppose you can play with the butter a bit more can't you. And I am going to need that again, and the fish is done in fact. I'm going to check the carrots again, and turn them on just a little bit. And now I'm going to take the fish out - I'll use this plate I think, to erm, put the fish on - are these pyrex, can you stick them in the oven. So its virtually all ready actually, the only thing I want to do is chuck in those almonds. I'll give this a quick wipe because its got a bit of oil on it. Put the knife over here, I'll wash it actually. And put the plate down. I'll get my knife and fork ready. And erm I think I'm going to put the trout on a plate, and then fry the almonds, and then get the potatoes out, and do it like that. I'll just check again what you said about it, but you didn't say an awful lot. (At what stage are you going to take the carrots out?) I'm going to leave the carrots till last now, because they need like another minute aah. (??). No, they're not under done, they're ready, they are
ready, no they are - they're soft to the bite, but I don't want to overdo them because I hate them when they are soft. A carrot is a very dodgy thing you know. So the almonds are in. I can see I'm not going to live these carrots down. Right, and now I'm going to get the potatoes out. And I'm going to turn the oven off! And I'm not quite sure what you are supposed to do with these - I don't know whether you are supposed to wrap them in kitchen foil a bit so that you erm (what do you normally do?) I haven't made roast potatoes in so long I'm not sure. Yes, I'll dry them off a bit with a bit of this I think. But they'll be ever so hot. Don't they look good! And I'm just thinking how I'm going to wash that because of the hot fat and everything (Leave it on the side) Yes, OK. Meanwhile, my meal is getting cold. They're ready, so I'll stick them on top. And I'm glad I'm not cooking this for 6. I suppose you couldn't - you'd either have to stick it in the oven on foil. Mm, oh its really good. The last thing is the carrots which are boiling again, so they are nice and hot. Everything is off except that the red lights on here still, and I don't know why, and I want either a saucepan lid or a sieve. A colander will do. That's great. I'm going to stick some hot water into the saucepan and wash it up afterwards, stick the carrots on a plate. Loads of carrots, They're just right - I have to say that. No they are. And I've finished.
2. Task Instructions

The following is a reproduction of the three instruction sheets used in the study. The first is a general sheet, used for all subjects. The other two are specific to the particular meal to be cooked.

Thank you for agreeing to take part in this experiment. The instructions are on two sheets - you will be given the second sheet when you are ready to start. This will tell you exactly what it is you will cook. You should feel free to ask any questions.

General instructions:

You are asked to cook the following meal for yourself (see below). There are very few constraints on how you do this, but it is very important that you try to have all the parts ready to serve at the same time, and that the whole meal should be cooked in what you think is the shortest possible time - you should not have anything ready too early, nor should you be waiting for one last part at the end. You should try to be efficient rather than rushed. Of course, you may not succeed, but you should have this as your main aim. In addition to cooking the meal, you should use any spare moments to clear and wash-up as you go along.

Additional instructions:

In addition to all the above, I need you to talk about what you are doing all the time - in other words, I would like a running commentary. I am particularly interested in hearing what you are planning to do next and what you think you will be doing after that, and so on. You should try to talk all the time (if there are periods of silence, I will prompt you), saying whatever you are thinking about or whatever you notice about the thing you are doing - I know it’s not easy. It might help you to provide me with what I want if you try to talk in the future tense some of the time ("I will do ...", or "I am going to do ...."). If you also have a reason why you are deciding to do several things in a particular order, then say it. In order to have something to say all the time, some of the things you say might seem to be pretty trivial - this is OK - you should say everything that comes to mind, and describe everything that you do or are going to do.

If you have no questions and you are clear about where everything is and how it all works, then you can proceed. As soon as you get the second instruction sheet the experiment starts and you should start "thinking aloud".
The task.

You are asked to cook the following meal for one person (yourself):

Fish, baked with butter, new potatoes, mangetout and a cheese sauce.

These 4 things can be cooked as follows:

**Baked Fish.** - this will be either cod or haddock, depending on availability.

The piece of fish should be washed under a running tap, dried with the paper provided, and then parcelled up with a knob of butter in a piece of foil. This parcel should then be placed on a tray ready to go into the oven. It will take approximately 25 minutes to cook at 190C/375F. The oven will need preheating, which will take approximately 10-15mins. On these times, you can be fairly sure that the fish will be cooked, but you can reassure yourself by looking to see whether the fish is opaque (it should be), and whether it offers any resistance to a knife being pushed in (it shouldn't).

**New Potatoes.**

Take as many of these as you think you need for one person, wash/scrub or scrape them as you prefer, and boil them as you would normally.

**Mangetout.**

Despite the name, you can't quite eat everything, and it is usual to cut both ends off before cooking. They need to be put into boiling water and boiled for 5 mins.

**Cheese sauce.**

This should be made as a roux, using butter/oil and flour with milk, and then grated cheese. Aim to make as much or as little as you would like, but it should be smooth and lump free when finished, and then pan should not be difficult to clean.

Although you will be the one that will eat the meal (if you wish), I will be checking to see that it was satisfactorily cooked - this will include not burning pans or over cooking things as well as under cooking them.
The Task.

You are asked to cook the following meal for one person (yourself):

Trout, fried in butter and garnished with almonds, roast potatoes, and carrots.

These three things can be cooked as follows, although it should be noted that these instructions are only intended as a guide.

Trout and almonds

The fish should be washed in cold water and dried with the kitchen towels provided. It should then be coated in a mixture of flour and herbs, taking care to coat the inside as well. A teaspoon of herbs to a tablespoon of flour should be sufficient. The fish should then be shallow fried in butter or a mixture of oil and butter for approximately 6 minutes each side. When the fish is cooked, it can be removed from the pan and a few flaked almonds tossed in to cook in the remaining butter etc, and will serve as a garnish on the fish. These need no more than just browning, and should take a minute or less. Care should be taken not to let the butter get too hot and burn.

Roast potatoes.

So that this meal does not take too long to cook, these should be par-boiled before roasting. In other words, the potatoes should be boiled as normal until almost cooked, and then but on the baking tray, covered with oil and put into a hot oven to give the final crisp and brown appearance. You should aim to produce slightly brown looking rather than pale potatoes - this will mean that they will need basting several times in the roasting stage.

Carrots.

These should be peeled/scraped/washed as you prefer, chopped and boiled.

Although you will be the one that will eat the meal, I will be interested to see that it has been satisfactorily cooked - this will include not burning pans or overcooking things, as well as undercooking them.
Appendix B

This appendix contains additional information relevant to the study reported in Chapter 5. The first item is a full protocol, from subject 2. This has not been edited in any way except to underline the sort of statements which were taken to be of interest in the analysis. Line numbering is provided so that the reader can identify the position of excerpts in the main chapter. The use of brackets is as described in Chapter 4 (Table 4.2.), with the addition of "(??)" which represents a comment made by the observer which is unintelligible in the transcript.

1. Protocol of Subject 2.

Right, (??) how long's that last, that tape? (That lasts about an hour) Oh that's alright - I hope it doesn't take more than that. Right, I've just logged in there and done the retries job which we always run before stopping the machines. Right starting to er - check there's no users on which there isn't. (??) Well if there are any ... there are none on there. There's just this one chap on rim E. Right meanwhile I'll go and close the hub link. Its called the hub link and its stops people coming on and messages coming through on the FTP link onto the machines while we're doing the backup. (??) Well I'm waiting for that bloke to get off (have you told him to get off?) Yes. Right, he's gone so I'll stop the machine. I'll stop the hub as well (??). Tearing off the console logs as I go. Now I just have to put messages on the DCX. (Have you stopped the hub yet?) Yes, this is oh bollocks these aren't logged on so I'll have to log them on. Just checking the right message to put on to all the address groups. So when a user tries to log on he'll get that message now. (Happy with that?) Yes, while I'm here, I could switch the Benson on as well. Right, now the system is in reset so I can start running down the relevant disks that I need, that I'm going to need to run the large backups. So we're running 2 on the hub and another on here, on B, so we've got to take OUR2 off. Keep that off 'cos we're not backing that up today. I'll be backing up USE1 into this drive here and OUR1 onto that drive over there. So while I do that I'll get the disks out. We've got two sets of Monday disks and I happened to know that today will be the Monday1 disks. Well I remember last week. Wednesdays and Fridays there's just one set of disks, Mondays there's 2 sets of disks so you keep a disk for a fortnight. Right, so I'll just be putting this one onto here. We don't want that one, so we'll stick this one on here. Right, so we've got the USE2 disks on rim B. The USE1MON1 one on this drive, drive 7 on the hub, (??) yes, this is the target one on that one, this is the target one on that one. And this is coming from that one on the right. So now I want the OUR1 backup to go on there. (Which disks are you backing up this morning?) Which? All the four big ones to the Monday1 backups. So I will do this one [removing it from a drive] after the others are done. I'll just get the machines to start
coming up, and whilst I do that I'll write enable the relevant drives that are being copied to. That's for that one, and that's for the far one. (?) The machines are starting to come up so I type in all the relevant rubbish. This, when it comes up, it automatically mounts the disks, because it thinks it has the ones for normally running so I have to dismount them as it were, and put them in a suitable state to be backed up to. Just wait for this - it'll come up asking me for a warm in a minute. (?) It takes it's time. Well, (?) ..., right, I type in warm, (?) Yes, these are the backup disks that have been mounted, but they don't want to be in a mounted state. They just need to be in an online state. [mutters commands as types] right, so those two disks should be in the right state. When I log on now. (So you've set the hub up now?) Yes, the hub is the only machine that is up at present. So the disks should all be in the correct state now and write enabled now, to just start to run the er. Right, so I'll start one of them off on here. Right, I've started backing OUR1 up on there. I'll just log on on this terminal as well and start the USE1 up - and you can see that is actually working. That's how far its got (Are you logging on to the hub again?) I'm logging into the hub again on this terminal when it comes up, and then I'll run USE1 from this, using exactly the same command, because it should be in exactly the same state. Right that's going. Right now that's important - I've got the two big ones going so now there are several other things that I've got to do, not least of all running the small ones. Hang on, I'll just avoid mass (?) as well (eh?) oh this stuffs trying come on this FDP and we're not supposed to have in running while we're doing the backups because it tries to put stuff onto OUR1. These are messages, file transfer messages that are coming from other sites, and when we're not running we're supposed to stop that - that's what I did on that far console in the corner there (??) well it does, but this stuff still comes out, but I've avoided now (??) no, there was a message that came up on there that there was an FTP message that had come through. I mean it shouldn't come through cos I've avoided that and closed the hub link over there. Anyhow, I'll finish bringing up rim B while I'm here. So I have to log on on here (??) and the disks and that will need to be put on line. So I'm logging on and while I'm logging on I can online the disks [mutters commands as types] so I'm trying to online the disks as I'm logging in. (the long tedious sign on message?) Yes, which is even longer on here. (Anything on those screens?) [has just glanced around] No, that's still that FTP message, I'm just keeping an eye on them. I mean they should run for about a quarter of an hour. (You're just going to keep an eye on them occasionally?) Yes, in case a message occurs. Right, disk 3 is online, disk 4 didn't work. Right, the disks are online and I'm logged in and I can start using the same command again. I don't need 'in use' because they're both online. (??) Yes, third out of four. so. (are you going to wait for the fourth one until one has finished?) Yes, I'll wait for those other two to finish and then I can run the other one on the hub and theoretically when that's finished that's when the backups are finished because I should have done the small ones and done all that other stuff as well. I've done that and now I have to write up the disks of course, so we know where we are (the system log?) No, not the actual system log, Bob has done the system
log, no, the er disk log - what disks are on what drives. Right, so the big backups are going. Bob has done the logs, what else are we doing, er the small backups - you can see it up there on Monday's schedule - the 4 big ones and those 2 small ones. (You know this in your head anyway? You look at that just to check) Well I'll just check the small ones because erm. Right, I'm going to rip this one off. Err, so I need the rim E disks - so rim E is in reset, so I'll bring it up, taking that second disk off and run the backup onto that second drive there. (??) Yes, the backup of rim E's system disk. You see, rim E is one of the rims that has got 2, no 3, small disk drives, so you can do it. If it is the C one you can't because there's not a spare drive (??) No, you see I could take, with C, we take the system disk off and stick it on the hub, but as E has got a drive, because then I can run the hub system disk one on the hub. So I'll bring up the machine. I'll run down this [drive] - now the system will try to mount this, but it doesn't matter. This disk is not essential to the system running but users like to have it up when they're working on new systems and that so we leave it up all the time, mounted, so we put a command in on the autocommands when you bring E up that automatically mounts it. So I'll have to override that - it helps if both arms are functioning properly when you are trying to take these things off. Right, so just leave that on there, put the E backup disk on there. (??) yes, ... onto drive 2, yes, write enable it, using the little red switch under there. So I'll finish bringing this up, online the disk (how come this one did not come up with the startup screen?) because I haven't actually logged on yet. I brought it up, and I'll just log on now. You see there its just tried to mount the disk that its taken off - it says unknown, so I'll try and online the proper one now. The backup disks, if they are a backup of the system disks, which is the most important disk to bring the system up, the backup will be internally labelled exactly the same as that, so that if there is some balls-up with that, you can just take the backup and stick it straight on, and bring it up with that, whereas normal backup disks have their own name internally, so if you wanted to call them the same thing you'd have to change the name internally, which means having a system which is already up and running (you knew you had to come over here and ???) well, no, it automatically fails - it tried to mount it, but its got the wrong name on it internally. It hasn't worked (so you're now going to online the disk that you just put in this one?) Yes, that's right. Its called internally 'system' but if I put it up as that, there's already a system on zero so it would get a bit confused - so I'll put in an alias. Its called system internally, but I'm going call it BKSYSRE because that's what it says on the label outside - on the external (right) label on the box as it where.

Right, so my backup disk is now online, so I know the system disk ?? is in a mount state. For backup, you should either have both disks in an online state. Or if the one you are coming on has to stay in a mount state, which that one does, it won't run if you try and dismount it, there's a command called 'in use' which you tag on the end, and the system can access it still. The command for the small disks is slightly different, so I'm doing backup copydisk from system to bkysysre inuse, and in this one, for the small ones, it'll ask me the name for to and from - but if I'm being a bit sharp, I can just see that I can be a bit
sharp and just type them in in anticipation. That's from system, to BKSYSRE, right.

(Even though you typed that on the top line, you have to...) yes, its a sort of fail-safe thing I suppose. Meanwhile I know we've got to do the hub ones, now this disk was up for a job on Friday night - we don't have to do anything about that one today, we once a month back that one up to tape - that's taking stuff of those two little winchester fixed disks, so I know that that's finished really - even if that's failed we won't be doing anything with that yet (that's not your problem?) No. not at the moment. So meanwhile I get out the system backup disks. System hub Monday and system hub Friday we've got - we're using Monday's today, one may deduce, by the fact that today is Monday. Right, so I'll stick this on here so we've got two drives on here that we can use. While that's running up, I'll take that down - your not really supposed to leave large disks on the back of other drives.

Right, so I'll sling this one back in the shelf. Right, I want to log onto the hub again, but as I'm using both those (??) [checks state of the two console monitors] - we're using both of those, we'll have to log onto this thing here - this Decwriter here. Right, whilst I'm logging in, I also type the online the disk command, which again is internally labelled 'system' but internally is labelled 'syshmon', so I'll online that as well. So I'm logging in there. OK, make sure that its write enabled - you see that one is still write enabled from that job on Friday, so I'll put that one to write protect, and I'll write enable the relevant drive for the backup. (??) I'll also have write in the small disk log - you see I haven't got the source on, but I've got the syshmon - ooh, I nearly had to go and get another sheet of paper, which will have to be done next time - a popular job that everyone loves to do (??)

Yes, so I've got that one up there, and also on E, instead of sysgen disk, I've got bksysrme - so its just those two different - well three, but that's because we've taken that other drive down. Check the big ones again - nothing going there. Right, I've issued a command here, and its 'inuse' because its mounted, and I know its going to ask me the from disk, and its going to ask me from disk, (??) so I've pre-empted it myself. So I'll just do this query ASEND user and I'll see that it is copying. I can run that command on here as well again. To see how far - yes, it goes up to about 170 something on disk 4 - there's 4 logical disks on them (about a third of the way through?) No, well about three quarters the way - there's one two three and four, I think. Alright Jane? Jane is hard at it as well as you can see (with the plotter?) Yes, that's the Versatec test as you can see. Its a bit what? [its slipping] is it? Everything else alright in there is it? - I'm just liaising with my shift now. (??) Right, er, well you can check on any of the other tasks that they'll have been doing. Bob has done the tape drives. Jane has done the printers and now she has just done the Versatec test ?? only a reminder thing, when they start up the Opus's out there it produces a print test so you have to stick your laser writer card in there, so we get that off just to check it printing. Right, these are the backups we're doing today - OUR1MON1 [around this point, he is going through the operator's check sheet] (??) No, this is a reminder because we found that a few years ago that there's a few things that you are supposed to do that get forgotten about conveniently - sort of the more mundane things its
quite easy to forget. So its OUR1, USE1 and USE2, SYSHUB and SYSRIME. Well I've
got all them going. The disk changes book which you saw me doing, checking there's no
spoolers ?? we done that ??.. Bob has done the tape test - this little, er to test the drives, this
little procedure that the old girl wrote to test the drives - it sort of compares a file on each
drive (?) and then - the HASP link is our link to ULCC [still going through the check
sheet] so obviously we don't start them up until we've got the system back up and running
properly. These other things are things that will come later. The weekly tape we will have
to do of SYSHMON when the backup has run successfully. When we've brought the
system back up, we'll take a tape backup of that disk there (you'll do that when you've got
the computer backup and given to the users?) Yes, that's right. We'll run those, yes, with
the users on. Right now you can see that has finished on RIME (RIME is the small one is
it?) Yes, the small one we've got running over there. I don't want to do anything else on
this machine so I should - I mean the proper way of doing it is to offline the disk and put
them back, but I mean as I'm finished using the machine, all we're going to do after we've
done the backups we reload all the machines anyway so there's nothing else I want to do
on here, so I just halt it. Just halt the machine - I now take this - I don't have to put the
command in to offline the disk (were you expecting that to have finished or did you just
look at it and it had finished?) Well I didn't realise that it would have finished so quickly (it
was quite quick) Yes, they're only little, so obviously I have to make sure that the write
able button is off so we can't accidently overwrite SYSGEN when we put it back on
there. You got my message then yesterday? (yes). Right, I'll just put that up - I don't have
to mount it or online it or anything because when we bring the system up it will
automatically do that (??) Yes, I've finished with E now. Also there's a print out - when
the backup finishes, the print out comes out on the printers - we're supposed to check, but
it just says the message here, so the same as comes up here, so as long as we've got a hard
copy kept in the files. ... has come to see us as he often does. But he's a bit shy this
morning [??] Yes, everything seems alright, no air conditioning problems or anything
[there was a power failure the other day] was there? Right, go and see if there is any print
out - lots of print out on Mondays. (Programmer?) Yes, that's ... - she's just come in to
check. Well we haven't had any problems recently, but if there's any system problems we
tell her, so she comes in and sees. As its been the weekend there may have been problems
with the air conditioning or something, but recently its been alright. Erm, right, I'm just
trying to think now (what about?) What I've got to do and what has been done. I mean
Bob has - that's a popular job - taken the log - you see that printer thing there, yes, that
make a log of all connections made over the DCX - it piles out all this paper, and its a sort
of joke that its a job that everyone tries to forget about, tidying that up, but Bob has done it
(??) well also we put the date on the top of them so when you've got them in the draw, for
easy reference when you are looking through, 'cos the old girl goes through them and, I
mean we'll have a look through them to see if there was any cock-ups over the weekend or
anything, erm there were no evident ones when we came in - the machines were all up and

236
running - they can - they are supposed to reload themselves if there are any problems but
sometimes they don't. (??) Yes, that's finished now (??) yes, but I ?? to display the batch
jobs. Erm I thought it probably would have finished by now anyway, so I'll leave that up
though, even though its finished (??) Yes, that is a copy of the system disk, now I'll leave
that up because we're going to do a tape backup of that as soon as we've brought the
system back up. But of course I will write protect it, (??) These are just - oh there we go
those are just finishing now. This is the OUR1 finishing, the other one should finish very
soon, so I'll offline the OUR1 backup disk - its said its successful, but I'll check it over
there because sometimes they're not - yes it is offline, right, then what I'll do is I'm going
to put this one back on here, and when that one's finished I'll take the backup disk and put
the USE3 on there - OK they should be, yes those are finishing now. (Rather than start
backing up on the two that had finished, you would rather wait until the other two had
finished?) yes, otherwise you would have to take that one off as well. The system is sort
of devised to simplify it - if you start these two together, around the same time, they should
finish around the same time - barring getting any errors of course, which can throw you
out. So I'll offline the backup of USE1MON1 - the backup of USE1 (??) yes, that's right.
?? Right, so I've offline the disk - there we are, so I'll run this one down, and as they are
both successful backups - there's no problems with them. I'm doing it this way because
then you leave it back in as near a state to normal running (??) yes, I mean, that's right, it
simplifies - that's (??) Yes, well I mean the more you take them on and off - I mean they're
not made I think for serious on and off, you know, all day so the more you minimise that
the better, really. Right, so I want the USE3 backup from here. And I'm taking the USE1
backup off of drive 7 now. Its difficult doing it with your wrong hand. Right, so I'll stick
that back - now I still want that write enabled, so I'll leave that as it is. Jane, get some tea!
Is that the only job in the queue? (??) And what happened? Well cancel it and change the
priority to about 200. (What happened?) Well the plotters started as you can see and it says
it is active, but that job - there's nothing happening on it, its stuck, so if she cancels it and
changes the priority up a bit, because there's other jobs in the queue, and see if they start
running (Are you doing plots before you start the machine again?) Yes, because they'll be -
what you mean before we reload? Yes, because they'll be little jobs. When this last
backup finishes, I'll avoid the plotter and the printers - so that you only interrupt them, but
if there's little jobs, we should be alright. Right, what am I doing now? That's not going
on properly - I'd either forgotten or hadn't pressed it on properly. Meanwhile, if the
system was reloaded, that disk would be automatically onlined - well mounted in fact, but
as erm we've stopped the system and taken that disk off, that disk now need onlining on
drive 9, so I'll have to online [Jane interrupts with a problem] Is that the same job? [no, its
another one] right, hang on. So I'll online the backup disks as well, so they've come up,
so then I just start that, and I don't need the 'INUSE' because it only onlined (The final big
backup?) Yes, the final big backup - the two small ones have run successfully - oh I
haven't checked this one on B - its come up completed successfully, right. Then, as I
Appendix B

don't want B anymore. I'll just stop it, and then in a minute I'll just take those disks off 
and put them back to normal. I'll just have a look at this plotter now, and see what Jane 
has done to it. So it hasn't done any yet has it [sound of paper and things] Bloody thing, 
right, so that should be right, I mean the plotter looks alright, there should be stuff going, 
but there's not, it won't start up, so (?) erm, well I'll display the queue of things that are 
on it - you see all those are only little jobs. There's the one that was initially in it, but Jane 
has changed the priority on it up to 200, so that one of the others could run, but they're 
not. Is the other Benson set up for normal plots do you know? Oh, it doesn't matter, we'll 
have a look. Its setup for ?? is it, oh. Erm, right well this is - get a cup of tea. We'll 
probably leave it until after we've reloaded and see - that'll hopefully cure it (?) Well it 
don't know, I mean there's no discernible problem really, these jobs are only little, there's 
nothing. This is all the stuff that has been put in the system over the weekend (?) it 
shouldn't take too long - I mean there's a few large print jobs to go and then there's these L 
ones, which are the very big ones so when those are gone we'll start running those, erm 
there's these little plots that we saw here, and then there's the ink plots - I mean these 
should all take about 10 minutes when they all get going, these little plots, but... [hears 
teletype] That'll be I think - yes its the hub logging us out because - we were still logged in 
from when we ran the small backup, but its just logged us out now, (?) yes because 
nothing was done on it, in logon state. Erm, so I've stopped rim B (so the hub is currently 
running the big backup?) Yes, I've got that, that's the only thing apart from the plotters 
(hub - VDU over there?) Yes, that's logged onto the hub from over there. Those two 
VDUs and the ones in the other room are all logged onto the hub, so that one is running the 
bib backup - that's the only thing apart from the printers and supposedly the plotter that are 
running on the hub, so this last backup should go fairly quickly. So in the mean time, I've 
stopped B because we've run that one successfully and I'll put the disks back to normal -
well I'll take this one off and leave it ready to come - USE2 off of here, and leave it ready 
to come back on drive 7 on the hub, but I'll have to run them down first. (?) I'd gone over 
there without running them down of course, so run them down, also what I didn't do was I 
didn't do the large disk logs when I changed to start running the USE3 backup. (What 
reminded you to fill that in?) well just that I was just going to take them off and have to 
write them up and I just thought that I hadn't written up .. (So whenever you change a disk 
you have to write it up?) Yes, so that if there's a head crash or something that damages a 
disk, we can trace what other disks have been on there - I mean this happened the other 
month - a month or so ago we had a problem with this drive and the engineer said it was 
really knackered so we checked about the last three disks that had been on there, 
surprisingly enough they were alright, so I'll put this down as being about 8:40 since it 
was a little while ago. We've still got OUR1 (?) On here I've got USE3MON1 - I've 
taken those two off of there - or I'm about to, and I've got USE3 up on here, still for the 
backup, right so those 2 have run down now. Ah, that Benson has come up with an error 
which will probably explain why its not doing those plots. I'll have a look at that again in a
minute [changing a disk at this moment] (Right now you're going to get everything ready for the restart?) Right, so as soon as that's finished the last backup has finished, I can just whip this back on and bring the system up. I've got to remember to put this back on here, which I haven't written up in the book, I've just remembered - I've just remembered that I've forgotten to put this in as going back up on here. We've only had this on for the past couple of months and it's still not ground into the system that its automatically there. Right, just getting this one off now. So as, when the other backup's finished all I've got to do is I'll stop the system, take that off, and put that on there, and then they're all ready back as normal. So I've finished with B, so I don't need the write enable. Stick that up there, check the console, (as you are going past?) Yes, that's about three quarters of the way through. I've put OUR2 up there [mumbles latter to self as if writing in the log] Going back to that Benson error, if I do a stop on it, I'm just stopping and resetting it. PNW3. Hopefully this should sort it out. Bollocks, its come up with another one straight away, I wonder what that is? It may be the user's job, because there's no - it should start to do something (??) No, I stopped and reset the process because it was a process error that it came up with, but as soon as I started it, it came up with it again. Ah, it seems alright now, maybe it just needed powering on and off. (You switched in off and back on again and it came back up) Yes, however, because we're getting near the end of that backup I'll avoid it so that er, otherwise, because as soon as that's finished I want to stop the system and erm (You've avoided the printers - but it will finish the job its doing?) Yes, I've avoided the Benson - by avoiding it, it finishes the end of the job. If you cancel it, it stops immediately, but if you avoid it. So I'll display the other writers and see what ones of those are going - the printers will probably be going (Do you need to avoid the ones you ?? the system back on?) Yes, you see if I'm in the middle of a large - oh shit, its just started that I think. It would be a pain and a waste of paper and everything to have to start it up again, so they're all avoided. I'll probably go in there and find that there's one of them got an error and is just sitting there anyway. That's nearly finished, but that one's got quite a long way to go though. Right, well I'm hanging about doing that I'll get out the er, this is something that we do on Monday mornings (So what's the current status?) the current status is that we're just - oh, its finished now, [the backup] no, I'm not going to wait, that's only about a sixth of the way through. I'm going to cancel that because it will start up again. (What was going to stop you restarting - the fact that you had that long printout that was only part way through?) Yes, otherwise I'd have to wait about 10 minutes for that to finish, which is a bit of a pain (It was the fact that the copy had finished and its more important to start the system up now) yes, because its just gone 9 o'clock and there's a user out there, waiting to get on - keen! - and there's probably hundreds more throughout the college just itching (Turns tape over) [somebody requests a print out] Well hang on, give us 5 minutes because I'm just about to reload [five minutes!] - if I started it now, it would get interrupted wouldn't it! Right so that's finished, and I could wait for the printers, and I could go in and run Bill's job, but I'm not going to, I'm going to stop the
system and put that disk back to normal and just bring the system up so that Bill can go and
put in a load more jobs for us to print off. [banter] Right so I'll take this USE3MON1 disk
off of here, and I'll be replacing it with USE2 [banter] Right, take this out of here, [banter]
Right, I'm putting the write enable switches off. This disk hasn't run up - this OUR2 disk
on B, and I'll stick this one away while I'm at it, and all the disks are back in position,
ready to start the system up for the new day's computing - high powered. Bloody hell,
he's here again - must be another lecture again [more banter] Well I'm starting them all up
- I mean you have to bring the hub up first, erm but I'm going to start bringing these up
simultaneously (in parallel, all together) although I've only got one pair of hands - its all the
same - its May isn't it, not April - its all the same commands, date and time (???) right, I
keep thinking its April still, (??) right, now the hub is ready to put in the warm, so that's
now up, erm, C doesn't seem to be doing anything, erm. Yes, C seems to be taking its
time in coming up - all the others are ready, I've done the warm on all of them except E and
C now, E will be coming in just now, so I'll put the warm in on E, but I'm still waiting for
something to happen on C. Meanwhile, while I'm waiting for that, I can take the message
of the switch here, and I closed the hub, so now I'll enable the hub, and while I'm here,
I'll tear the log off, finally. Screw it up in a ball, and throw it in the bin. [a man interrupts
again] Right, Eric wants his HLP - he's got his important management report. Right, start
that up - I mean this is a break with the normal scheme of things, I'll log on, we've got this
macro which starts up all the peripherals together, so I'm going to run that, its called
devices (would you have run it anyway) Yes, but I would probably have run it from in
there though because it would just start the printers up - we've got to wait for all this crap
to come up on the screen again as usual, about these evaluations and all this nonsense that
we've got going here, erm, this will start up - the printers will get going again, and then -
what wheel is it you want Eric - [??] its probably on there anyway. Oh it is on there - do
you want a film? [??] right, I'll check the queue to see where Eric's job is, come on, right,
I'll change that over to class K

and so the day continued...
2. Task Instructions
The following is a reproduction of the instruction sheet used in the study.

I would like you to give a running commentary whilst you are going about your job. You should try to talk all the time (if there are periods of silence, I will prompt you), saying whatever you are thinking about or whatever you notice about the thing you are doing - I know it's not easy. In order to have something to say all the time, some of the things you say might seem to be pretty trivial - this is OK - you should say everything that comes to mind, and describe everything that you do or are going to do. I would rather you said too much than too little.

It might help you to provide me with what I want if you try to talk in the future tense some of the time ("I will do ..."); or "I am going to do ..."). If you also have a reason why you are deciding to do several things in a particular order, then say it. I am particularly interested in hearing what you are planning to do next and what you think you will be doing after that, and so on. Other than that, I am interested in why you might do something now rather than later, and equally why you might decide not to do something immediately but put it off until later.

Any questions??

Thanks,

Miles.

P.S. Try to remember to wear a shirt with a pocket for the tape recorder!
Appendix C

This appendix contains additional information relevant to the study reported in Chapter 6. The first item is a full protocol, from subject observed on the Wednesday. This has not been edited in any way except to underline the sort of statements which were taken to be of interest in the analysis. Line numbering is provided so that the reader can identify the position of excerpts in the main chapter. The use of brackets is as described in Chapter 4 (Table 4.2.), with the addition of “(??)” which represents a comment made by the observer which is unintelligible in the transcript.

1. Protocol of Wednesday's Subject.

(All is quiet?) At the moment. I've only pulled so far for that Buxton, because if he doesn't move (...) that's the one on platform 4, I can't get across you see with 1M32 to 14. (You'd rather have 1M32 onto 14 than this one out?) Well, he's due out now - he should have been out at 7 minutes, you see, that fella's not due here until 15, but, you know - they should have been off for several minutes and then they'll come on and say oh, we've got no guard or we've got no driver or there is something wrong with the set (So what you thought is that there is something wrong with that..) We've got to er... It doesn't do to pull off right through with a train - it doesn't do to complete a move, let's put it that way. [PHONE rings] "Hello,... ". I never tend to pull off for any move until the train starts moving (because if he is late, he could be very late?) Well, yes, you get let down too many times - and once you have committed yourself you see, its too late to go back (If you commit a route, if you set a route up, can you undo it?) Oh, you can do, but it takes a couple of minutes to time a route out, once you have pulled the signals back you're waiting 2 or 3 minutes for the route to drop out, and you can't do that until you have informed somebody on the station that you are doing it - you can't just pull it up willy nilly. (So your best policy is only to commit so far?) That's what I do. (Different people do it in different ways?) Well, yes. I don't believe in committing myself, I've been let down too many times. (So you've just set this 1E62 off on 3) Well I've pulled off for him because he's indicated out of the station, plus the fact that there is nothing to go that way, so if he is late, he is not going to stop anything. (Brilliant) But, coming across from where he is, 1M32, across the junction, across all 4 running line (the guy on 14?) Yes, so it doesn't pay to be too eager. (??) I've got one coming across the top - that 2F89 - he's due to follow that in you see. (So you're holding 1M32 at 315 signal?) until he's clocked the time - he's early at the moment. Up the slow Tim. He's not due on the station - on platform 14, until 16 minutes past (That's 4 minutes yet) That's plenty of time yet - they always tend to be a bit early getting there, because they get recovery time from Stalybridge. (2H80 will be out of the way by then?) Oh, yes, he's on the way now. (That route clears behind him quite
quickly doesn't it?) It does. He'll be running on greens you see, so he'll be moving now. (The station looks very quiet) It is this time of day unless something goes a bit haywire - it could happen couldn't it? (That LELS is a light engine for Longsight?) Yes, that's it, Light Engine for Longsight (It's been there every day this week - same time, same place) Well they're preprogrammed you see, the movements of these trains, these locomotives are all preprogrammed, although sometimes they will alter them, if there's a shortage of locomotives, or there's a fault on one. We can come across now with 32 you see - its only 14.13 now and there's plenty of time - he's only got to get across to 14, he's got a couple of minutes in hand now. (He's following straight on the back of 2F89) He'll follow right behind 2F89 (Which is just pulling into 14) Yes. (And he is being followed by 2A63) yes, that comes on the Eastern side and that comes onto platform 2 (So rather than going over the top, he's staying along the bottom side here?) Yes, that's right. (Which is a simpler manoeuvre, except you've got 1E62 in the way at the moment?) Well he should be out - but he's not, 2H63 is not due in until 21 minutes past, which is another 7 minutes, (But this 1E62 has not moved yet has he?) no, he's not out until 15. (so you may be lucky) Well I hope I'll be lucky, let's put it that way.

I've pulled off 7 now - I've pulled right through with that one (The 1G61?) Yes, up the fast for that one please. And I'm pulling off now to the slow for the express. Up the slow please. Again, both those moves aren't conflicting with anything that is likely to happen so there is no reason why they shouldn't be given green signals right through to.... (So you've passed it right the way along to the Longsight man?) Yes, (Longsight: Someone xxx on 13) 1L70? (Longsight: Yes) Oh well there you are you see! This is what happens. I've gone and pulled off all the way up there, and they tell us now there's no driver - so that's er now blocking platform 13 (You cleared him all the way along to Longsight - have you undone that route now?) Yes, I've pulled the signals up on that route now. So we can't get of 10 and 11 now (Background - there's no driver at Stockport - they don't want him to leave here until they've got a driver) Right, so he's to stand there? (??) Oh, OK. Did you get that? The train is to stand here until they've got a driver at Stockport to give him a relief. (Background: - Bloody Hell) Exactly, my sentiments exactly. This is the sort of thing which takes place you see. You see we've got no access - with that train standing now, we've got no access to 10 and 11, he's blocking the route. (So you've cancelled his route, but it hasn't timed out yet?) No. It'll take 2 to 3 minutes to time out. (Is there no platform 12 at Piccadilly?) No, they did have, but its gone. (Background: So that means that you have to work up round 14 now - so you've just got to juggle them a bit). So this is where the messing about starts now - because we've got to go both directions through platform 14, but obviously they can't all go (Background: so its just a case of fitting them in when you can) Its fitting them in "We're stuck on platform 13". [Discussion about Christmas cards] That route has dropped out now, that 382 - it doesn't mean anything really, we still can't get on platform 10 or 11. We're alright now off the main station,
68 whereas with that route up we couldn’t come out of ?? ... - I’ll just have to play it by ear
69 now, I won’t be able to plan anything up, it’s just going to be one through each way most
70 of the time now - I’ve just got to play it by ear sort of thing, you know. You can’t plan
71 your routes when it’s like this, because it depends when the trains come and how long they
take when they’re in there (Otherwise you’d be planning) Well, yes, (Background: No,
you’ve just got to weigh up what you’ve got at Oxford Rd and what’s coming down and)
72 You try and keep em as near as possible to time on their proper path, but it doesn’t always
work out that way - you’ve got to keep a route open for a class one more so that a class 2.
(Background: Oxford Rd?) That’s what you’ve got to think of - if he gets stuck there and
73 we end up getting full down here, it’s moving them - we’re going to be stood here for a
xxxx driver. He’s not away from you is he until 29? (That’s 2H29 coming from
74 Longsight?) Yes. (You’ve cleared him through first?) He’s going through first, (the next
75 one in the other direction will be this 1M32 will it?) No, it’ll be 1M65 - but he doesn’t leave
76 Oxford Rd him until 14.24 (3 minutes) there’s no reason why he shouldn’t have a path
77 clear by then. What have we got coming down? 2H15 (and he is going to what?) 7,
78 platform 7.
79 Got a Hadfield coming off the Eastern side (2H77) we generally pull right off for them
80 (Where is he going in to?) He’s going into number three. (And 2H63 is going into number
81 2?) That’s correct (and 2H77 is following him but going into number 3?) That’s it. (Now,
you’ve set that 2H77 up to 321 signal) Yes (Which is the one immediately behind where he
82 is cleared) That’s right. As soon as the 2H63 is out of the way, I’ll pull off right through
83 (And he’s clearing that bit now isn’t he).
84 (1M65 is in Oxford Rd now isn’t it?) It is (and 2H29 is nearly through) uh hu,
85 (Background) Oh, that’s handy - we’re going on 13 (That’s cleared the
86 block?) It has. There you are then, up the slow. (So that means you have set 1M65 to
87 follow on to 13 where it ought to be?) Yes, it can go on its proper platform now (So that’s
taken your worry away?) It has, its reduced the workload considerably, because when you
88 get a platform block like that on the junction it can be a pain, especially during the peak
89 hours - these fellas will tell you. If you get one fail or something like that, you are in
trouble. (These are your only 2 through lines, aren’t they?) That’s all. And you’ve got a
90 train through there every 2 minutes in the next hour or so, and obviously they are going to
91 build up. I’ll alter this number now for 7.
92 And that one wants altering - 2H63 ... (Background: - no track machines working this
93 weekend) Does that mean we’ve got a nice quiet Sunday? (Background: No, there’s
94 nothing on this weekend at all) I wouldn’t think so, it being Christmas Eve (Background:
95 Nothing at all).
96 Hadfield is coming into number 3 now (Very smooth - Hadfield is the one that you set up
97 from Ashburys a couple of minutes ago?) That’s the one. Normally, on the Eastern side,
you can, you know, complete the move, and pull them right the way through onto the
station (Only when you have to cross them to 13 and 14) Yes, they’re the ones you’ve got
to watch. This Eastern side is notorious for trains failing, or guards missing and stuff like
that (Background: - more than the rest of the station put together) (Oh) And somehow, it
always happens to me - if I commit meself, summat will always go wrong (Which is why
you’re learned not to commit?) (Its very quiet up there at the moment) It is at the moment,
yes. (Have you got anything in mind for the distant future?) The next one out is, erm the
14.30 - we’ve got 2 out at 14.30, we’ve got an a Cardiff off number 4 and a Rose Hill of
number 2. At the moment there’s nothing (That 2H54 is the Rose Hill is it?) Yes. (That
will go straight down the Eastern side again?) It will, straight down the Eastern side. (The
one on number 4 is the 1V14?) Yes - he’s going via Stockport, straight up the fast (So they
won’t conflict?) No. We’ve got one to come in, that 2H01 up there - he’s running late
(his’ not quite made Slade Lane yet) No, and he’s supposed to go in on platform 7 - but I
can’t put him on platform 7 because he’ll block the 2K10 in you see. So what I’ll have to
do here ... (In other words, he should have gone in before the 2H10, but because he is
late) He should have done - so what I’ll have to do now is put him on platform 6 and start
him from there you see - that train on platform 6 isn’t out until 20 past 3 - and he is only
going to be in 10 minutes. (Do you always try and put things on a pair of platforms so that
passengers aren’t inconvenienced) Yes, if possible - if you can’t get him on his own
platform, try and stick him on the other side of the island, then the passengers are on the
proper platform and they have only got to go from one side of the island to another. (It
would be a bit mean to swap from 1 to 10) Yes, you’d have passengers running around all
over the place, and apart from anything else, its causing delay to the trains, if you’ve got to
start shifting passengers round. So I’ll put him on number 6. [SHOUTS: “14.37 to
Buxton will go from number 6”] (That was the announcer asking you - but you would
have told her anyway in a couple of minutes?) I would have told her, yes. Now they’re
indicating out on 4, so we’ll pull off for the Cardiff (And that’s straight up the fast?)
Straight up the fast, and we’ll pull off for him because he ??? And we’ll pull off number
2, and if he doesn’t go, he’s not going to stop anything. (So you are safe making that
move) We’re safe making that move (You can commit and nobody cares). Up the fast for
the Buxton Tim. I’m knackered again for the moment until that takes place (He’s pulled
them both out, and 1M65 coming in) He’s gone - he came from Oxford Rd (Oh right
through 13).

This one’s late, this 2H22 going to platform 13 is er something like 7 minutes late.
(Longsight (Tim): Here you are, down the slow for Trafford Park). (So there’s a lot of
routes set up, but not a lot to worry about) Not at the moment, we’ve just got to let them
come and go as they do now - once you’ve set the route up for them all, you’ve just got to
leave them to go (Right, so your 2H01 has pulled into 6 - that’s the Buxton going out?)
that’ll be going out again at 14.37. (The 2K10 is going to go before then?) He’s due out at
14.33 - (Right, so you’ll give him a route soon?) Yes, I’ll give him a route soon.
(anything else?) No, that’s it at the moment, once that Buxton is in, I’ll pull off for the one
off number 7, but you can’t do the two moves at once you see, because they are coming
through the same section of line (2369) Yes, into both platforms you see. 2369, that’s a
point number. Now he’s safely in, we can set up the route for him - oh no we can’t, we
haven’t come up with the Cardiff have we. (The Cardiff is the 1V14 on 4?) That’s him,
he’s coming out now (You thought he had gone?) I did actually, I wasn’t taking a lot of
interest. There is another way out - I could have come out over the down fast, but erm,
(You’d rather go through the up fast?) I’d sooner keep it straightforward where possible
(where would you cross him over - which points?) I’d come out through 2369 through
2367 points (And you’d bring him down there) Yes, onto the up fast. He’s not due out
until 33 (You’ve got up fast and down fast, and you can use them in both directions, but
you’d rather keep it by the book) I prefer to keep it straightforward, where possible.
(You’d rather go through the up fast?) I’d sooner keep it straightforward where possible
anyway when it’s not necessary (So brought it across) You see, some of them will use
the alternative routes - when you’re stuck you’ve no alternative have you, but when there’s
plenty of time, there’s no need.

Fast line for Stoke please. (Very quiet down here on the Eastern side) It is at the moment -
the next one out is off number 3, which is out in 2 minutes (And he’ll follow that 2H54,
which is clearing now) We can pull off for him anyway, and forget about him then. Right,
they’ve indicated that one out on platform 7, and we’ve pulled off and its 14.34 and its still
not moved - and its due out at 33 (So you have committed a line and you may be caught
again) I may be caught again, because that Buxton is out of number 6 at 37 and he can’t get
out until the other fella’s gone (So you’ve got 3 minutes) It looks like he’s moving now -
the departure end of that platform, the track has shot up, so it looks like he could be on the
move (here he goes) yes. I’ll pull off up the slow for the Style line train. Up the slow Paul
for Style (That’s the one off 13) That’s it. So the next one out now is off number 6 (That’s
the Buxton service) Yes (That comes in as 01 and goes out as 00) Yes, (always?) No, they
come in with different number, but always drop down a number - they come in with odd
numbers and go out with even - the next one might come in as 2H09 and he’ll go out as
2H08, or if he comes in as 07, he’ll go out as 06. That’s usually the pattern that they
follow. Then you’ll get one that’s slightly different to kid you you see. (So you have to
keep your eye on the book?) You’ve got to keep your eye on the book. (Are you just
scanning over it, making sure its alright?) Everything is going according to plan - there you
are, he is indicating out on platform 6. You’ve got to keep scanning. Fast Line Buxton.

The next one now is off number 1. (The Buxton has got a line, but it hasn’t moved yet)
No, its not due out until 37 and its not quite that yet. Some of them will come out pretty
well to time, others are always a minute or two down, for various reasons. (I suppose they
can sometimes make that minute up) Well, his first stop is Stockport you see, so .. (If he
gets a good run he might make that) Oh aye, I can’t see why he shouldn’t (So long as its
Appendix C

not early) Well they shouldn’t leave early - its up to the platform staff to ensure a right time
departure you see. (You generally give them a clear aspect a couple of minutes in advance
don’t you) normally, you’re supposed to have your signals clear a couple of minutes before
the train is due to depart, but the platform staff should indicate them out you see.

How have you found it so far, interesting? (Very interesting... There goes the Buxton train
off 6 - that’s now showing LELS - so that actually had three different items on there?) Yes,
there was three on there, there was a locomotive on the buffers, the train itself, which
forms the 15.18 Birmingham, and the Buxton train went on top (The Birmingham is the
V97) That’s right - he hasn’t got an engine on it yet - that’ll come up eventually from
Longsight.

[talks to someone saying ‘anywhere you like’] (This 0V20 you’re talking about, which
platform is that going to go to?) That’s going to go to platform 10 - to V20 on platform 10.
(So the V20 and the 0V20 match do they?) Yes, that’s the locomotive for the train that’s on
10 - its a train of vans, parcel vans (Going what, London way?) Its er Bristol. The one
behind it, G69 is Birmingham.

(Why did you particularly say that he could send it anywhere he liked?) Well, he could put
it down the slow line or down the fast - I can get into platform 10 from either of those
routes you see (you don’t think there is anything going to come out to conflict with that?)
There’s nothing to come out at all - nothing to come out to conflict with that move at all, the
only thing thats going to er ?? The only thing that will stop me getting onto platform 10
from either of those routes is coming off platform 13. So what I’ll do with this is I’ll bring
the engine straight down the main to the station throat, as we call it, and as soon as that
2D43 has gone from platform 13, that engine can go in on platform 10. (That 2D43 is
sitting there at the moment) Its sitting there at the moment - its coming out this way - its
going to Chester via Stockport. (I’ve still not got the directions right - now it’s moving
nicely) Up the slow for you again for Chester. There’s the engine showing up again (So
the 2D43 is the Chester - that you have just passed on to the Longsight man, and that has
started to move now) He has probably had to crawl into platform 13 and come to a stand at
the end of the platform - you know where they have extended them with the new extension
(13a?) Yes, that’s what they call 13a. (Longsight man ???) Yes, I’ll have him to 303. (This
is 0V20) - no that’s 0V97 coming out now - to 303 (And where’s he going then) He’s
going on platform 6 (To V97) uh hu (and he’s going out again) Yes, the 15.18
Birmingham. (And you can only bring him in that far because you have put 0V20 to that
crossing point?) I can’t go any further at the moment with him. (And now you have also
put the 2M45 in there) That’s coming round now, and when I can find it on here, that’s er
going on platform 2. 2M45. And there’s nothing to come out of the Eastern side, at all, all
those trains are - the first one there is not due out now until 10 past 3. (That’s why you’ve

247
Appendix C

got times on them) Yes, when they’re going to be stood there for any length of time, I always put the times up to remind us what they are - because if you put the reporting number up, you’re looking all the time ‘what’s that, what’s that’, but if you put the times in, (So, the guy from Longsight gets preference, then you are going to bring in the er 0V97, but you can’t do that until the 0V97 has cleared) No, well I can do, I can come on the other route because I want to come onto the main station, but I wouldn’t clear the route until he has come off Longsight depot (0V20 is moving in now -) Yes, he’s going into platform 10 now. I’ve pulled off right through now for 2 for that 2M45, incoming Sheffield. (??) You see that engine 0V97 has still not come from the depot yet - if I’d have pulled off I’d have had all those three running lines blocked. (Right - so you set him up to 303) That’s why... (You committed far enough for him to start moving or he never would?) No, I’ve only given him one signal, to bring him out. (And will he know how far he has got?) Oh yes, he knows, he’ll be a poor driver if he doesn’t. (0V20 has got in there now) He’s just going in platform 10 now.

Where’s the express Eric? (Eric: I’ve no idea) (Longsight: Top line for you) (Have you and Eric lost an express?) Yes, there should be an express up from Liverpool, which should follow that Chester that came up - you know that 2D43 (which came through 13) It should come through platform 13 at 14.45 - 1E63, a Liverpool-Scarborough, and its not even appeared on the scene (Not even at the far end?) No, so at the moment we’re missing a train. (So he’s stood at 303 now, the 97) So we can bring him down this way (He saw it first so he reminded you?) uh hm. Now he can go in there like so you see. Straight onto his train. (And that’s the Birmingham coming out again?) That’ll be 1V97 coming out there in a minute. (2M45 has just come in) He’s just arrived over there now - this Buxton coming down is for number 7. But we can’t get in at the moment because we’ve already set for the engine to go into number 6. (How come there’s a line cleared through..) Those signals are on automatic - we leave them like that - those signal can be left on automatic, so as soon as the trains clear the overlap of the one in advance, the one in the rear will clear again you see. (So that means that another train coming up will have that line?) Got another one coming round there now. (Eric: its not left Liverpool) Oh. (Its supposed to have arrived by now?) Yes, well he was ... in another 5 minutes there’s another one out of Liverpool - 52 past the hour they are. Be a waste of time running it. (Maybe they’re short of guards as well?) Well, you never know.

You never know. You always keep the signals clear, through the junction, you know, 13 and 14, to keep them moving otherwise they build up. (So V97 has just pulled in - oh he’s not quite pulled in yet) No, he approaching the station. (His bay has just changed - that means you can now get the 2H03) Yes, he’s got to be right in on the platform though, behind 336 signal (2H03 is a Buxton again) Buxton again (There’s a lot of Buxton trains) every half hour (They seem to be more frequent) well they are, ever since last May - they
were an hourly service up until last May (Now they run these new sprinters don’t they)

That’s it, yes. You see, he’s not due in until 50, so he’s not doing so bad, the Buxton,

he’s only going to be a few seconds late. 1M34 - we can pull right off for him - he comes

in to platform 3. (That’ll pull neatly into there) Yes, that’ll pull neatly into there (He’s just

cleared behind this 2H55 into 1) uh hu. There’s two ways of going into platform 1. You

can go in the way that I have gone - what they call the top route, or you can go in the

bottom way, which is through the crossover road 2352s and straight in that way (So you

cross over late? - when you could have crossed early) Yes. (In fact you crossed the other
guy at that early point) uh hu. Right, we’ll alter a few numbers. 2M45 becomes 2E48.

You can’t find the damn things half the time there’s that many. Oh there it is. [keys in

numbers + mutters]. I’ll carry on up the slow now for that Crewe - up the slow for the

Crewe please. Got a Hadfield coming. (The Crewe came out of what?) Out of platform 13

- see that 2K28, that’s the Crewe. (1M34 coming into 3 where you set up before) nothing

else at the moment now - the next train due out in not until 1500, that’s the London off 9,

and the other fella of number 1 I think it is yes - the New Mills off number 1. (Not much

happening?) Not much happening at all. PHONE rings “Hello - right OK, well the Buxton

is on there now and he’s there until 15.07, but your train’s not in yet anyway - follow it up

when he comes, yes - OK” (That was a guy on 6/7 for what?) That’s the engine on the

block of 7 - see 0V62? (Yes) the driver is ready to bring the engine up now to back onto the

train, well the train is not in until 10 past three, and he can’t get out anyway cos that

Buxton is there until 7 minutes past, so as soon as the Buxton goes he’ll follow it straight

up and he’ll go into the sidings then you see. (Then the train will come into 7 again) That’s

right, the train will come into 7 again, and he can drop straight onto it. Its due in at 15.09

(1H63) That’s it (The one you are going to put the 0V62 on?) Yes - that becomes 1V62, its

return working. (Right, I’m getting the hang of this) It takes a while. The Hadfield

coming in 2H79 goes straight on top of that thing on 3. Now we always complete the

move for that because there is nothing else to stop it (That’s the Eastern side again) Yes.

(On the Eastern side you do a lot of completing don’t you?) You can do more or less on

that side - the only thing you have got to watch is for anything on platform 4 - the way that

I have come into platform 3 now, I can’t get out of platform 4 (But if you had crossed over

at the earlier signal) I would have been able to. That thing on 4 is there until 10 past 3 I

think, so there’s no panic for that. (That’s Eric passing you something else from Oxford

Rd - 2H94) uh hu, that’s another one for Stockport (and it will follow 2K28) Yes (Will that

stop at 13?) Yes, they all stop in Piccadilly. There’s only one or two services that run

straight through Piccadilly - I think there’s only like 2 in a 24 hour period - I can’t

remember what they are, I never can - can you remember what they are? - the ones what

run straight through Piccadilly, I’m sure I can’t remember - (Eric - we have a Blackpool-

Nottingham, and a Nottingham-Blackpool - and then you have the Edinburgh-Paddington)

It runs straight through (Is that the middle of the night?) Its getting on - its not the middle of

the night, its late on at night. (I could understand if it was the middle of then night. Is
there anything in there that sort of strikes you?) Nothing at the moment - everything at the moment is going according to plan (You use that same book everyday do you?) Oh yes, just keep the one book - these go from midnight to midnight, Saturday excepted ?? the service. (Longsight?: There’s your Cardiff) OK. If there’s any alterations, then they notify you - they give you a sheet with any alterations on it (Sometimes you write in these books) Yes, I write in them - well they all do - so that you know that when a train comes in you know what an engine does if its bringing a train in, and you know what the train forms when its on its return journey out again you see, otherwise if you didn’t do that, you’d have to start looking through the sheets to find out it does, and you haven’t got time - especially at the peaks, you haven’t got time to be sorting through, messing about like that. There’s your Stockport up the slow. That one on 3 we’ll alter his number now. Thats out again at quarter past three (thats the 1E64) mm. That’s back to Hull - its an hourly service to Hull (its a class 1 as well) - they’re only sprinters (So some sprinters are class 2 and others are class 1?) Yes (It doesn’t go on what the engine is like) Oh no, it doesn’t go on the train at all it depends on the service. There’s the light flashing on platform 9 for the London. So we’ll pull off for there. Up the fast for the London Paul. They have a I think its a half hourly service to Blackpool here (Doesn’t that go through Stockport now?) Yes, that’s coming through from Stockport. (Eric: That Cardiff said ‘give me a run and I’ll go’) Did he begod - well we’ll see what we can do for him. (You’ve just changed platform 3?) Yes, that’s the Hadfield come in, 2H77, 79 rather (Its working that number out [74]) Yes, its going back to Hadfield (Do they always go back where they came from?) More or less - Hadfields always do their own, and Buxtons do their own return working, as do the Blackpool Buxtons, er Blackpool-Stockport I should say. The London stock more or less stays in the same circuit. PHONE: (platform 10) ‘Yes Please’ That’s the driver on the engine on 9, that OA13, wants to follow the train up, which he can do (So when 1A59 goes out, he’ll follow it) Yes, he’ll follow it, nip it into the sidings. (And he’ll back up on - you haven’t got it yet?) No, we haven’t got it yet - it looks like his train is late as well, I think 1A13 (?1H13) is the 4 o’clock Euston - it is, and that train should have been in at half past 2, and that train has gone astray as well. So we’re missing an incoming London 1H10, he should have been in at 14.32 and there’s no sign of him yet (You hadn’t noticed that yet?) To be quite honest, no. I don’t worry about them until they hit the panel (Right - do some people?) Some may (do some go more on the book that on the panel?) some may. It depends you see - what you’ve got to be careful of you see is when you put an engine in the engine sidings, you don’t trap it in - that’s what you’ve got to be careful of. It happens when you’re ... with a locomotive, and you’ve got one in the engine sidings and you put one on top of it, and er, you’ve trapped it in. So in this case, what we want now - we’ve got 2 engines to come up now (The 62, and the 13) Yes. Up the slow Paul. So we want the engine off 9 in the engine sidings before the engine off 7 because the engine off 7 is going out. If you do it the other way round, you’re knackered. You can still get them out, but it means that they have to drop out the other end, and come out the other end. Its a lot
of messing about, and to do that, you’ve got to contact the driver, and you can’t always do
that. (On the one hand you just worry about what you can see up there - if you can’t see it, it’s not your problem) I do, yes (But on the other hand you are thinking ahead about things you can’t see) Well, yes, in that way you are. That train has shown up now, that 1H10 - so it must be round the Stockport area now. (??) It’ll come straight into platform 8, I think, that one goes into. What time is he due in - yes, he goes into number 8. That’s the only thing really, most of the time I don’t worry about the train until he’s actually hit the panel - you’ve got to plan your station - that’s the only thing you’ve got to watch, is your
platforming. You end up trapped otherwise. (You are planning both your platforms and your engine sidings so you don’t get stuck in the far end) Yes. In that respect you’ve got to think ahead a little bit - and if a train is going to be late, you’ve got to find out how late its going to be before you commit yourself (Longsight: Down the slow please) (That one going through platform 13?) That’s a freight that - 6V46 - or is it that one there you’re thinking of - that one’s going to Trafford Park that 6H13, where the other fella’s come from. He’s indicated out on platform 3, going back to Hadfield, and he’s ready on platform 10, but we can’t get him out until that other one has come out of his way (13 and 10 clash). (What’s the flashing light on 10?) That’s him wanting to come out, 3V20. But we can’t shift that because we’re still coming through platform 13 with that 6V46 you see - he’s got to be clear at 332 (??) signals before we can give him signals off 10 - we can set the route up, but the signal won’t come up until he’s clear. (Right, so he’s got a route off now) And he’s now ready on number 7 - (He’s got the green light, [BEEP] and the 2H02 for Buxton is clearing out again) Yes, he’s ready to go, so we’ll pull off for him, and see if anything happens. (You’re putting him down the up fast again - because that’s the way you like it?) That’s the way I prefer to do it, yes. (You might have taken him on the down fast?) Yes, but you see if you do that, you’ve got 1H10 coming down, and if he doesn’t move, we can’t get 1H10 in. (1H10 is going into?) 1H10 is going into platform 8. He’s on the move now, out of platform 10. Up the slow Paul, 3V20. Now if I’m not mistaken, that one on platform 4 is out at 15.10 [checks sheets] (That’s the 5H58?) uh hu, it is. (So he’s going to want to be cleared out in a minute?) He’ll be looking for a road in a minute yes. He’ll have to come out on the down East - he’s going round the corner, the same way as that Hadfield is going (It looks pretty clear for him?) Yes, well at the moment he can go as far as 342 - I’ll just give him one signal, then put him onto the cross over onto the up East you see. (So he’ll come out down East, and then he’ll go over to the up East, which is what he should be on) Yes. (The up East has only got a one-directional arrow, whereas the Down East has got a bi-directional arrow) All the lines are bidirectional bar the up east. Up the slow for Style. (So up East, you’re only allowed to go right to left) That’s it, yes, one direction only. When you look at the diagram though, its the only way you could go because there’s no way to get onto the up East from the left hand direction, you see, going left to right. You couldn’t get on it anyway. It looks like the Buxton is on the move out of 7, yes. That engine on 7 said he would follow up, lets hope he does. [Down the slow].

251
What I'll do now, I won't bother putting that engine on 7 if he comes up, I won't put him into the engine sidings, because 1H10 is there, so I'll run 1H10 in, and then put the engine straight out of the engine sidings onto the train, and then I'll put him in the engine sidings.

I may as well back straight up onto the train and then its done with, rather than block the engine in waiting for another train which hasn't yet appeared on the scene (so you're leaving the guy on 7 where he is) I'll leave him there for another couple of minutes until the other fella's in (And then take him straight out and put him back in?) Yes. (So the one going on this incoming train is the one in the engine sidings?) That is correct. (So you may well put 0V62 in the engine sidings, but at the moment he can stay where he is?) He can stay where he is. He's not doing any harm where he is, and I've not got to get him out of the way for any particular reason so there's no point in shifting him just yet. He's ready on platform 4, so we'll take a chance with him. (He's gone to 342 points) that's right, he'll go over and return to his right line of travel. We'll put an engine number in for 1H10 - oh we'll knock that one off there (Where's that one you've just put an engine in?) On to platform 8 - the engine that brings in 1H10 will work 1A09 back out again (So he's not come up on the berth again) No, he's not in yet - as soon as 1H10 comes in, it'll push that to the second berth - the buffer end. (So 0A09 is actually an engine?) Yes.

THE END.
2. Task Instructions
The following is a reproduction of the single instruction sheet given to the signalmen in this study.

I am interested in what you think about when you control the movements of trains in and out of the station. To help me understand this, I would like you to talk about what you are doing as you go along.

In particular, I am interested in hearing what you are planning to do next and what you think you will be doing after that, and so on. You should try to talk all the time (if there are periods of silence, I will probably prompt you), saying whatever you are thinking about, or whatever you notice about what you are doing - I know it's not easy. It might help you to provide me with what I want if you try to talk in the future tense some of the time (“I will do...” or “I am going to do...”). If you also have a reason why you are deciding to do several things in a particular order, then say it. In order to have something to say all the time, some of the things you say might seem to be pretty trivial - this is OK - you should say everything that comes to mind, and try to describe everything that you do or are going to do.

Naturally, it is more important that you concentrate on the job in hand, and so talking to me about it should take second place.

I would like to reassure you that although I will be recording what you say, this will be used for my own research only, and at no time will any person's identity be used, nor will any of the recordings be used as an official assessment of performance.

Finally, I would like to thank you for agreeing to help me in this way. I am very grateful.
Appendix D

This appendix contains the instructions and information sheets provided for subjects in the
carwash simulation study of Chapter 7. In addition, there are two colour photographs
showing a the experimental setup and an example of one of the screens.

Welcome.

Thankyou for agreeing to participate in this study.

The study is divided into 4 parts (plus an initial part in which you will be shown what to
do, and have the opportunity to practice), and should take approximately 1 hour 30 minutes
altogether. You will receive the sum of £5 for taking part.

There will be a video camera recording the study, but this will be used only by me for
analysing the results. Anything you do or say during the study will be treated in the
strictest confidence. Your name will not be used in any report.

The task you will be asked to do will be described to you shortly, but requires no special
experience.

Do you have any questions so far?
Introduction
In this experiment you will play a computer game in which you have the role of a Car Wash Manager. This job involves controlling many things happening at once - the aim of the experiment is to find out how efficiently people can do this.

The Car Wash which you will manage employs an excess of people, but unfortunately has only very few of the other things it needs (such as Buckets and Hosepipes - these are called resources). The efficiency of the Car Wash, and the speed with which cars are dealt with, is down to your skill in deciding how the resources should be shared.

Your goal, then, is to make the Car Wash as efficient as possible - by processing all the incoming cars in the shortest possible time.

The Car Wash has a queue of incoming cars, which you can choose to deal with in any order you like. The cars are of different size and will want one of the 3 levels of service that your Car Wash offers (Standard, Extra, and Deluxe). You will be informed about each car's size and level of service as it arrives.

In the Car Wash itself, there are 3 Bays. Each Bay can hold one car, and a car must be in a Bay before you can do anything to it. Unfortunately, due to a recruitment oversight, you are the only person who can drive the cars in and out of the bays. This means that you will have to spend some of your time away from your control centre, and you should bear this in mind.

All of the above details are summarized in the form of tables which you can refer to as you work.

Different Levels of Service
There are a number of different things which can be done to a car - the particular set depends on the level of service that the owner of the car has chosen. The Standard, basic, level of service consists of just Washing and Rinsing. The other levels of service add extra processes on top of this.

Processes and Resources.
Different processes require different resources - so Washing might require a Bucket, whilst Rinsing would require a Hosepipe.
Processes and Processes.
Naturally, some things can only be done in a sequence - so it doesn’t make sense, for example, to Rinse a car before you have Washed it. Similarly you can’t drive it out of the bay until you have finished everything else on it.

Processes and Cars
Cars vary in size - your Car Wash groups them as Small, Medium, and Large. The time taken by some processes, such as Washing, varies with the size of the car, but others (such as emptying the Ashtrays) take the same time whatever the size.

The Computer
Although it might be fun to have a real Car Wash for you to manage, this is not practical. Instead, a computer is being used to simulate the main features (except the water).

The middle screen in front of you is the Car Wash itself (don’t worry about the other two screens for the time being). The top half of this screen is divided into three - representing the 3 bays in which you can put cars.

On the right at the bottom is a list of all the cars waiting, together with their sizes. You will know what level of service they require and you can choose to do them in any order you wish. Once you load a car into a bay however, you cannot unload it until it has finished.

On the left at the bottom of the screen is a box representing your store of resources (called the “Resource Store”). This lists all the different resources that you have, together with an indication of how many you have in total and how many of those you are currently using.

The other details of the screen will be explained to you in the course of a training trial.

Making Things Happen
Since there is an abundancy of people standing round waiting to do something, simply allocating the necessary resources to a particular car and process is sufficient for that process to start being done (assuming, of course, that any processes upon which it depends is complete).

To make life easier for you, the experimenter will act as your dumb assistant - you tell him what you want to happen, and it will happen. However, he is quite stupid and only understands a few simple commands. These are described below. (Slight variation on these are allowed).
To allocate a resource you simply tell the experimenter what you would like to happen, in the following way:

USE resource TO process car

eg, in the case of allocating a Bucket to Wash the Escort, you would say something like:

USE a Bucket TO Wash the Escort

If you try an allocate eg more Buckets than you have available in the store, then nothing will happen.

Once a process has been allocated all the resources it needs, its ‘Time’ counter will start decreasing. This shows you how much of that process is still remaining. If you think that a resource allocated to this process would be better somewhere else, you can give a command to remove it:

TAKE AWAY resource FROM process car

eg to undo the previous allocation:

TAKE AWAY the Bucket FROM Washing the Escort

As soon as you do this though, the process will halt until you give it back the resources it needs.

When the Time count reaches zero, the process is complete - it disappears from the car’s list and any resources it has are automatically returned to the store (so you can use them again).

The final item on each car’s list of things to be done to it, is to Drive it Out of the Bay. The resource required by this is ‘You’ (ME?). Allocating ‘You’ to this process results in the car being removed from the Bay. You can then use this bay for another car in the queue.

In the same way that you drive in a car, the following command can be used to drive out a car when it is finished:

DRIVE OUT <car>
Note that this has exactly the same effect as the command:

USE me TO drive out <car>

Summary of commands
DRIVE <car> INTO <bay>
USE <resource> TO <process> <car>
TAKE AWAY <resource> FROM <process> <car>
DRIVE <car> OUT

Finally
If you have any questions or don’t understand something, please ask.
The Car Wash

Services offered:

<table>
<thead>
<tr>
<th>Service</th>
<th>Wash</th>
<th>Rinse</th>
<th>Ashtrays</th>
<th>Vacuum</th>
<th>Polish</th>
</tr>
</thead>
<tbody>
<tr>
<td>Standard</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Extra</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Deluxe</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>

Details of the processes:

<table>
<thead>
<tr>
<th>Process</th>
<th>Time</th>
<th>Needs</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wash</td>
<td>20</td>
<td>Bucket</td>
<td></td>
</tr>
<tr>
<td></td>
<td>30</td>
<td>Sponge</td>
<td></td>
</tr>
<tr>
<td></td>
<td>40</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rinse</td>
<td>20</td>
<td>Hosepipe</td>
<td>Must be WASHED first</td>
</tr>
<tr>
<td></td>
<td>30</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>40</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ashtrays</td>
<td>10</td>
<td>Dustbin</td>
<td></td>
</tr>
<tr>
<td></td>
<td>10</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vacuum</td>
<td>20</td>
<td>Hoover</td>
<td></td>
</tr>
<tr>
<td></td>
<td>30</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>40</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Polish</td>
<td>20</td>
<td>Wax</td>
<td></td>
</tr>
<tr>
<td></td>
<td>30</td>
<td>Cloth</td>
<td></td>
</tr>
<tr>
<td></td>
<td>40</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Driving IN</td>
<td>10</td>
<td>You</td>
<td></td>
</tr>
<tr>
<td></td>
<td>10</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Driving OUT</td>
<td>10</td>
<td>You</td>
<td></td>
</tr>
<tr>
<td></td>
<td>10</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>10</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Summary of commands

DRIVE <car> INTO <bay>

USE <resource> TO <process> <car>

TAKE AWAY <resource> FROM <process> <car>

DRIVE <car> OUT
Subject self rating sheet:

Subject No:   Trial No: ______

How efficient do you think you were?

I would like to know how efficient you think you were in washing the last set of cars. Please circle the appropriate number below.

<table>
<thead>
<tr>
<th>Not efficient at all</th>
<th>Very efficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 1 2 3 4 5 6 7 8 9 10</td>
<td></td>
</tr>
</tbody>
</table>

How difficult did you find the task? Please circle an appropriate number on the scale below.

<table>
<thead>
<tr>
<th>Not difficult at all</th>
<th>Very difficult</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 1 2 3 4 5 6 7 8 9 10</td>
<td></td>
</tr>
</tbody>
</table>

If you have any comments, you can write them in the following space:
The Cognitive Failures Questionnaire (CFQ).

The following questions are about minor mistakes which everyone makes from time to time, but some of which happen more often than others. We want to know how often these things have happened to you in the last six months. Please circle the appropriate number.

<table>
<thead>
<tr>
<th></th>
<th>Very often</th>
<th>Quite often</th>
<th>Occasionally</th>
<th>Very rarely</th>
<th>Never</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Do you read something and find you haven’t been thinking about it and must read it again?</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>2. Do you find you forget why you went from one part of the house to the other?</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>3. Do you fail to notice signposts on the road?</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>4. Do you find you confuse left and right when giving directions?</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>5. Do you bump into people?</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>6. Do you find you forget whether you’ve turned off a light or a fire, or locked the door?</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>7. Do you fail to listen to other people’s names when you are meeting them?</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>8. Do you say something and realize afterwards that it might be taken as insulting?</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>9. Do you fail to hear people speaking to you when you are doing something else?</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>10. Do you loose your temper and regret it?</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>11. Do you leave important letters unanswered for days?</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>12. Do you find you forget which way to turn on a road you know well but rarely use?</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>13. Do you fail to see what you want in a supermarket (although it’s there)?</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>14. Do you find yourself suddenly wondering whether you’ve used a word correctly?</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>15. Do you have trouble making up your mind?</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>16. Do you find you forget appointments?</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>17. Do you forget where you put something, like a newspaper or a book?</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>18. Do you find you accidentally throw away the thing you want and keep what you meant to throw away - as in the example of throwing away the matchbox and putting the used match in your pocket?</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>19. Do you daydream when you ought to be listening to something?</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>20. Do you find you forget people’s names?</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>21. Do you start doing one thing at home and get distracted into doing something else (unintentionally)?</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>22. Do you find you can’t quite remember something although it’s “on the tip of your tongue”?</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>23. Do you find you forget what you came to the shops to buy?</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>24. Do you drop things?</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>25. Do you find you can’t think of anything to say?</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>5</td>
</tr>
</tbody>
</table>
In the next trial, you will have an extra screen to help you. The screen on the left will have a list of all the resources in your Resource Store, together with all the processes which still need that resource. As soon as you allocate the particular resource to this process, it is removed from this list. This is provided to simplify your job, since you can now tell at a glance what processes are competing with each other for a resource.

Another way of thinking about this, is that you are aiming to remove all the entries from this screen - as soon as you have done this, then you only have to sit back and wait for everything to finish (or more cars to arrive!).

Only details of the cars which you have driven into a bay are shown here, cars still in the queue are not.

Do you have any questions? (I will demonstrate this to you before the next trial).
Additional Instructions (S)

In the next trial, you will be able to queue-up uses of resources. This means that if the store of resources shows that you are using all the Buckets, for example, you can say what you would like to be done with a bucket as soon as one is available. The screen on the right shows a queue of such future allocations - you can have up to 3 things waiting to use a resource. They will be done in the order they appear on the screen, but you can change this. The extra commands that you need to use for this are described below. In a moment, this system will be demonstrated to you.

You can only queue-up processes for cars which are already in a bay - if you try to queue-up a resource to be used on a car which is still in the car queue, nothing will happen. (This also means that you can’t queue-up an intention to Drive In a car to a Bay).

Extra commands:

QUEUE <resource> TO <process> <car> NEXT
QUEUE <resource> TO <process> <car> AFTER <process> <car>

eg if the Bucket is currently being used to Wash the Mini, and you would like it wash the Escort next, you might give the following command:

QUEUE the Bucket TO Wash the Escort NEXT

The first item in the Bucket’s queue on the right-hand screen would then be to Wash the Escort.

If you would like to follow Washing the Escort by Washing the Jaguar, the you could now say:

QUEUE the Bucket TO Wash the Jaguar AFTER Washing the Escort
Additional Instructions (SP)

In the next trial, you will have two extra screens to help you. The screen on the left will have a list of all the resources in your Resource Store, together with all the processes which still need that resource. As soon as you allocate the particular resource to this process, it is removed from this list. This is provided to simplify your job, since you can now tell at a glance what processes are competing with each other for a resource.

Another way of thinking about this, is that you are aiming to remove all the entries from this screen - as soon as you have done this, then you only have to sit back and wait for everything to finish (or more cars to arrive!).

Only details of the cars which you have driven into a bay are shown here, cars still in the queue are not.

The screen on the right allows you to queue-up uses of resources. This means that if the store of resources shows that you are using all the Buckets, for example, you can say what you would like to be done with a bucket as soon as one is available. The screen on the right shows a queue of such future allocations - you can have upto 3 things waiting to use a resource. They will be done in the order they appear on the screen, but you can change this. The extra commands that you need to use for this are described below. The system will be demonstrated to you in a moment.

You can only queue-up processes for cars which are already in a bay - if you try to queue-up a resource to be used on a car which is still in the car queue, nothing will happen. (This also means that you can't queue-up an intention to Drive In a car to a Bay).

If you queue-up a use of a resource, then this use will disappear from the screen on the left.

Extra commands:

QUEUE <resource> TO <process> <car> NEXT
QUEUE <resource> TO <process> <car> AFTER <process> <car>

eg if the Bucket is currently being used to Wash the Mini, and you would like it wash the Escort next, you might give the following command:

QUEUE the Bucket TO Wash the Escort NEXT
The first item in the Bucket's queue on the right-hand screen would then be to Wash the Escort.

If you would like to follow Washing the Escort by Washing the Jaguar, you could now say:

**QUEUE the Bucket TO Wash the Jaguar AFTER Washing the Escort**
Photograph showing a subject and the three screens of the carwash simulation.

<table>
<thead>
<tr>
<th>Resource Store</th>
<th>Name</th>
<th>Used</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Bucket</td>
<td>B</td>
<td></td>
</tr>
<tr>
<td>2. Cloth</td>
<td>C</td>
<td></td>
</tr>
<tr>
<td>3. Dustbin</td>
<td>D</td>
<td></td>
</tr>
<tr>
<td>4. Hose pipe</td>
<td>H</td>
<td></td>
</tr>
<tr>
<td>5. Sponge</td>
<td>S</td>
<td></td>
</tr>
<tr>
<td>6. Vacuum Cleaner</td>
<td>V</td>
<td></td>
</tr>
<tr>
<td>7. Wax</td>
<td>W</td>
<td></td>
</tr>
<tr>
<td>8. You</td>
<td>Y</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Resource Store</th>
<th>Name</th>
<th>Size</th>
<th>Service</th>
</tr>
</thead>
<tbody>
<tr>
<td>1: Rainier</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2: Escort</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3: Uno</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Photograph of the main carwash screen