An Exploration of the Role of System Dynamics in the Analysis of Disruption and Delay for Litigation

by Susan M. Howick

Submitted for the degree of

Doctor of Philosophy

Department of Management Science
University of Strathclyde
Glasgow, UK

2001

ACKNOWLEDGEMENTS

There are a number of people I would like to thank for their support throughout the work reflected in this thesis. However, there are a few that I would particularly like to recognise.

I would like to thank Colin for his many useful comments, his guidance and his continual encouragement.

I would like to thank my parents who have supported and encouraged me, not only throughout the work carried out to produce this thesis, but also throughout the rest of my life.

This thesis is dedicated to Matt without whose continual support it would never have been completed.

ABSTRACT

System Dynamics (SD) is a modelling approach that has been used to support litigation cases that are investigating overruns on large engineering projects caused by Disruption and Delay (D&D). However, the role that SD can play in the analysis of D&D in large projects is not fully understood.

The first aim of the research is to explore the appropriateness of SD as a modelling approach in the analysis of D&D for litigation. Criteria on the suitability of SD to model a situation are taken from the SD literature and explored to understand their level of contribution to the research. Experiences from the researcher's involvement in two litigation cases are then used to test how empirical data performs against the criteria. The explorations lead to a revised set of criteria being proposed. These criteria should be used to assess whether or not SD should be used to analyse D&D for any specific litigation case. Testing the data against the criteria also results in lessons for the modelling of D&D. This includes a proposed method of assessing the level of D&D in a project through an analysis of managerial actions.

The second aim of the research is to explore the issues that are involved in using SD to analyse D&D for litigation. The approach taken uses the empirical data to test the degree to which SD can meet the purposes of modelling D&D for litigation. This process leads to a number of conclusions. It highlights limitations of using SD in this environment; emphasises the importance that the audience plays in the modelling process; explores the difficulties encountered in gaining audience confidence in the model; provides an appreciation of the validation process required when modelling in this environment.

The research provides an initial understanding of the role that SD can play in the analysis of D&D for litigation. It is hoped that this can be built on with future experiences of modelling D&D for litigation.

TABLE OF CONTENTS

CHAPTER 1: INTRODUCTION

1.1 Introduction

Large projects have been undertaken all over the world for many centuries. For example, the design and construction of the pyramids in ancient Egypt and the design and construction of the Scottish Parliament buildings today. The use of projects throughout the centuries has been to implement change in society. Therefore, if society is to advance effectively, the successful implementation of projects is crucial. From an individual organisation's point of view, the successful implementation of projects is also crucial since "The pace of change... has been increasing at an ever-faster rate. Effectively and efficiently managing change efforts is the only way organizations can survive in this modern world" (Webster 1993, p5).

Although the implementation of successful projects is crucial to both society and the survival of organisations, many large projects have ended in failure (Morris and Hough 1987, Kharbanda and Pinto 1996). Failure for a project may mean completing the project late, spending in excess of the budget, the final product not meeting its required specifications or a cancellation of the project prior to its completion. In particular, cost and schedule overruns are very common in large projects (Morris and Hough 1987). However, understanding the reasons for complex projects overrunning in cost and/or schedule is by no means straightforward, but it is necessary. For organisations to improve in the implementation of projects, they need to be able to learn from their mistakes. Therefore, when projects fail, organisations need to gain an understanding of the reasons for the failure.

When organisations attempt to understand why a project failed, the process may result in some of the blame being placed on other parties involved in the project. For example, the project may have been disrupted or delayed due to actions taken by a client or contractor. As a result of this conclusion, the organisation may attempt to seek compensation for the outcomes that resulted from the actions by the client or contractor. However, if the client or contractor is not willing to compensate the organisation, the process may result in the organisation commencing litigation proceedings.

Litigation can be a long process and one where organisations may be requesting millions of dollars of compensation. The plaintiff will treat the litigation process as a project itself. Although the investment in the project can be large both in time and in cost, the potential return on the investment may also be huge. It is unlikely that any other project that they may participate in would have such a large potential return. However, litigation is risky. The plaintiff may invest a large amount of time and money in the process with no return. The plaintiff also runs the risk of impacting future business. The defendant, as well as other organisations, may not wish to do business with the plaintiff in the future since they may be concerned that any problems in future work will result in the plaintiff litigating against them. Due to the risks involved, it is vital that the plaintiff has the best possible support for their case before proceeding with litigation.

To help support claims for compensation, mathematical models have been used in litigation processes. In particular, computer simulation models have been found to be useful in this environment. "The advantage of a simulation model is its ability to portray a complicated situation better than any verbal description could hope to do. A model can disclose relationships between various events which might not otherwise be apparent. Also, a model makes it possible to consider all relevant factors simultaneously in the solution of the problem" (Fleming 1980, p874). For these reasons "... litigants have successfully used the results of computer simulation as the basis for expert testimony" (Fleming 1980, p874).

When considering simulation modelling approaches, System Dynamics (SD) has been one approach that has been used in a few litigation cases to explain the reasons for time and cost overruns on large engineering projects (Cooper 1980, Weil and Etherton 1990, Ackermann et al 1997). SD is a simulation modelling approach that was specifically designed to model and explore feedback. SD was introduced as *Industrial Dynamics* (Forrester 1961) since its first application was to explore the behaviour of industrial systems. However, the approach has since become known as SD to reflect the wider applications of the modelling approach.

The use of SD in a litigation environment has been to model the effects of Disruption and Delay (D&D) (Cooper 1980, Weil and Etherton 1990, Ackermann et al 1997). D&D in a project can be triggered by a simple delay or disruption. However, the ramifications of D&D in the project can be complex. Therefore, a simple analysis of the direct consequences of each disruption and each delay will not cover the full impact of D&D. Due to the complexity of D&D, it becomes difficult to attribute any project outcome directly to any one disruption or delay. Any approach used to model D&D needs to be able to model the paths between an initial disruption or delay and the final outcome of that event in terms of both a schedule and cost overrun.

In each of the litigation cases that SD has been used to model D&D, the SD model has formed the basis of a claim for compensation for many millions of dollars. The effects of D&D for each of these cases have been estimated to represent at least 40% of the overall claim. For any organisation, the gain or loss of millions of dollars in this way represents an extremely large amount of money. Indeed, it could be a sum of money that either bankrupts the organisation or, at the other extreme, provides it with its most profitable venture. These potential outcomes mean that there is a need to ensure that if the SD approach is used to model the effects of D&D in project overruns, it is done in a way that ensures that the organisation has the optimum chance of gaining or rebuffing the compensation that is being sought.

Large projects are important to society. Therefore, projects will keep on occurring in society. However, if projects keep on overrunning in cost and schedule, organisations will only be able to survive if they attempt to seek compensation for the losses they incur. Litigation is therefore going to continue, as other organisations are unlikely to provide the compensation that an injured organisation is seeking. However, litigation is a risky project to undertake. Therefore, to ensure the optimum chance of success, organisations need to fully understand the approaches that are available to support their case. SD is a modelling approach that has been used to support litigation cases for overruns caused by D&D on large engineering projects. Therefore, if SD is to continue to be used successfully, there is a need to fully understand the role that it can play in the analysis of D&D in large projects. If there is a lack of knowledge of the capabilities and the limitations of SD during litigation, then the modelling approach will not be used in a way that provides optimum support for a claim. An organisation may therefore suffer through reduced compensation, or a complete failure to either gain or rebuff compensation. This may then have a detrimental effect on the survival of the organisation.

Although an understanding of the role that SD can play in the analysis of D&D for litigation is argued to be important, there is nevertheless a lack of literature in this area. This thesis aims to improve this situation.

1.2 The Aim of the Thesis

The overall aim of the thesis is to explore the role of SD in the analysis of D&D for litigation. This aim brings together two topics; SD and the modelling of D&D. The thesis therefore explores both of these topics in an attempt to discover what each topic can inform us about its use alongside the other topic.

Before SD can be used in the analysis of D&D, a decision has to be made whether or not it is a suitable approach to use to model D&D for any individual project. Therefore, the first aim of this thesis is:

- *to explore the appropriateness of SD as a modelling approach in the analysis of D&D for litigation.*

The first resource used to investigate this issue is the SD literature. This should provide examples of the types of situations that SD has been used to model in the past. Also, it is hoped that this will give a general indication of the types of situations where SD *should* be used. If criteria can be extracted from the literature on the suitability of SD to model a situation, then this could be used to assess whether or not SD should be used to model D&D for any given situation.

Assuming that the SD literature does not banish the use of SD in the analysis of D&D, then the second aim of this thesis is:

- *to explore the issues that are involved in using SD to analyse D&D for litigation.*

An exploration of the second topic; the modelling of D&D, can help achieve this second aim.

It was noted above that to enable SD to support a litigation claim for an overrun project, there is a need to fully understand the role that SD can play in the analysis of project overruns. Therefore, throughout the explorations, if any lessons can be gained about the use of SD in the analysis of D&D, then this would be beneficial. Also, if any knowledge can be gained separately about either SD or the modelling of D&D, then those people using SD to analyse D&D would also benefit from an increased understanding of the two topics.

1.3 The Approach used for the Research

1.3.1 Data Opportunities

To aid the explorations undertaken in this thesis, the author draws upon the experiences of a number of litigation cases. These experiences come from the author's association with eight large engineering projects where D&D has been analysed for litigation purposes. Although all eight projects inform the explorations in this thesis, two of the projects provide a more detailed input. For these two projects the author has been integrally involved in the analysis of D&D. This involvement has meant that the author has been able to gather a large amount of data regarding the analysis of D&D for litigation and hence use it to help to develop the explorations throughout the thesis.

For both of the litigation cases that the author was integrally involved in, compensation was being sought for a time and cost overrun that occurred on a large aircraft modification project. One of the litigation cases involved the main contractor of the project seeking compensation from the client for the project whilst the other litigation case involved a sub-contractor seeking compensation from the prime contractor for the project. In both cases the plaintiff for the litigation case believed that the defendant had caused disruptions and delays to the project that had contributed to the project overrun. The author was involved in both of these cases as one of the consultants that the plaintiff hired to analyse the causes and effects of D&D in the projects. The work that the author was involved in has, so far, spanned four years, however neither claim has yet been settled. A fuller understanding of the two projects can be gained in chapter 2 when they are described in detail.

The involvement in the two projects has provided the author with a privileged opportunity, rarely made available to researchers. The reasons why this opportunity is so rare are as follows:

- Litigation is not a process that is entered into lightly. Although an organisation may believe that a client or contractor is to blame for a project overrunning in both time and cost, they need to be sure that they have sufficient evidence to prove their case. The gathering of this evidence may take a very long time and cost a great deal of money. For example, the cost of each of the litigation processes for those projects that the author has been associated with have been in excess of £1m. However, as well as the direct cost of, for example, lawyers and consultants' fees, resources within the organisation are also tied up during the period of litigation. Senior management will need to spend time directing the claim, whilst at least one member of staff may be tied up managing the claim. Staff will also be required to manage the gathering of data and recording of all relevant documentation. Also, those who worked on the project may be regularly asked to provide witness statements during the litigation process. The use of all these resources during the litigation process mean that the organisation cannot use these resources on other projects and therefore other projects may be disrupted.

  If one organisation litigates against another organisation, this can have an extremely detrimental effect on future business between the two organisations. The defendant may be put off doing future business with the plaintiff since they may be concerned that they are under threat of litigation if anything goes wrong with future projects. Other organisations may also be put off doing business with the plaintiff, since by seeing them go ahead with litigation proceedings on one organisation, they may feel threatened that the same process may be used in future projects they are involved in with the plaintiff.

  The above highlights the immediate costs associated with litigation as well as the future costs of potential lost business. Therefore, although many projects can fail, the number of these projects that actually result in litigation proceedings is far fewer. This reduced number of potential projects means that there is a lack of projects from which to gather data.

- If one organisation does decide to litigate against another organisation, the process is extremely confidential. The plaintiff will want to keep their discussions confidential, so that the defendant is not made aware of the strategy they decide to adopt during the litigation process. Also, confidentiality is important so that any individuals that are a part of making any decisions about compensation, for example a judge, or potentially a jury, are not influenced by any prior knowledge about the project. For these reasons, any consultants used during the litigation process are required to sign confidentiality agreements. Therefore, if researchers were to request access to data from a litigation process, an organisation is likely to be very wary of granting such access.

- As organisations are extremely wary about allowing a researcher access to a litigation process, the only access that is potentially open to the researcher is through organisations where the researcher has a long-standing relationship and therefore a level of trust has built up between the organisation and the researcher. However, this level of trust may take many years to build. For this reason, immediate access to organisations for this type of research material is generally not possible.

- Due to the importance of a litigation process, i.e. the potential gain or loss of millions of dollars and the impact it can have on the future business for the plaintiff, the plaintiff is likely to want to minimise any potential interference in the process. This means that they are unlikely to agree to researchers solely being a part of the process so that they can gather data for research purposes. This may cause some of their staff's time being used up answering questions for research purposes, whereas it should be used to meet important deadlines for the litigation process. This produces a further hurdle for a researcher to gain access to a litigation process. However, one way of getting around this hurdle is for the researcher to be a part of the litigation process, for example a consultant to the plaintiff. This means that they can gather data for the dual purposes of the consultancy work as well as the research. However, opportunities to get involved in a litigation process as a consultant to the plaintiff are not open to every researcher. A plaintiff will normally seek consultants who have experience in this field of work. This therefore places even further restrictions on access to researchers.

Each of the above points places restrictions on the number of projects that a researcher can gain access to if he or she wishes to gather data from projects where D&D has been analysed for litigation. Therefore, although *two* projects does not appear many from which to draw conclusions and hence calls into question the generalisability of the conclusions from this thesis, the author has actually been provided with a privileged opportunity which rarely arises. The lack of opportunities of gaining data regarding the analysis of D&D in large projects during litigation precludes a wider study. Therefore, the two projects that the researcher was involved in actually provide very valuable data for the explorations covered in this thesis.

The data made available to the researcher is also very valuable due to the large amount of time and money that a plaintiff is willing to invest in the litigation process. This is an exceptional resource that is not normally made available to researchers. Hence, the access the researcher had to data during the litigation processes has provided the researcher with a rare opportunity.

1.3.2 Data Collection

The data gathered from the two projects came from three sources:

- Interviews with members of the plaintiff's organisation who had either been involved in the project or were a part of the plaintiff's claim team.
- Documentation produced or received by the plaintiff's organisation. This included letters, memorandums, reports etc each of which may include recorded data.
- Observations made by the researcher during the role as a consultant in the litigation process.

Throughout the data collection, the author had dual objectives to consider. As well as the objectives of this research, the author had to collect data as part of her role as a consultant who was analysing D&D that occurred in the projects. In some circumstances these objectives meant that there was a need for the same data requirements. However, often the researcher's role took the data requirements beyond those for the consultant's role. Indeed the data collected from simply observing the litigation process was mainly gathered to progress the research objectives. However, this would not have been possible if the author had not already been a part of the litigation process.

Due to the two separate roles that the researcher had to undertake, it was important that throughout the process the researcher bore the different objectives of the roles in mind. The other consultant working with the author on each of the projects took the role of project manager for the consultancy work and hence ensured that the consultancy objectives were met. However, when collecting data, the researcher still had to place an emphasis on the obligations of the researcher as a consultant to the organisation. The objectives of the consultancy role were given priority, since the research opportunity only arose due to the existence of the consultancy role. However, if opportunities arose to collect additional research data whilst consultancy data collection was being carried out, these were taken as much as possible. This meant that the gathering of research data was often dictated by the need for data in the consultancy role.

A further complication that the dual role as researcher and consultant produced was the awareness of the researcher being a part of the situation being researched. The second aim of this thesis is to explore the issues involved when using SD to analyse D&D. This was carried out through an exploration of the purposes of modelling D&D for litigation. The purposes of modelling a situation are dictated by the audience for the model and the modeller is always an audience for the model. Therefore, an exploration of the issues involved when using SD to analyse D&D will include the modeller's, i.e. the researcher's views on the model. On one hand, this may give the researcher a deeper understanding of the issues that need to be addressed. However, the researcher also has to bear in mind the potential for bias by the researcher when exploring the issues. For example, the modeller may be more supportive of the modelling approach undertaken than a researcher who is external to the process. It would be impossible to eliminate this potential bias without actually becoming an external researcher. However, the difficulties of being permitted to take on such a role in this research environment have already been discussed. The researcher therefore had to be aware of the potential for such issues to arise throughout the research process and attempt to avoid them.

The issues of the dual role of researcher and consultant also arise during any form of action research. Although the research covered in this thesis did not take the form of action research, similarities exist between the two. The cyclical process of action research can be seen in figure 1.1 below.

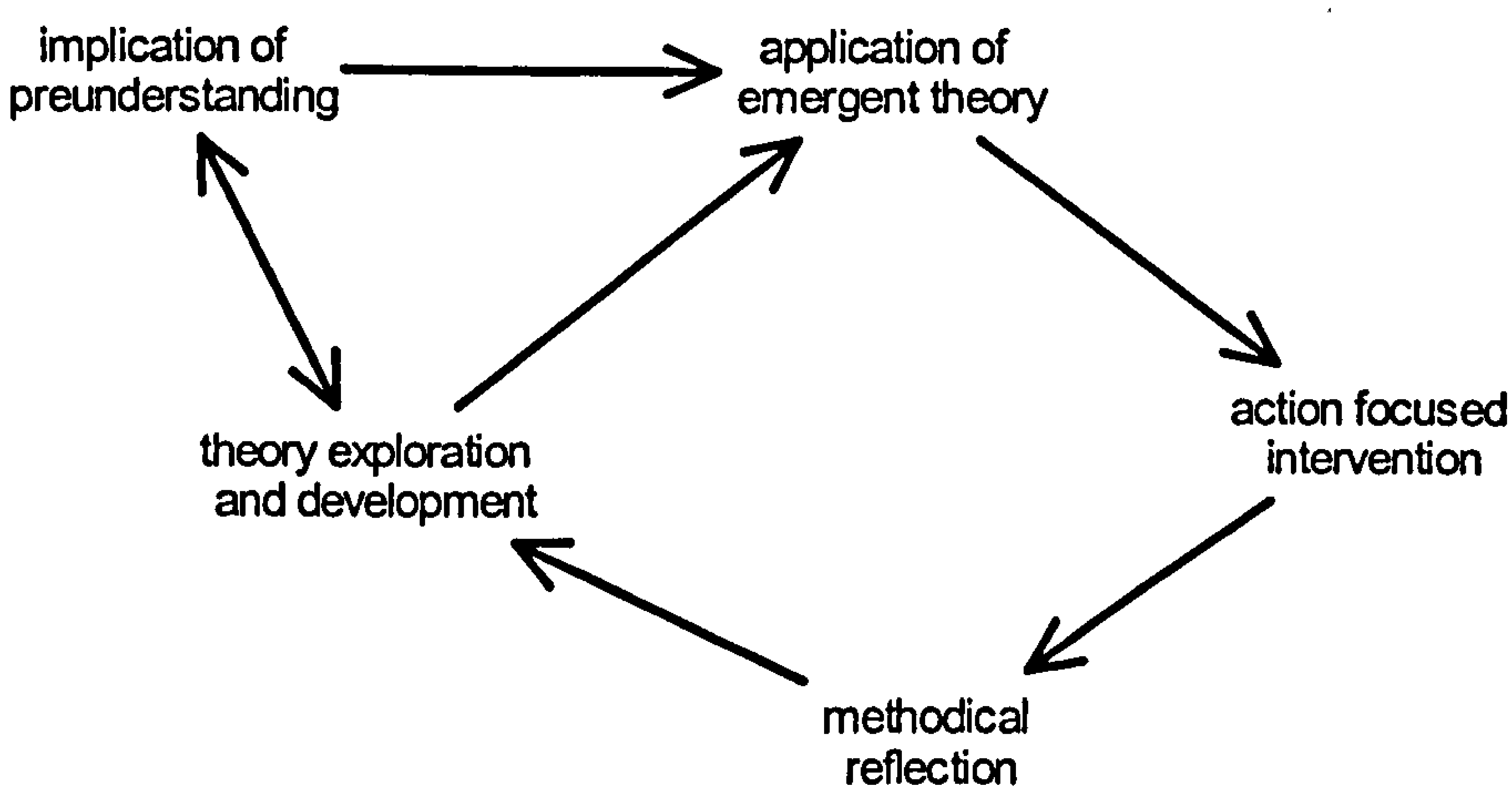


Figure 1.1 – The Cyclical Process of Action Research taken from Eden and Huxham 1996

The explorations in this thesis develop theory based on data gained from practice. However, the theory developed in this research has been developed as a reflection on action, rather than being used as an intervention on action. Therefore the action research cycle of theory-to-action-to-reflection-to developing theory is not sustained due to the lack of action focused intervention once theory has been developed.

1.3.3 Using the Data in the Explorations

Before commencing the research, the researcher did not develop a theory for the research to test. Instead, the objective of the research was to explore the data collected together with existing literature to discover what conclusions could be arrived at to further the knowledge of the role of SD in the analysis of D&D for litigation. Therefore, an exploratory approach was adopted.

The first aim of the thesis is to explore whether or not SD should be used in the analysis of D&D for litigation. This involves extracting criteria on the suitability of SD to model a situation from the SD literature and exploring the criteria to understand their level of contribution to the explorations. Experiences from the researcher's involvement in the two litigation cases are then used to test how empirical data performs against the criteria.

The research then goes on to explore the issues that are involved in using SD to analyse D&D for litigation. The approach that is taken is to use the data to test the degree to which SD can meet the purposes of modelling D&D for litigation.

During the explorations, both the literature and data were called upon as appropriate in an attempt to see if they could add any insights to the explorations. For this reason, the data from the two projects are not discussed together in one section. Instead, the data is used throughout the thesis in small sections to aid many different minor explorations. Each of these minor explorations aims to further the overall explorations, one step at a time.

1.4 Structure of the Thesis

This thesis has been structured in line with the two main aims that were highlighted in section 1.2.

The thesis initially focuses on the first aim:

- *to explore the appropriateness of SD as a modelling approach in the analysis of D&D.*

To provide the reader with sufficient background information to enable a full understanding of the explorations covered in this thesis, an introduction to SD and a discussion on the definition of D&D will be covered in chapter 2. The SD literature is then used as a basis of the explorations to achieve the first aim of the thesis. In carrying out these explorations, three different viewpoints are highlighted in the SD literature. Each of these are given a chapter of the thesis in which to explore how the viewpoint can contribute towards a set of criteria that could be used to assess whether or not SD should be used to analyse D&D in any individual project for litigation. These explorations are carried out in chapters 3,4 and 5. The data that was gathered from the researcher's experience in the two projects introduced in section 1.3 is also used to aid the explorations. By using the data from the two projects that the researcher was integrally involved in, any assertions made by the SD literature can be tested using the practical experiences from modelling D&D. To aid the reader's understanding of the data taken from the two projects, a detailed background of the projects is covered in chapter 2.

A summary of chapters 3, 4 and 5 is given at the end of chapter 5. This summary includes a revised set of criteria that represents the most informative criteria taken from the three viewpoints in the SD literature. This is intended to form a practical set of criteria to help modellers decide whether or not SD should be used to analyse D&D in any individual project for litigation.

The thesis then focuses on the second aim:

- *to explore the issues involved in using SD to analyse D&D for litigation.*

This involves gaining an understanding of what SD can achieve when modelling D&D and an appreciation of the limitations that exist when SD is used for this type of analysis. For these explorations, the purposes of modelling D&D are defined and considered. Any modeller needs to understand the purposes behind modelling a situation. These purposes will provide a target for a modelling approach to achieve. Therefore, by exploring how far SD can go to meet the targets set by modelling D&D, a modeller should be able to gain a good appreciation of the issues involved in using SD to analyse D&D. This process should be able to highlight the limitations of SD in this environment and the particular areas that the modeller needs to be aware of during the modelling process. These explorations are carried out in chapters 6, 7 and 8. Chapter 6 explores each of the modelling purposes and highlights those that need particular attention. The highlighted purposes are then considered further in chapters 7 and 8.

Once both of the main aims of the thesis are explored in chapters 3 – 8, chapter 9 summarises the conclusions from the research. This chapter details how the research has helped to further the knowledge in the area of the use of SD in the analysis of D&D for litigation. Of course, the conclusions to any research can only be regarded as being useful when the limitations of the research are understood. The limitations to this research are therefore also covered in chapter 9. Finally, chapter 9 concludes with areas of potential further research that have been highlighted from the conclusions reached.

Based on the above structure, one of the aims of the next chapter is to provide the reader with a detailed understanding of the two projects that the author was involved in and provide the data that is used to further the explorations carried out in this thesis. Chapter 2 will also cover an introduction to SD and a discussion on the definition of D&D so that the reader can fully appreciate the issues that are covered in later chapters of this thesis.

CHAPTER 2: BACKGROUND INFORMATION

2.1 Introduction

The aim of this chapter is to provide the reader with background information that is required to enable a full understanding of the remaining chapters of this thesis. Two separate areas of information are covered in this chapter:

- Background information on the two projects that provided the data for the explorations carried out for this research. Chapter 1 discussed the author's involvement in these two projects as a consultant hired to analyse D&D that occurred during the projects. These two projects provided a privileged opportunity for the collection of data to aid explorations into the issues involved in using SD to analyse D&D for litigation. It should be noted that pseudo-names will be used when referring to the two projects for confidentiality reasons.

- An understanding of SD and D&D. In section 1.2 it was stated that the aim of this thesis is to explore both SD and the modelling of D&D in an attempt to discover what each topic can inform us about its use alongside the other topic. Therefore, before exploring each of these topics, some background information will be provided. An introduction to the SD modelling approach will be given. Also, since *D&D* can be used to represent many different complex events that can occur during a project, there will be a discussion to clarify what is meant by this term.

2.2 The Pirate Project

The first project that the author was involved in was the Pirate project.

The Pirate project involved the modification of 28 aircraft. The client for the project hired the organisation that originally manufactured the aircraft that were to be modified as the main contractors for the work. The contractor then sub-contracted the main airframe modification work to a family run business who had experience in the modification of similar types of aircraft. A further reason for selecting this sub-contractor was that it was situated in an area of low employment and political intervention persuaded the client and the contractor that the project would be beneficial for the area.

The bidding of the project did not follow standard procedures. Although the contractor carried out an estimate of the work involved in the project this was not submitted for consideration by the client. Instead, the client informed the contractor of the amount of money they were willing to spend on the project. The contractor accepted the client's offer without a full appreciation of the nature of the work that was required in order to complete the project. The agreed contract was a fixed price contract. These circumstances were one of the main reasons for an underestimation of the budget for the project.

Throughout the project, the aircraft were purchased from various owners. These aircraft were then transported to the facility where the modification work was carried out. Due to the different owners and usage of the aircraft, the condition of the aircraft varied immensely. This meant that the amount of individual work required on each aircraft was far larger than had been anticipated. Although there was provision made in the contract for the cost of additional work, the individual work on the aircraft meant that the amount of learning that could be gained between each aircraft was reduced.

Retaining labour during the project proved to be difficult. Originally, one of the reasons that the location was chosen was due to the low employment history in the area. However, it was also an area that had not proved attractive to skilled workers. For these reasons, if employment became available in other areas of the country, workers had no hesitation in moving to the other work. Labour turnover was therefore far higher than in other parts of the country. This made it difficult to retain individual worker learning on the project.

Ten months into the project, the sub-contracted family business decided that they could no longer continue as a business and announced that they would be selling their facility. In an attempt to minimise the disruption to the contract, the contractor purchased the facility. Due to the inadequacies of the sub-contractor, the contractor had to put a lot of effort into trying to get the project back on track. For example, they had to bring in some of their own experienced staff, revisit schedules and bill of materials for the project. Also, a general re-organisation of the material and production workstations was required.

Based on the original contract, the project was scheduled to last for 36 months. However, after 48 months of work, the contractor was projecting a further 10 months of work.

By this time attempts to negotiate a schedule extension and additional finances for the project from the client had failed. The contractor had therefore begun litigation proceedings on the client in an attempt to seek compensation for their losses. A claim team was set up to drive the litigation process. This team consisted of:

- A senior manager from corporate office who was ultimately responsible for the decision-making during the claim process.
- A claim manager who was responsible for the day-to-day running of the claim.
- Lawyers who acted as the legal representatives for the organisation.
- The head of the group responsible for the work within the plaintiff organisation.

The senior manager involved in the claim team had been involved in a previous litigation process where an explicit analysis of D&D had been carried out. Based on this experience he felt that a similar analysis should be carried out for the Pirate project. For this reason a team of two consultants, including the author, were hired to analyse the D&D that occurred during the project. This analysis highlighted the following areas as the main disruptive triggers during the project:

1. An underestimation of the scope of work when estimating the number of hours required to complete the project.
2. Delayed corporate learning. This represents one element of the learning expected to be gained between products in production. Corporate learning represents the learning expected to be gained by the organisation rather than individual workers.
3. The unpredictable nature of the project meant that the management for the project was unable to gain a normal level of reduction in the time to produce the first product through planning actions carried out before the commencement of the project.
4. Learning was lost between products in production due to a higher than expected level of staff turnover.
5. Increased individuality between products. This caused a reduction in the learning that could be achieved between products in production.

6. When corporate learning was achieved between products in production, it was at a rate that was slower than had been expected.
7. An increased number of non-routines (NRs) than had been expected. A NR represents additional work that is required on a product in order for it to meet the specifications laid down in the project contract. NRs were normally identified by the plaintiff when either stripping or rebuilding the product. These were caused by the condition of the product being of a poorer quality than had been expected at the beginning of the project.
8. Delays in approval of NRs by the client.

Other consultants were also hired to analyse the data recorded throughout the project. Their aim was to provide a more traditional analysis of the project overruns for the litigation case.

The final claim for the project was based on a 22 month schedule overrun and an overrun in costs of 113% of the original contract.

The litigation case for the Pirate project remains unresolved. The defendant and plaintiff have had informal discussions on the way forward for the litigation process, however no conclusions have been reached. To date, the plaintiff is awaiting the defendant to propose dates to proceed with more formal discussions in an attempt to resolve the claim for compensation.

2.3 The Castle Project

The second project that the author was involved in was the Castle project.

The Castle project involved the modification of three aircraft. The prime contractor for the project hired a sub-contractor to modify the three aircraft to enable the contractor to fit new equipment into the aircraft. The parties negotiated and signed a fixed price sub-contract. The schedule for the work was expected to cover a period of 57 months.

The three aircraft to be modified were owned by the client for the project. One of these aircraft had been designated as a prototype for the work. However, due to problems with each of the aircraft, the aircraft that was designated as the prototype was altered on a number of occasions. As a result, the sub-contractor was unable to reap the benefits of a prototype since the three aircraft in effect ended up being virtually worked on in parallel.

A major issue that the sub-contractor had with the project was the constant changes in the work required. The prime contractor requested an unreasonable number of changes throughout the duration of the project. For example, there was in excess of 1100 configuration change requests (i.e. changes to the designs) of which 80% were triggered by the client. This compares to an expectation of approximately 50. This meant that workers had to regularly undo and redo work. Interviews with staff, who had many years of experience in aircraft modification, showed that the staff had never worked in a project that had required such extensive rework to be carried out. They stated that at times they were unwilling to carry out work since they were sure that a change would occur which would mean they would have to do the work all over again.

Fifty-eight months into the contract, the sub-contractor was projecting an overrun of 25 months on the project. At this time, the client for the project decided that the project could not be completed based on the agreed scope of work due to the cost overrun that was being experienced. For this reason, the client decided that a re-scope of work was necessary. This re-scope replaced the original design with one that would mean that the capabilities of the final product would be significantly reduced. This re-scope resulted in the sub-contractor requiring to undo work and carry out further design work, before then progressing onto the new work.

The sub-contractor believed that a major cause of the cost and time overruns on the project were triggered by actions taken by the prime contractor. After the prime contractor refused to accept any blame for the project overruns, the sub-contractor began litigation proceedings on the prime contractor in an attempt to seek compensation for their losses. The senior management for the sub-contractor set up a claim team to drive the proceedings. This team consisted of:

- A senior manager from corporate office who was ultimately responsible for the decision-making during the claim process.
- A claim manager who was responsible for the day-to-day running of the claim.

- Lawyers who acted as the legal representatives for the organisation.
- The president of the group responsible for the work.

Consultants were also hired to analyse the data recorded throughout the project. Their aim was to report on the areas where the additional labour hours and time overrun had occurred and provide supporting data for the litigation case. They also provided a more traditional analysis of the project overruns through the use of Critical Path Analysis (CPA). This tool can be used to deconstruct a project into a network of activities before proceeding to evaluate the expected project duration (Lockyer and Gordon 1991).

Based on the senior manager's previous experiences in litigation cases he felt that an explicit analysis of D&D should be carried out for the Castle project. For this reason a team of two consultants, including the author, were also hired to analyse the D&D that occurred during the project. This analysis highlighted the following main areas as the main disruptive triggers during the project:

1. A reduction in expected learning benefits between products. This was due to an unexpected amount of individuality between the products and various disruptions in the workflow resulting in a reduced ability to gain learning from one product to the next. This also occurred due to the lack of a prototype on the project.
2. The defendant was late in supplying customer-furnished information which was vital to enable the plaintiff to progress design work as expected.
3. The defendant was late in supplying customer-furnished equipment. The plaintiff was then required to accept temporary substitutes that were on loan from the defendant for production purposes. However, differences between these substitutes and the final equipment resulted in the need for changes late in the production process.
4. Delays in the resolution of problem trouble reports (PTRs). A PTR occurred when the plaintiff came across an issue in production that required additional information in order that production could proceed. The plaintiff waited unreasonable lengths of time for information from the defendant and therefore the production work could not proceed as expected.
5. An excessive number of changes causing both disruption and an increase in work scope by more than 30%.

6. The disruptions caused by each of the above five categories were of great concern to the plaintiff as their requests that the project schedule be lengthened were refused. Instead, the defendant told the plaintiff to use work-arounds throughout the engineering and production processes. Work-arounds meant that the plaintiff was required to make assumptions about missing information or equipment. Some of these assumptions later proved incorrect when the correct information or piece of equipment arrived from the defendant.

The final claim for the project was based on a 13 month overrun in schedule (the re-scope had reduced the original 25 month overrun) and an overrun in costs of 131% of the original contract.

To date, the litigation case for the Castle project has not been resolved. The plaintiff has tried, unsuccessfully, to negotiate a resolution with the defendant. Currently the claim team is working towards a hearing that will be presided over by an arbitrator.

The aim of this chapter is to provide the reader with background information that is required to enable a full understanding of the remaining chapters of this thesis. Two separate areas of information were identified as follows:

- Background information on the two projects that provided the data for the explorations carried out in this research.
- A discussion on the meaning of D&D and an introduction to SD modelling.

Since background information has now been provided on the two projects, this chapter will now turn to a discussion on D&D and will end with an introduction to SD modelling.

2.4 What Is D&D

The aim of this research is to firstly explore whether or not a specific modelling approach, i.e. SD, should be used in the analysis of D&D and secondly to then explore the issues involved in using SD to analyse D&D. Therefore, to carry out a thorough exploration, it is essential that a full understanding is gained of what is actually meant by the term *D&D*.

2.4.1 D&D in the Literature

The two words *disruption* and *delay* have their own meanings. King and Brooks (1996) define *disruptions* as events that preclude the contractor from completing the work in a manner in which the work was bid. They also define a *delay* as an increase in time needed to complete the project beyond what was contemplated at the time the contract was signed. However, a simple analysis of the direct consequences of each disruption and each delay will not cover the full impact of D&D in a project. Instead, each of these events will act as triggers of D&D.

Eden et al (2000) fully discuss the nature of D&D. A summary of the main points of the discussion are as follows:

- The direct impact of a disruption or a delay will not cover the overall cost of D&D. The ramifications of D&D are complex. Indeed, apparently insignificant disruptions to a project can have significant ramifications. It becomes difficult to attribute 'knock-on' events and delays directly to any one disruption or delay.
- Some disruptions are expected at the beginning of the project, they may simply occur in excess during the project, whilst other disruptions are completely unexpected.
- The acceptance of change orders is probably one of the most common disruptions to large complex projects. This is due to:
  - The difficulties in estimating the true cost of D&D associated with a change order.
  - The impact of a gathering sequence of change orders.
  - The client's lack of appreciation of the impact of D&D from change orders means that the contractor will have a difficult time when attempting to ask for additional time or money.
- CPA has traditionally been used to identify the impact of delays to a project. However, when evaluating D&D, there are a number of shortcomings of CPA. For example, CPA does not take account of management actions to deal with D&D.
- D&D can create a portfolio effect. This occurs when the impact of many disruptive events taken together is much greater than the impact of the sum of each individual impact.
- The most significant circumstance for major D&D occurring is when there is a disruption that causes an expected delay in the project and management then take

action in an attempt to accelerate the project. The action taken then has the serious consequences of causing more disruption and delay, which itself causes more disruption and delay, and so on. This creates a feedback phenomenon in D&D.

- There are difficulties in drawing boundaries between *direct* and *indirect* consequences of D&D. One possible definition given is that direct consequences can be easily thought through, whereas indirect consequences cannot be easily thought through due to the human difficulties with identifying and thinking through feedback loops.

Eden et al (2000) attempt to provide the reader with an appreciation of the different elements of D&D and the complexities involved in attempting to analyse it for any project.

In the above summary, it was noted that one of the most significant circumstances for D&D arose from a feedback phenomenon. For this reason, D&D analyses often focus on feedback. A feedback loop is formed when causal relationships form a loop with all the relationships flowing in the same direction. Feedback loops are well documented in the SD literature (for example Richardson 1991, Forrester 1961, 1968a). Sterman (1989a, 1989b) shows that people find it very difficult to grasp the complexities of such loops and so tools are normally required to aid their understanding.

2.4.2 D&D in Litigation

The term D&D has been used in litigation cases for large complex engineering projects to explain the reasons for time and cost overruns which cannot easily be explained through direct cause and effects (Cooper 1980, Weil and Etherton 1990, Ackermann et al 1997). The cause and effects are often significantly separated in time and space (Weil and Etherton 1990) and there is often a complex, long route of causality linking the final outcome with the initial reason. For this reason, when attempting to analyse D&D, it is important to gain an understanding of the causal paths that link the cause with the effect.

One of the objectives of litigation for projects that have overrun in both time and money is to attempt to attach monetary figures to the initial causes that triggered the cost overruns. Therefore, the onus is on the plaintiff to demonstrate that they can account for £x due to cause y. This means that the initial cause of each outcome has to be identified. However, the feedback nature of D&D means that it can prove complex when attempting to identify which event initially triggered which outcome. Also, if one feedback loop interacts with another feedback loop, the overall outcome will be greater than the summation of the individual outcomes. As each loop interacts with another loop, the overall dynamic behaviour results in an exponential growth in the overall cost to the project (Rodrigues and Bowers 1996). Complexity levels rapidly increase as the number of feedback loops that exist also increase. Indeed when D&D is rampant, managers will report that the resulting complexity is akin to chaos (Eden et al 2000). When modelled, large complex projects can show hundreds of feedback loops and thus involve great complexity.

The complexity of all the feedback loops interacting with one another will cloud attempts to trace the initial cause of an outcome. Eden et al (2000) noted that a disruption can cause a delay that can go on to cause a further disruption and so on. Thus it can prove very difficult to discover exactly which delay or disruption actually initiated the resultant D&D. Also, as already noted, when a delay is expected to occur, management may take actions to compress the project to avoid the delay (Cooper 1994, Eden et al 2000, Howick and Eden 2001). However such actions usually cause further D&D and exacerbate the disruption to the project. Thus, the resultant D&D may be blamed on the managerial decision-making. However, it could be argued that this only occurred due to the initial disruption and delay that caused the expected delay to the project. Therefore, any analysis has to be able to trace D&D back to its *initial* trigger.

Any analysis of D&D needs to be able to provide some form of evidence linking cause y with the eventual outcome of an overrun of £x. Due to this need to provide evidence of linkages between causes and outcomes in a litigation case, no matter what analysis approach is used, it is reasonable to assume that the modeller should begin by gaining a clear understanding of the logic that drives the project overruns. This will therefore mean gaining a full understanding of the D&D that occurred during the project.

2.4.3 The Purposes of Modelling D&D for Litigation

For both the Pirate and Castle projects, the purposes of the modelling process were discussed between the modellers and the claim team early in the claim process. The purposes that were agreed were consistent with the negotiated purposes that have been agreed on a number of other projects with which the author has been associated. The purposes are also not in disagreement with the literature describing specific cases of modelling D&D for litigation (Cooper 1980, Weil and Etherton 1990). Therefore, based on these experiences, the general modelling of D&D for litigation can be summarised as follows:

1. Demonstrate that a part of the time and cost overruns were caused by D&D through particular triggers of D&D.
2. Replicate over time the hours due to D&D which were over-and-above those that were contracted, but were required to carry out the project.
3. Demonstrate the extent to which the client management of the project was reasonable and the extent that overruns could not have been reasonably avoided.
4. All the above have to be demonstrated in a way which will be convincing to the several stakeholders in a litigation audience.

Defining the purposes of modelling a situation is an important step in any modelling process. Therefore, the above modelling purposes play an important role in this thesis. They will be referred to throughout the thesis and, in particular, are the main focus of the second half of the thesis when the second aim is considered i.e. the exploration of the issues involved in using SD to analyse D&D for litigation.

2.4.4 D&D Analysis in the Castle Project

The first piece of analysis carried out by the consultants brought in to analyse the D&D for the Castle project was the modelling of the events of the project. To do this, the consultants used information provided by the plaintiff in the form of a draft claim document for the project. The consultants used a *cause map* for this purpose. Cause maps are directed graphs and they follow the same principles as cognitive maps (Eden et al 1983). Whereas cognitive maps are intended to map the thinking of an individual, cause maps do not 'belong' to any single person but rather are an aggregate of the thinking of many people (Eden and Ackermann 1998). They are therefore characterised by an hierarchical structure which is most often in the form of a means/ends graph. However, the maps may contain circularity in which loops are created. When presenting a cause map, short pieces of text are linked with arrows. A statement at the tail of the arrow is taken to cause the statement at the arrow head (Eden et al 1992).

From interviews with various staff involved in both the project and the litigation process, the cause map was validated. Analysis of the cause map enabled the map to be collapsed and the main feedback loops to be highlighted. The main triggers of these feedback loops were also highlighted.

When a cause map is collapsed in this way, an influence diagram may be developed (Ackermann et al 1997, Eden 1994). In this context an influence diagram focuses attention on the concepts that are "... amenable to quantitative judgement and important in constructing a model of the feedback dynamics" the concepts also have "...a tighter phraseology and closer to that of variables" (Eden 1994, p269).

The term influence diagram has been used in SD to represent another form of qualitative diagram (Coyle 1977). However, Wolstenholme and Coyle (1983, p569) note that the term influence diagram appears in other fields of systemic analysis (Eden et al 1979) other than SD. Indeed, Eden et al's (1979) meaning of the term influence diagram which Wolstenholme and Coyle refer to is the type of diagram described in this work. The use of Eden et al's type of influence diagram as an appropriate pre-cursor to SD modelling was identified by Eden et al in 1983.

The cause map that was constructed for the Castle project was relatively small (127 concepts) compared to cause maps constructed for other litigation cases with which the author has been associated. The size of the cause map meant that it was more manageable than is normally the case. Therefore, the modellers did not construct an influence diagram but, instead used the collapsed cause map as the link between the quantitative and qualitative models. The concepts contained in the collapsed cause map are more descriptive than those normally contained within an influence diagram and the cause map enables quicker access to background information contained in the original cause map.

The collapsed cause map was used as a pre-cursor for a quantitative SD simulation model. Once the cause map was translated into a computer-based SD model, the causes and effects of the triggers of D&D could be quantified. The SD modelling approach is discussed in section 2.5 below.

It should be noted that the cause map for the Castle project was created with the use of a software package called Decision Explorer[1]. Extracts from the cause map created with this software will be used as part of the explorations throughout this thesis. When reading the material contained within these extracts the following points should be noted:

- Each concept that appears on the cause map will commence with a number. This number is a reference number used by the software and has no significance to the material to which it is attached.
- Some concepts will appear in different fonts. These fonts relate to different styles used by the software so that the concepts can be grouped by the type of information they are conveying. For example, one style may be used for an event that represents a trigger of D&D whilst another style may be used to represent contextual information.

2.4.5 D&D Analysis in the Pirate Project

During their involvement in the Pirate project, the consultants brought in to analyse the effects of D&D created a cause map reflecting the events that occurred during the project. For this project, the cause map was created entirely from cognitive maps created from interviews with staff involved in the project and members of the claim team.

When exploring the causes of the overrun in the Pirate project, many of the disruptive triggers were found to be due to an underestimation of the budget for the project. If the modellers were to consider the use of a SD model to model D&D in this project, the hours of underestimation, due to the triggers in the Pirate project, would need to be treated as inputs to the model. The consequences of D&D from the additional input hours could then be discovered from the output of a SD model. Therefore, before any SD modelling could be considered, an analysis of the amount of additional input hours was required.

A spreadsheet analysis was used where the additional hours required on each aircraft due to each disruptive trigger were estimated. This approach resulted in a number of additional hours for each disruption for each aircraft. When these additional hours were compared to the actual hours spent on the project it was noted that a large amount of the overrun hours were captured by this approach. The spreadsheet analysis therefore concluded that a large proportion of the cost overrun could be taken account of without any analysis of the

[1] Decision Explorer is a proprietary product of Banxia Software Ltd, Glasgow, UK.

indirect effects of D&D. As it appeared that the amount of hours due to the effects of D&D may have been minimal, it was concluded that the effort involved in constructing a simulation model to fully explain the effects of D&D outweighed the potential benefits.

The analysis approach taken for the Pirate project was therefore different to that for the Castle project. The projects therefore provide an opportunity for exploring the differences between the two projects as well as providing very different environments to explore the questions that will be raised during this thesis.

Now that a discussion on D&D has completed, this chapter will end with an introduction to SD modelling.

2.5 System Dynamics

The SD modelling approach evolved from work at M.I.T. in the late 1950's (Forrester 1958). In 1961 Jay Forrester published *Industrial Dynamics* (Forrester 1961) in which he describes a new computer-based modelling approach as "... the investigation of the information-feedback character of industrial systems and the use of models for the design of improved organizational form and guiding policy" (p13). The principles behind the computer-based *Industrial Dynamics* modelling approach were based on control engineering.

Forrester saw a need for this new approach because he believed that management science had previously failed to assist top management. He therefore saw a need for an approach that would aid management in their task of designing and controlling systems (1961). The approach should be able "...to analyze the principal interactions among all the important components of a company and its external environment" (1961, p8).

Industrial Dynamics (Forrester 1961) was defined as a modelling approach to explore the behaviour of industrial systems. Since this early publication, the modelling approach has been used to explore the behaviour of many different types of social systems. Its name has therefore become known as *System Dynamics* to reflect the wider areas of applications.

When using the SD modelling approach, Forrester proposed that the following steps should be undertaken:

1. Identify a problem - this helps to form the purpose of the model.
2. Describe the system - the elements of the system relevant to the problem should be identified verbally along with their interrelationships.
3. Construct a mathematical model - the verbal descriptions should be translated into mathematical formulae.
4. Simulation - the mathematical model should be simulated to generate the behaviour through time of the system as described by the model.
5. Interpret the output of the simulation - compare the results of the simulation against knowledge about the real system. Revise the model until it is deemed to be an acceptable representation of the real system.
6. Redesign the system - make changes to the model in an attempt to gain improved system behaviour. Simulate the model after each change to discover the change in system behaviour.
7. Implement the changes - translate the changes that indicate improved system behaviour into the real system.

Since Forrester's early publications on SD, the steps in the modelling approach have been expanded upon by many other authors in an attempt to provide fuller definitions (Forrester 1968a, Coyle 1977, Randers 1980, Richardson and Pugh 1981, Roberts et al 1983, Wolstenholme 1990, Coyle 1996, Sterman 2000).

One of the propositions made for the SD modelling approach has been a split in the approach between qualitative and quantitative analysis (Wolstenholme 1982, 1990). This advancement was aimed at enhancing the qualitative phase of system description. Forrester's original convention for system description (the use of sources and sinks and valves) shifted to the use of diagrams known as influence diagrams (Coyle 1977), causal loop diagrams (Randers 1980, Morecroft 1982) or signed digraphs (Wolstenholme 1982).

The use of qualitative tools for system description has resulted in differing opinions in the SD community. There are authors who believe that the qualitative analysis phase is a separate stage to the quantitative analysis stage and can provide its own conclusions (Coyle

1996, 2000; Wolstenholme 1990, 1999). However, other authors believe that qualitative tools are important, but are not a requirement for quantitative simulation and, if used without quantitative tools, can in fact be misleading (Sterman 1994, Richardson 1996, 1999).

During the projects discussed in this thesis, the system description stage of the analysis has been carried out using cause maps. The use of this type of qualitative tools as an appropriate precursor to SD modelling is discussed by Eden (1994), Ackermann et al (1997) and Bennett et al (1997).

2.6 Summary of Chapter

This chapter provides background information on two projects that will be continually referred to throughout this thesis. A discussion on the meaning of the term *D&D* and an introduction to SD modelling have also been covered. Discussion of this background information was deemed necessary so that the reader may fully appreciate the explorations that are undertaken in the remainder of this thesis.

Now that the reader has sufficient background knowledge, the next chapter can commence explorations that help achieve the first aim of this thesis. Section 1.2 defined the first aim of this thesis as:

- *to explore the appropriateness of SD as a modelling approach in the analysis of D&D.*

The SD literature was highlighted as a resource that could be used to progress these explorations. Chapter 3 will now focus on this work.

CHAPTER 3: EXPLORING CRITERIA DETAILED BY FORRESTER IN THE SD LITERATURE

3.1 Introduction

Following section 1.2, the aim of the following three chapters is:

- *to explore the appropriateness of SD as a modelling approach in the analysis of D&D for litigation.*

The resource chosen to investigate this issue is the SD literature. The literature will be used to explore what criteria, if any, have been defined as those that should be used to identify whether or not a situation can be modelled using SD. Data will also be used to explore these criteria. Section 1.3 discussed the author's privileged position of being directly involved in the modelling of D&D for two litigation cases. This privileged position means that a mass of valuable data, which is rarely made available to researchers for the reasons discussed in section 1.3, will be used to support the line of arguments made in this thesis. In following this approach, the next three chapters will enable an exploration into the types of situations the SD modelling approach was intended for and will further the investigation into whether or not it should be used to model D&D in a litigation environment.

In searching the literature, three different views of criteria have been chosen which need to be satisfied in order that a SD approach can be used to model a situation. The first two sets of comments are taken from people who have devoted the majority of their working lives to the use of SD rather than any other modelling approach. Early comments from the founder of SD, Jay Forrester, are used to form the first set of criteria. These criteria define the focus of this chapter. The second set of criteria contains more practical elements than the first and was laid down by Geoff Coyle in 1977. These will be explored in chapter 4. Finally, criteria set down by Robert Flood and Michael Jackson are examined in chapter 5. The third set of criteria is a contrast to Forrester and Coyle's criteria as the authors adopt a more multi-methodological approach in discussing their criteria.

The three sets of criteria have been chosen to cover a wide as possible set of views on the use of SD. A lack of literature covering this issue in detail narrowed the choice of criteria to use. Therefore, the traditional view of Forrester, a practical view by Coyle and a wider multi-methodological view by Flood and Jackson have been chosen. Section 3.2 will demonstrate that many other authors in the SD field hold a similar view to Forrester. Therefore, the views considered in the next three chapters certainly go well beyond the views of only three sets of authors.

3.2 Forrester's Criteria

As this thesis intends to explore what, if any, criteria have been defined as those that should be used to identify whether or not a situation can be modelled using SD, it is necessary to begin exploring the literature that introduced SD to the world. In *Industrial Dynamics* (1961), Forrester introduces SD (or Industrial Dynamics as it was originally named) as a method used to explore the behaviour of industrial systems in general. Since this early publication, the description has been revised to the application of feedback concepts to social systems (Forrester 1968c) to reflect the wider applications of the modelling approach. However, Forrester does not provide much detail in any of his work on criteria that could be used to assess the suitability of modelling a situation using SD. One of Forrester's views is that "everything we do as an individual, as an industry, or as a society is done in the context of an information-feedback system" (1961, p15; 1975, p54). As Forrester describes the first and most important foundation for SD as the concept of information-feedback, he is suggesting that virtually any situation involving individuals, industry or society can be modelled using SD. Another of his views (1968a, p4-5) is that "the feedback loop is the basic structural element in a system. Dynamic behaviour is generated by feedback". Also that "every decision is made within a feedback loop" (1968a, p4-4). Therefore, it can be concluded that we have dynamic behaviour in any situation where a decision is made. As this thesis is interested in managed projects, the managerial decisions involved in these projects would suggest the existence of dynamic behaviour. This suggests that SD *could* be used to model any of the managed projects of interest to this work. Indeed, the above discussion suggests that the SD modelling approach could be used to investigate *any* managed situation within an organisational setting.

If the above is accepted, then it could be concluded that it is not necessary to define criteria to consider when the SD modelling approach should be used, as it would appear that there are no *organisational* situations where the SD modelling approach cannot be used. Even if this far-reaching statement were true, it does not mean that the SD modelling approach *should* be used in all situations as it may not be the *most appropriate* modelling approach to take. An alternative conclusion is that the extracts taken so far from Forrester's work lack sufficient detail to enable them to be used in any meaningful way to further the explorations that are of interest to this chapter.

The difficulty in identifying a set of criteria is not only observed in Forrester's work. For example, Richardson and Pugh (1981, p2) also avoid a detailed discussion on the suitability of SD as a modelling approach. They address the issue of the suitability of the SD approach by stating that it "... applies to dynamic problems arising in feedback systems". This definition does not provide the reader with any additional information beyond that already discussed at the beginning of this section. It therefore does not enable any further insights to be gained into possible criteria to be used to assess the suitability of SD.

At this point, it is interesting to note Vennix's (1996, p104) point of view on the difficulty of defining a set of criteria to assess the suitability of SD to model a situation. He states that one of the most difficult questions asked, even to an experienced model-builder, is when to use SD. This implies that he believes there are situations when SD should not be used, but it is difficult, practically, to assess the criteria to be used. The difficulty that Vennix describes in defining criteria to assess the suitability of SD to a situation may be an explanation for why there has been a lack of it appearing in the literature. Legasto and Maciariello noted such a lack of criteria in 1980. They stated that "The present situation is such that nonadversaries, i.e. the neutral observers, find it difficult to assess the value of system dynamics (and competing methodologies) to them. Because no criteria have been developed for evaluating the comparative advantages and drawbacks of each methodology, it is also difficult to select, on rational grounds, an *appropriate* methodology for a given problem" (p23).

Thus, Forrester's work is no exception to the conclusion that the SD literature is lacking in criteria to assess whether or not SD should be used to model a given situation. However, when pushed by Ansoff and Slevin (1968) to demonstrate that SD could be considered as a body of theory, one criterion Forrester was asked to fulfil was to provide a "statement of its limitations, an implicit definition of areas of experience to which the theory does not apply" (p394). In reply to this, Forrester (1968b) attempts to define situations when the SD modelling approach should *not* be used. Forrester suggests that SD does *not* apply to problems that *lack systemic interrelationship.* It does *not* apply to areas where the *past does not influence the future.* It does *not* apply to situations where *changes through time are not of interest* (1968b, p605). Although Forrester introduces these criteria, he makes no attempt to fully explain any of them to the reader. In fact, he does not appear to return to this issue in any of his later writing either. This brief set of criteria is the closest Forrester comes to considering the issue of criteria to assess the appropriateness of using SD to model a situation. One reason for this may be because, by introducing SD to the world, he believed that a set of criteria describing the appropriate use of the SD approach was not necessary. He had introduced SD as an approach for the analysis of *any* managerial problem and therefore did not see any need to consider the suitability of its use in a situation.

The above three criteria (i.e. SD does *not* apply to problems that *lack systemic interrelationship,* where the *past does not influence the future* and where *changes through time are not of interest*) represent Forrester's view on when SD should be used to model a situation. However, many other authors in the SD field hold similar views regarding the use of SD to model a situation. For example:

- Roberts (1978, pxi, Italics added) defines SD as "...the application of *feedback control systems principles* and techniques to the modeling of social systems".
- Richardson and Pugh (1981, p2, Italics added) summarise the appropriate use of SD as a method that "... applies to *dynamic problems* arising in *feedback systems*".
- Meadows (1980, p55, Italics added) notes "If the problem is centred on *generic dynamic behavior* of a mostly *closed system,* if the variables include *motivations and goals,* if the validation includes *assessment of the realism of the model structure,* then it is a system dynamics model."
- Vennix (1996, p105-106) states that SD can be used to model problems that are *dynamically complex, long term,* and where it is *possible to generate a reference mode of behaviour* which may represent a problem.
- Barlas (1998) states that a suitable problem for a system dynamics study is a *good dynamic feedback* problem that is *not a static or open-loop problem.*
- Sterman (2000, p41-42) states that SD can be applied to any *dynamic system,* with any time and spatial scale.

Each of the above authors refers to systems that are *dynamic* and/or contain *feedback*. These concepts will be seen to be embedded in Forrester's criteria. In addition Meadows notes that the problem should be centred on mostly *closed systems* and that the variables should include *motivations and goals*. Also, Vennix notes a practical point that it should be *possible to generate a reference mode of behaviour* to represent the problem.

Due to the repetition of the concepts used to define the types of problems that SD should be used to model, Forrester's criteria are taken to represent a variety of views in the wider SD literature for the purposes of this chapter.

On first impressions, the three criteria laid down by Forrester do not provide much detail for the reader. However, as this is a set of criteria set out by the founder of the SD modelling approach, it is important to explore their usability as criteria to assess whether or not it is appropriate to use SD to analyse D&D for litigation.

3.3 Criterion 1: Systemic Interrelationship

After introducing the limitation that "SD does *not* apply to problems that *lack systemic interrelationship*", Forrester does not explain what he means by it. However, an insight into its underlying meaning can be gained from *Industrial Dynamics* (1961) where Forrester comments that, "...we can expect that the interconnections and interactions between the components of the system will often be more important than the separate components themselves" (p6). This quote highlights the systemic nature of the relationships between elements of a system. It could be concluded that in using the phrase *systemic interrelationship* Forrester is referring to a system for which the relationships between elements of a system are more important than the elements themselves. However, as highlighted by Roberts et al (1983) "the systems approach to studying systems emphasises the connections among the various parts that constitute a whole". Therefore, if a modeller chooses to take a systems analysis approach to a problem, it is the relationships between elements of the system that will be focussed upon and become the variables of interest in the analysis. As Forrester introduced SD as "a method of systems analysis for management" (1961, p9), by its very nature, SD will focus on the *relationships* between elements of a system.

Based on the above, any organisational problem will only *appear* to "lack systemic interrelationship" if an analyst of the problem *chooses* to study it using an approach that does not take a systemic view of the problem. Therefore, the only conclusion that can be made is that, by taking a systems analysis view, any organisational problem *could* be modelled using SD. However, in stating that "SD does *not* apply to problems that *lack systemic interrelationship*", Forrester surely should have had situations in mind that he considered would lack systemic interrelationship. In fact, by introducing SD as "a method of systems analysis for management" (1961, p9) it is reasonable to assume that Forrester should have had *managerial* situations in mind which would lack systemic interrelationship. However, he is not clear in stating what these situations would be.

It was suggested above that any organisational problem would only *appear* to "lack systemic interrelationship" if an analyst of the problem *chooses* to study it using an approach that does not take a systemic view of the problem. Therefore, when considering the modelling of D&D in a litigation environment, one way in which to consider whether or not SD is a suitable modelling approach is to ask the following question:

- *Is it necessary to adopt a systems analysis approach when modelling D&D for litigation?*

If the answer to this question is 'yes', then it is unlikely that the problem will "lack systemic interrelationship". To consider this question, non-systems analysis modelling approaches used as part of the Pirate and Castle projects will be considered. A systems approach has been defined by one author in the systems analysis field as "an approach to a problem which takes a broad view, which tries to take all aspects into account, which concentrates on interactions between the different parts of the problem" (Checkland 1981, p5). Each of the modelling approaches that will be discussed from the Castle and Pirate projects are concluded to *not* have taken all aspects of the problem into account and *not* concentrated on the interactions between the different parts of the problem. For these reasons, they have been concluded as non-systems analysis approaches.

It should be noted that all but the last of the modelling approaches discussed in this section were carried out as a separate exercise from the modelling carried out by the author. The modellers involved in these first few analyses that will be discussed *made no conscious* *acknowledgement* of whether or not they would undertake a systems or non-systems analysis approach. Indeed they had no formal training in systems analysis approaches. Their remit was to undertake analysis that they believed to be of benefit to the litigation process.

If it can be shown that non-systems analysis approaches were capable of modelling D&D in the Pirate and Castle projects, then it may be concluded that it is not *necessary* to adopt a systems analysis approach when modelling D&D in a litigation environment.

3.3.1 Pirate Project - Analysis of Non-Routine Work

During the modification of the aircraft in the Pirate project, non-routine work was continually discovered on each aircraft. This work represented unanticipated work to an individual aircraft due to a part of the aircraft being of poorer condition than was expected and therefore requiring replacement. The extent of the disruption caused by the discovery of a non-routine (NR) was highly dependent upon when it was discovered. For example, the defendant became increasingly fussy about the quality of the final product and often parts of the aircraft which the plaintiff believed to be of a sufficient quality were identified as a NR by the defendant only in the last stages of the production process. Such late discoveries of NRs caused large disruptions in the production process created from a large amount of additional work from undoing and redoing previously completed work. The impact of the discovery of a NR is similar to the impact of the arrival of a change order (CO) in many other large projects. A CO occurs when a client requests a change in the work being undertaken by a contractor. COs are one of the most common disruptions to large complex projects. However, the work required to understand the dynamic ramifications of a CO is extensive therefore its cost is usually underestimated (Eden et al 2000, p293). The type of ramifications from a CO are normally rework, the effect of compression caused by the additional workload and the portfolio effect from a sequence of COs (Eden et al 2000, p294).

It was expected that many of the NRs in the Pirate project had been under-bid, therefore during the analysis of the project, the plaintiff carried out an analysis of the hours of D&D associated with NRs. Initially, the NRs that occurred on the project were split into twelve categories. These were as follows:

- The NRs were initially split into three classes:
  - High impact (defined as being discovered during the testing period of production) – these were described by the plaintiff as normally stopping progress and deemed as having a critical impact on work.
  - Medium impact (defined as being discovered during the middle sections of production) – these were described by the plaintiff as normally interrupting scheduled work, causing out of sequence work and disrupting the labour force.
  - Low impact (defined as being discovered early in the production process) – these were described by the plaintiff as normally causing tasks to be rescheduled and causing some disruption to the labour force.
- Each of the above three classes were further split into four categories based upon the number of hours bid by the estimator for the NR (up to 3 hours, 3 to 8 hours, 8 to 30 hours and over 30 hours).

Once every NR had been placed in one of the twelve categories, a sample of the NRs that had been completed was randomly chosen. For each of the sample NRs, the following hours were compared:

- The actual hours spent dealing with the additional production work required due to the occurrence of the NR. These hours were taken from records logging work completed against a code identifying why the work was carried out.
- The expected hours to carry out the additional production work which was expected to occur due to the arrival of the NR. The estimator would have calculated these hours when the NR was first discovered.

The difference between the two hours would represent the percentage of labour hours under-bid for the NR. After determining the percentage of hours under-bid for each of the NRs in the random sample, an average percentage under-bid for each of the twelve categories was calculated. Each NR was then multiplied by the percentage under-bid for its category in order to calculate the overall number of hours under-bid for NRs for the project.

The hours that were included in the actual hours calculation in this exercise were those directly observable as a consequence of the NR. For example, the report detailing this work comments on the type of hours that are considered for a high impact, low hours NR

"...although the time to replace defective components is usually up to 3 hours, there is an associated requirement in most cases to re-function the system and re-close the areas that were opened for access. This was estimated to increase the overall time by a factor of 3".

The problem with using this type of approach to analyse NRs is that it will only include those hours which can be directly linked to the NR at the point in time that it was carried out. For example, the total amount of additional hours that will accrue due to project compression caused by the NR may not be captured. To help explain this, suppose that a number of NRs occurred without any additional time extension being given to the project. If these changes affected the critical path for the project, the overall length of the project may increase. However, if the project management team did not want the project to overrun, managerial actions would be needed in order to compress the project. Although actions would be taken to aid the progression of the project, various side effects may occur which can have a detrimental effect on productivity levels (Howick and Eden 2001, Cooper 1994). This would result in a slower rate of productivity in work carried out on both:

- tasks associated with carrying out the NRs and
- other tasks which are not associated with the NRs.

Since the NR analysis described above only looks at the tasks associated with the NRs when summing up the total hours actually spent on the project, it would not capture the additional hours associated with the reduced level of productivity on tasks which are not associated with the NRs. The analysis would therefore underestimate the additional labour hours that were used due to the occurrence of the NRs.

The inadequacy of the NR analysis in not taking account of the increased hours due to tasks which were not associated with a NR, was noted by the plaintiff. In an attempt to take these additional hours into account, the plaintiff undertook further analysis. This analysis considered a sample of the aircraft being modified in the project. For each aircraft in the sample, the percentage overrun in hours on each production work station was compared to the number of NRs dealt with in the work station. The plaintiff hoped that this analysis would show that as the number of NRs increased, the percentage overrun in each work station also increased. This could have provided evidence to support the statement that the time to carry out all the tasks in the work station was affected by a NR and not simply those tasks immediately associated with a NR. Not surprisingly, a conclusive correlation could not be found. There were many disruptions occurring throughout the life of the Pirate project and although it is likely that the level of NRs in a work station affected the number of hours overrun by the project, this would not have been the only variable to influence the overrun. The disruptive effect of NRs would need to have been a dominating influence among all the variables disrupting the work undertaken in a work station in order for a significant correlation to be found.

In conclusion, neither of the two analysis described above were capable of capturing all the effects of D&D caused by the discovery of the NRs.

3.3.2 Pirate Project - Productivity Levels

Another short piece of analysis carried out for the Pirate project, which focussed on levels of productivity, was an attempt to find a correlation between the percentage of work carried out outside the work station where it was intended to be performed and the productivity levels of the labour force. Using a sample of the aircraft modified in the project, this analysis was carried out for the total production work on an aircraft and for the work completed at each production work station during the modification process. This piece of analysis was carried out by a consultancy firm who has a specialism in the analysis of large projects for litigation purposes. In the analysis, productivity was defined as earned hours of work divided by actual hours of work. The *earned hours* (also known as earned value (Fleming and Koppelman 2000)) for an activity can be defined in simplistic terms as the "percentage completion of an activity multiplied by the budgeted cost (of the activity)" (Raby 2000). Therefore, the earned hours for a project at any point in time can be represented by the number of hours that were expected (based upon the budget) to have been carried out to reach the current level of progress in the project.

Correlation calculations such as these can aid in the demonstration of relationships between the events and outcomes of a project. These are used as an approach to build up a picture of a variety of relationships between causes and effects during the project. However, the problem arises that one-to-one relationships such as these are neither easily observed nor statistically proven as there are so many different events occurring during the project which, in this case, could be affecting the productivity levels of the labour force. In the Pirate project, the correlations could not be successfully proven. This analysis was therefore not capable of explaining the cause and effects arising due to D&D. However, correlations have been used as part of modelling undertaken for litigation cases in the past and have proven successful in helping convince a courtroom of cause and effect relationships in a project (Nahmias 1980).

3.3.3 Castle Project - Earned Hours Analysis

The term *earned hours* was introduced as a part of the Pirate productivity levels analysis discussed above. An earned hours analysis also formed part of the analysis of the Castle project and was undertaken by a consultancy firm who has a specialism in the analysis of large projects for litigation purposes. Through discussions with the consultancy firms used on both the Pirate and Castle projects, it is clear that the earned hours analysis approach is routinely used in the analysis of large projects for litigation purposes.

The calculation of earned hours involves a comparison of the actual labour hours that have been spent on a project at any point in time and the labour hours that were expected to be completed based on the progress of the project. When choosing the point in the project to make such comparisons, appropriate milestones are often chosen where identifiable disruptions occurred during the period leading up to the milestone. These disruptions are then used to explain the overrun in hours over the specific period in time.

The disadvantage of taking this type of approach is that all that it actually proves is that there were additional hours worked during a period of time and that there were also disruptions during this period of time. However, a large amount of documentation between the defendant and the plaintiff is normally also required to help support the causality between the disruptions and the additional labour hours. Nahmias (1980) discusses a similar piece of work that was carried out in respect of a shipbuilding project. This work looked at the existence of a correlation between variables such as levels of worker efficiencies and the number of apprentices working on a project. Nahmias (1980, p8) noted that the correlations he discovered do not themselves establish causality "However, the analysis provides a means of corroborating other testimony from shipyard personnel indicating that, in fact, this was the case." Although documentation and testimonies can be used to support the argumentation that certain disruptions contributed towards the overrun in labour hours determined by an earned hours analysis, this does not prove that the disruptions caused *all* of the additional hours. In particular, it does not take account of purpose 2 of modelling D&D for litigation i.e.

- *Replicate over time the hours due to D&D which were over-and-above those that were contracted, but were required to carry out the project.*

There is certainly scope for the defendant to argue that many of these additional hours were due, for example, to mis-management of the labour force by the plaintiff or to an initial underestimation of the scope of the work involved in the project. The earned hours approach was therefore not capable of *fully explaining* the effects of D&D in the Castle project.

However, discussions with the consultancy firms involved in both the Pirate and Castle projects have indicated that the earned hours approach *has* been successful in courtrooms in proving litigation cases. This is argued to be due to the relative simplicity of the approach that means that an audience can easily follow and understand the underlying logic. The approach can therefore aid the achievement of the following purpose of modelling D&D for litigation:

- *All the above have to be demonstrated in a way which will be convincing to the several stakeholders in a litigation audience.*

3.3.4 Pirate Project - Spreadsheet Approach

For this example, the reasons for the modellers choosing to carry out a systems analysis approach changes. For previous examples, it was stated that the modellers made no conscious acknowledgement of whether or not they would undertake a systems or non-systems analysis approach. Indeed they had no formal training in systems analysis approaches. Their remit was to undertake analysis that they believed to be of benefit to the litigation process. However in this example the modellers, of which the author was one, were fully aware of the differences between systems and non-systems analysis approaches and came to an informed decision that a non-systems analysis approach would be taken.

As described in chapter 2, the background to the Pirate project is very interesting. For political reasons, the bidding of the project did not follow standard procedures. The contractor did not carry out a detailed estimate of the work involved in the project to be submitted for consideration by the client. Instead, the client informed the contractor of the amount of money they were willing to spend on the project. The contractor accepted the client's offer without a full appreciation of the nature of the work that was required in order to complete the project. These circumstances were one of the main reasons for an underestimation of the budget for the project.

When exploring the causes of the overrun in the Pirate project, many of the disruptive triggers were due to an underestimation of the budget for the project. A full list of the triggers is detailed in section 2.2. If a SD model was to be constructed to model the D&D in this project, the hours of underestimation due to the triggers in the Pirate project would normally be treated as inputs to the model. The consequences of the D&D from the additional input hours could then be discovered from the output of the SD model. Therefore, before any SD model could be considered, an analysis of the amount of additional input hours was required.

Since the analysis only needed to consider the additional input hours caused individually by each disruption, a non-systems analysis approach was considered sufficient for the analysis. A spreadsheet approach was used where the additional hours required on each aircraft due to each disruptive trigger were estimated. This approach resulted in a number of additional hours for each disruption on each aircraft. When these additional hours were compared to the actual hours spent on the project it was noted that a large amount of the overrun hours were captured by this approach. The modellers were surprised by this outcome. The reason for their surprise can be observed in the cause map shown in figure 3.1 below. It should be noted that in this figure the term *O&A* refers to *Over and Above* work. This is another term for Non-Routine work.

This cause map demonstrates that the people who were involved in the project and who had contributed to the information contained in the cause map believed that D&D had been trigged by various events. The outcome of this was an anticipated slippage in the project schedule. Management then responded to this by taking managerial actions in an attempt to compress the project. The outcome of these actions were then believed to lead to a cost overrun of the project. Indeed, the most senior member of the claim team strongly believed that a large proportion of the overall overrun was due to the effects of D&D. However, the spreadsheet approach concluded that a large proportion of the cost overrun could be taken account of without any analysis of the effects of D&D. Section 3.4 considers why there may have been a lack of D&D in the Pirate project. As it appeared that the amount of hours due to the effects of D&D may have been minimal, it was concluded that the effort involved in constructing a simulation model to fully explain the effects of D&D outweighed the potential benefits.

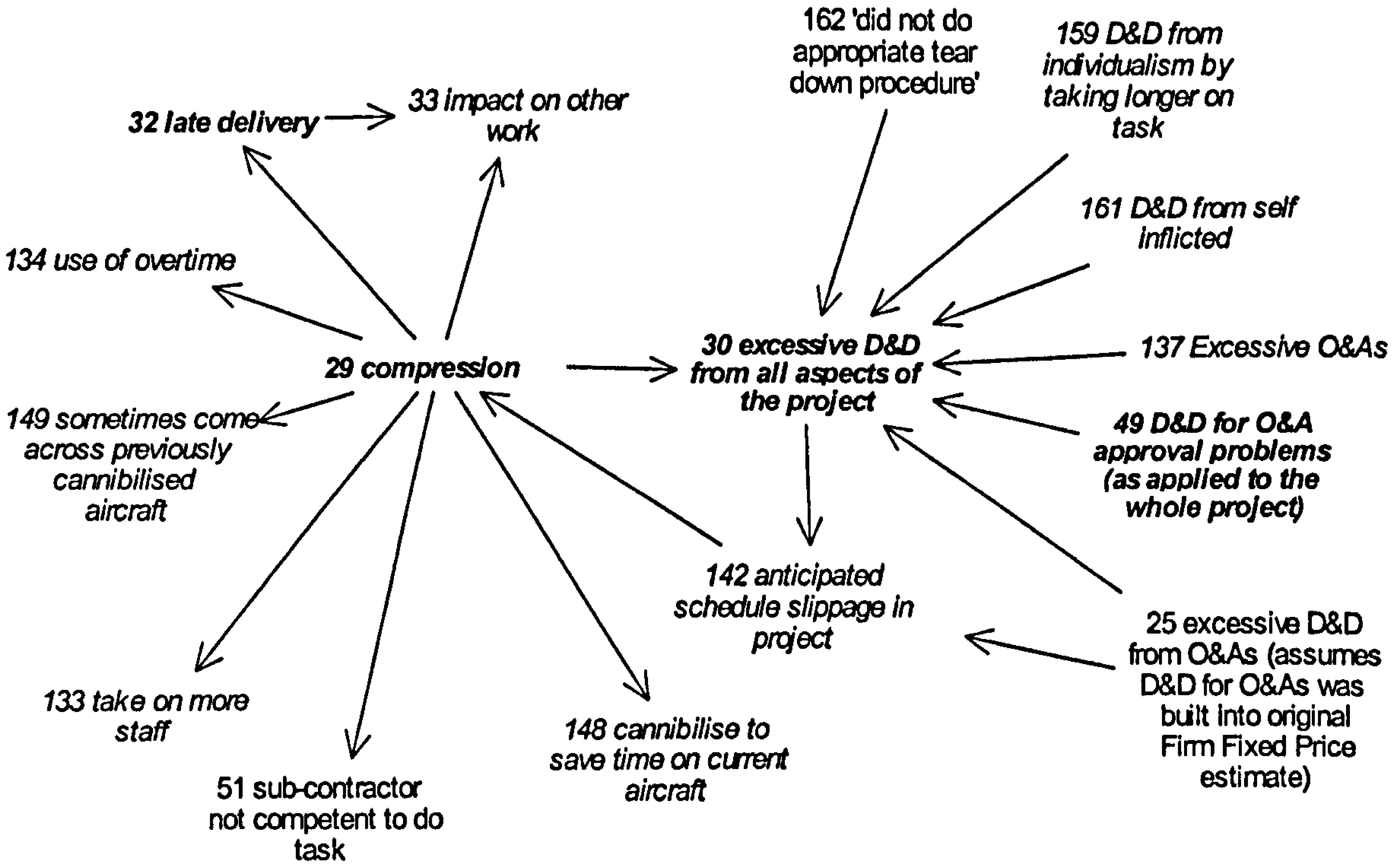


Figure 3.1 – The Effects of D&D in the Pirate Project

Since this section is considering whether or not non-systems analysis is capable of modelling D&D in a litigation environment, it is worth considering if the modellers *had* continued with the modelling of D&D in the Pirate project, whether or not D&D could have been modelled as part of the spreadsheet approach used to model the overrun in labour hours.

If the modellers had taken a quantitative *non-system* analysis through the use of a spreadsheet, one issue to consider is that they had already undertaken a qualitative *systems* analysis through the use of cause mapping. Cause mapping was used to summarise how those involved in the project believed the different elements of D&D were related to one another. The only way in which all the effects captured in the map could be taken account of in a quantitative model would be by taking a *systems* approach. Therefore, if the modellers were to attempt to model this with a *non-systems* approach, the modelling would either have to make an approximation to the impact of the various system relationships, or omit the impact of some of the relationships. For example, in the spreadsheet approach taken above, one of the causes of additional hours which was modelled was trigger category 4 which is detailed in section 2.2 as:

- *Learning was lost between products in production due to a higher than expected level of staff turnover.*

When this was modelled in the spreadsheet, data on the actual turnover experienced on the Pirate project was used as an input to the model. However, as can be seen in figure 3.2 below, labour turnover was believed to be an endogenous element of the system and therefore a systems approach would have attempted to reproduce the actual data, rather than use it as an input to the model.

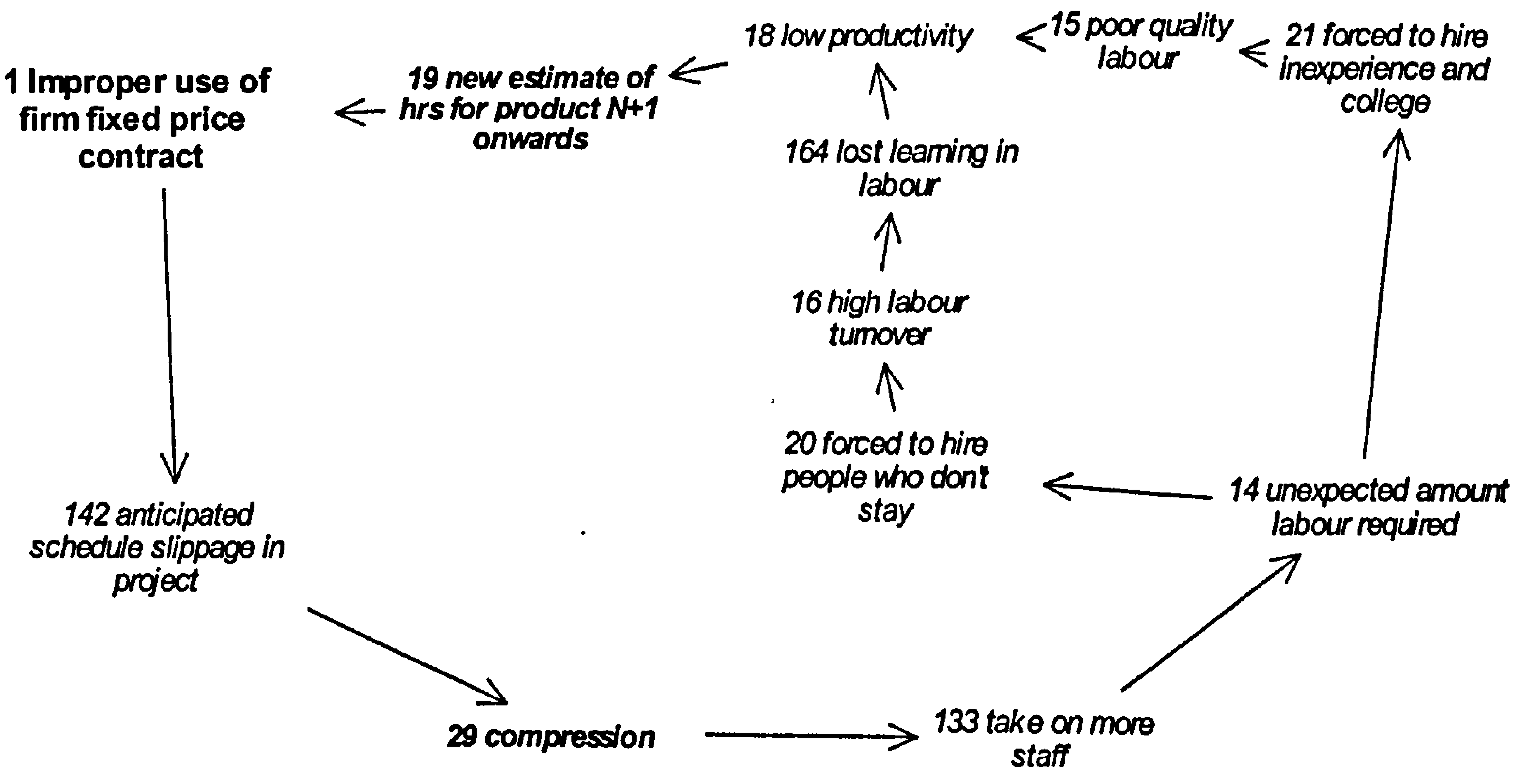


Figure 3.2 - The D&D Effects of Labour Turnover in the Pirate Project

To fully reproduce the effects of labour turnover that were experienced in the Pirate project in a quantitative model, the effects of project compression experienced on the project would also require reproduction. The spreadsheet approach mentioned above used the actual data for labour turnover as an approximation to the modelling of the relationships that are shown in figure 3.2. If an attempt was made to model the full effects of the relationships shown in figure 3.2 in the spreadsheet model, then the impact that compression has on the other disruptions taken account of in the spreadsheet would also need to be considered. For example, a reduction in productivity levels will affect the labour hours determined for the other disruptions modelled. These hours would therefore need to change every time that productivity changed. In fact all elements of the spreadsheet would continually require changing every time the feedback loops shown in figure 3.2 were traversed. Based upon its current design, the spreadsheet model only allows one entry per aircraft per type of disruption. However, the feedback loops in figure 3.2 are likely to be reinforced many times during the life of one aircraft. If the spreadsheet design was to remain unaltered, then an approximation which allowed for the effect of compression on each aircraft would be needed. However, if a fully explanatory model was required that provided a transparent analysis of D&D, then the only way in which this could be modelled in the spreadsheet is to alter the design of the model so that it becomes more time-dependant. The spreadsheet would need to be capable of determining the hours spent on each aircraft as time progresses, thereby capturing how the labour hours alter as productivity levels alter. This type of approach would mean that the spreadsheet analysis would in fact become a systems approach as it would be taking account of the full effects of the relationships described in the cause map.

This section has explained how a non-systems analysis approach was used to capture the majority of the labour hours that accrued over-and-above those contracted in the Pirate project. The additional hours modelled represented disruptive *triggers* of D&D rather than the *outputs* of D&D. As discussed above, this non-systems analysis approach would not have been capable of modelling the outputs of D&D in transparent manner. Only an approximation to the impacts of the various system relationships or even an omission of the impacts of the relationships could be achieved by taking the non-systems approach. However, it was noted that an issue with the modellers taking a quantitative *non-systems* approach is that they had already undertaken a qualitative *systems* approach through the use of cause mapping. In taking this qualitative systems approach, the only way in which all the relationships captured could be transferred into a quantitative model in a transparent manner would be by also taking a *systems* view of the relationships.

3.3.5 Summary of the Explorations of Criterion 1

In discussing Forrester's first criterion, it was suggested that *systemic interrelationships* would be highlighted if a *systems approach* were taken when analysing a situation. The focus of the section then turned to considering:

- *Is it necessary to adopt a systems analysis approach when modelling D&D for litigation?*

In order to consider this question, different forms of non-systems analysis approaches were discussed. Each of these was undertaken to analyse the project overruns experienced during either the Castle or Pirate project. It has been argued that each of these approaches does not either fully explain or capture all the effects of D&D in the projects discussed in this chapter. There is therefore evidence to suggest that it *is* worth pursuing a systems analysis view in these types of modelling situations. However, it is important to note that at least one of these approaches has been sufficiently convincing to a litigation audience that it has proven successful in a court of law. It has therefore been able to fulfil one important purpose of modelling D&D for litigation. This is in contrast to systems approaches that can be difficult for an audience to fully comprehend. Indeed, the complexity of SD as a modelling tool was a hurdle for the modellers in both of the projects discussed in this thesis. The modellers had to explore different approaches in presenting models to various stakeholders in order to aid their understanding.

The above suggests that both systems and non-systems analysis approaches have weaknesses when modelling D&D for litigation. However, if an approach could be developed to adequately explain the complexities of a systems analysis approach such as SD to the various stakeholders, then this would aid its use in a litigation environment.

3.4 Criterion 2: Past Influences the Future

After introducing the limitation "SD does *not* apply to problems ... where the *past does not influence the future*", Forrester does not explain what he means by it. However, based upon other work, it seems reasonable to suppose that Forrester is referring to the existence of feedback loops in a system which cause a variable in the system to influence its own future behaviour. For example, Forrester states that "A feedback system ... is influenced by its own past behavior. A feedback system ... brings results from *past* action of the system back to control *future* action" (1968a, p1-5, Italics added). This indicates how the results of previous actions can be fed forward to influence future actions. Forrester was particularly interested in the feedback that surrounded decision-making. He stated that "every decision is made within a feedback loop" (1968a, p4-4). Thus, Forrester claims that if we have the need for decisions, then we have a situation where the past influences the future through a feedback loop.

However, *any* organisational setting involves decisions being made and therefore a situation where the past influences the future. This general feature of feedback within decision-making is also noted by Coyle (1973, p400) when he comments that "... it is impossible to think of even the simplest operation, such as opening a door, which is not a feedback process." However, he continues by suggesting that "Whether it is worth explicitly modelling the feedback structure instead of merely treating the dynamic behaviour as a stochastic process as in stock control is quite another matter. The key lies in the *purpose of the model* ... If one is interested in a wider view with objectives of controllability and stabilization then a structured control model has to be built" (Italics added). This suggests that Forrester's criterion should not simply be SD does *not* apply to areas where the *past does not influence the future*. Instead, a more helpful criterion would be one that was altered to:

- *SD not only applies to situations which contain feedback loops, but where an explicit modelling of the feedback structure is appropriate when considering the purpose of the modelling process.*

The importance of modelling purpose has also been identified by other authors such as Legasto and Macairiello (1980, p36). When discussing the view that "System dynamics advocates assert the generality of feedback structures in social systems – that no real decision or policy can be made outside a feedback structure", Legasto and Macairiello highlight the impossibility of proving whether or not this view could be proven as a general theory by stating that "Unfortunately, neither view (pro- and anti- "generality theory") can be proven positively ... *modeling purpose* will be the major determinant of the propriety of a viewpoint" (Italics added).

The importance of the modelling purpose in considering the appropriateness of a modelling approach is not only restricted to SD. This is a vital element of any modelling process and is considered in many modelling texts. For example, Rivett (1972, p5) states that "...formulating the problem without formulating the objectives may mean we solve for symptoms only... objectives cannot be treated separately from the model formulation". Also, more recently, Pidd (1996, p13) comments on Ackoff and Sasieni's (1968) definition of a model "a model is a representation of reality" by pointing out that "... it ignores the question of why the model is being built. This aspect is crucial... This simple definition must be expanded to consider the purpose for which the model is being built." The purpose of the modelling process plays an important role in this thesis and, for this reason, forms the focus of chapter 6.

In the case of the modelling of D&D for litigation, the purposes of the modelling processes are detailed in section 2.4.3 as follows:

1. Demonstrate that a part of the time and cost overruns were caused by D&D through particular triggers of D&D.
2. Replicate over time the hours due to D&D which were over-and-above those that were contracted, but were required to carry out the project.
3. Demonstrate the extent to which the client management of the project was reasonable and the extent that overruns could not have been reasonably avoided.
4. All the above have to be demonstrated in a way which will be convincing to the several stakeholders in a litigation audience.

To be able to consider whether or not modelling the feedback structure is appropriate to the purposes of modelling D&D in a litigation environment, data from the Castle and Pirate projects will be explored.

In chapter 2 the description of the Castle and Pirate projects included a list of the main disruptive triggers that occurred during the life of both projects. The following explorations will discuss the disruptive triggers that occurred in both the Castle and Pirate projects and the feedback structures that resulted from them. These explorations will then look at the appropriateness of modelling these feedback structures to the modelling purposes for the two projects.

3.4.1 Castle Project

Based upon the information gathered on the Castle project, the main disruptive triggers discovered during the analysis of the project can be summarised as 6 categories:

1. Reduction in expected learning benefits between products. This was due to an unexpected amount of individuality between the products and various disruptions in the workflow resulting in a reduced ability to gain learning from one product to the next.
2. The defendant was late in supplying customer-furnished information (CFI) which was vital to enable the plaintiff to progress design work as expected.
3. The defendant was late in supplying customer-furnished equipment (CFE). The plaintiff was then required to accept temporary substitutes that were on loan from the defendant for production purposes. However, differences between these substitutes and the final equipment resulted in the need for changes late in the production process.
4. Delays in the resolution of problem trouble reports (PTRs). A PTR occurred when the plaintiff came across an issue in production that required additional information in order that production could proceed. This meant that the plaintiff waited unreasonable lengths of time for information from the defendant and that production work could not proceed as expected.
5. An excessive number of changes causing both disruption and an increase in work scope by more than 30%.
6. The disruptions caused by each of the above 5 categories were of great concern to the plaintiff as their requests that the project schedule be lengthened were refused. Instead the defendant told the plaintiff to use work-arounds throughout the engineering and production processes. Work-arounds meant that the plaintiff was required to make assumptions about missing information or equipment. Some of these assumptions later proved incorrect when the correct information or piece of equipment arrived from the defendant.

Category 1: The first of these categories resulted in tasks taking longer to carry out than originally expected. This type of delay is often not immediately noticed by management since the flow of work is uninterrupted. However, as work is taking longer than expected, the project will gradually fall behind schedule. The only reason that this type of delay would create a more complex disruption to the project, and indeed cause feedback to arise, would be if managerial actions were taken in order to compress the project in an attempt to avoid the schedule overrun.

Categories 2 – 4: The result of categories 2, 3 and 4 was that, unless managerial actions were taken, either engineering or production work would grind to a halt. Since it was not in the plaintiff's interest to allow the flow of work to stop, due to category 6, managerial actions would be taken if the critical path for the project was in danger of being lengthened. These managerial actions would also prevent the labour force from being idle whilst waiting for work to recommence.

Category 5: Figure 3.3 highlights the effect of the category 5 disruptive triggers. The excessive changes in the Castle project (concept 77) caused 'pressures of an increased work load and of schedule slippage against final delivery and other milestones' (concept 42) due to the disruptive trigger detailed in category 6. Managerial actions were then required in order that the project could be progressed. The additional scope of work created by these changes meant that although the flow of work was uninterrupted, management were required to schedule the new work. This may or may not have led to a delay in the overall project schedule, depending on whether or not the critical path for the project was affected. However, if a delay did occur, management could immediately take managerial actions if they wished to attempt to compress the schedule.

Category 6: Instead of being given additional time in the project schedule in response to the disruptions from categories 1 to 5, the defendant was expected to take managerial actions (such as work-arounds) in an attempt to compress the project schedule.

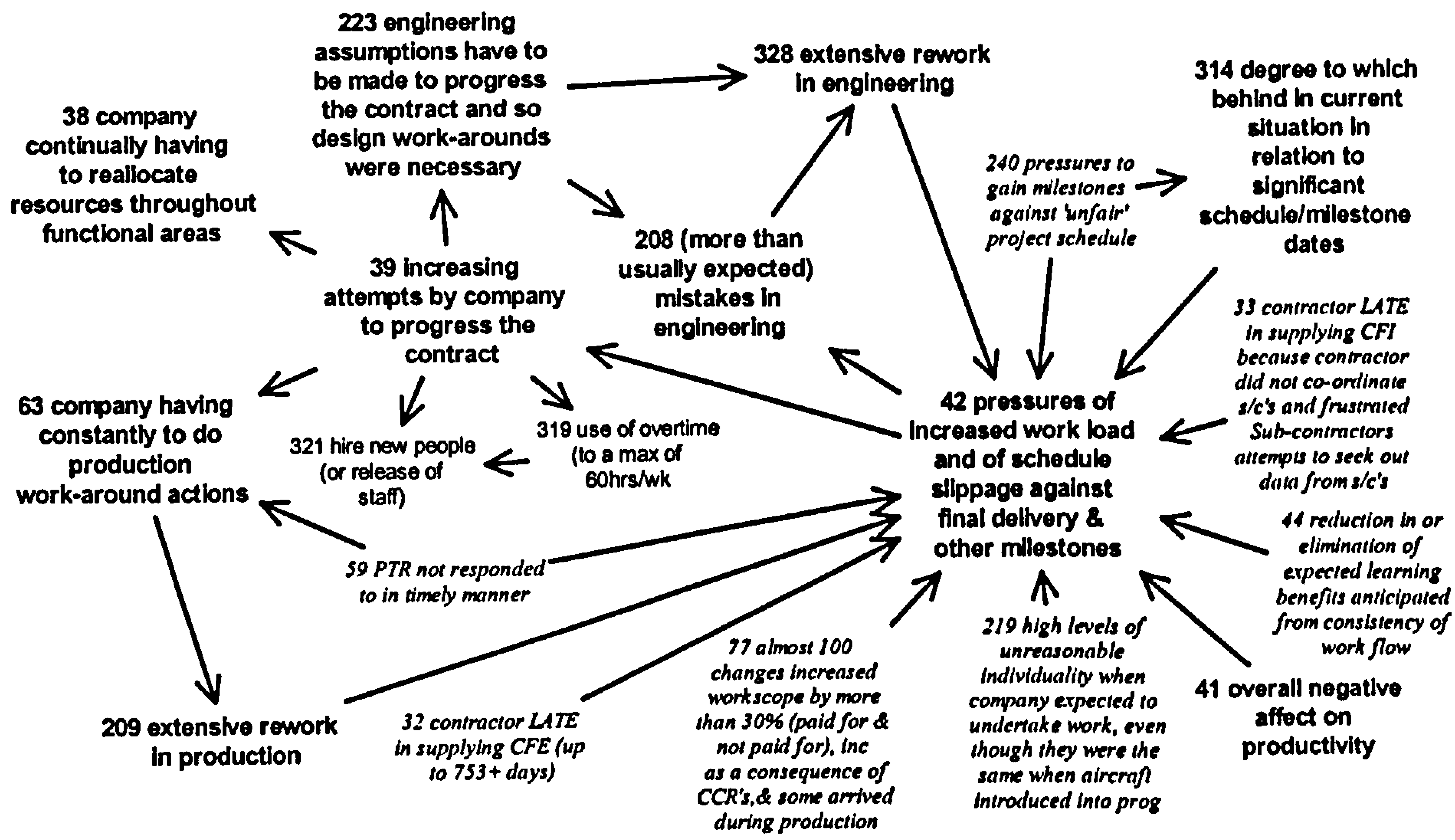


Figure 3.3 - Causes and Effects of Pressures from Increased Workload and Schedule Slippage in the Castle Project

From the above, three groups of disruptive triggers have emerged:

- Disruptions that halt the flow of work. Managerial actions are then immediately required if there is a danger that this halt in work may delay the critical path for the project.
- Additional tasks are created which do not interrupt the workflow, but require scheduling. This would immediately highlight any effect the tasks would have on the critical path and therefore to the overall project length. Management may need to respond to this through managerial actions.
- Tasks that take longer than expected to complete. However, the flow of work is uninterrupted and therefore an immediate management response is not very likely.

The above three groups are listed in descending order of the expectation that a management action will be taken in response to the disruption. For the Castle project, the following summarises the managerial responses taken to disruptive triggers:

Disruption	**Likelihood of immediate managerial action being taken in response to disruption**	**Disruptive trigger categories for the Castle project**
Halts the flow of work	MOST LIKELY	2,3,4
Additional tasks		5
Tasks take longer than expected	LEAST LIKELY	1

Table 3.1 – Categorisation of the Managerial Actions Taken in Response to Disruptions from the D&D Triggers in the Castle Project

The focus of managerial actions is important in this section since they are normally the main cause of feedback in a project (Eden et al 2000). Table 3.1 shows that it was very likely that managerial actions were immediately taken in response to many of the disruptive triggers in the Castle project. This means that it is very likely that feedback loops will have been created. Also, figure 3.3 demonstrates the causes and effects of the 'pressures of an increased work load and of schedule slippage against final delivery and other milestones' (concept 42). It also shows the effects of 'increasing attempts by (the plaintiff) to progress the contract' (concept 39) as a result of the previous concept. If the concept 'increasing attempts by (the plaintiff) to progress the contract' was removed from the overall cause map the only feedback loops that would remain would be those due to:

- Rework causing a cross-impact effect (i.e. rework on one drawing impacts other drawings and causes further rework) and thereby feeding back into rework.
- 'Pressures of an increased work load and of schedule slippage against final delivery and other milestones' causing mistakes that therefore cause rework and hence cause a further increase in workload.

The managerial actions resulting from 'increasing attempts by the plaintiff to progress the contract' account for the majority i.e. 462 out of 513 of the feedback loops created in the cause map.

From the above it can be concluded that not only were feedback loops created from managerial actions (as seen in the cause map), but they were also reinforced through the continual use of managerial actions as highlighted in table 3.1.

The feedback loops created from managerial actions potentially provide strong support for the use of SD in modelling this project. The reason for this is that SD was introduced as a modelling tool to be used to investigate the controlling of negative feedback loops that arise from managerial decision-making. It has been shown above that feedback loops arise from the main disruptive triggers in the Castle project. However, this section is particularly focussed on whether or not the investigation of such loops can also be seen to play an important role in the purposes of the modelling of D&D in the Castle project. Modelling purposes 1 and 2 state that there is a need to:

- *Demonstrate that a part of the time and cost overruns were caused by D&D and identify the triggers of D&D.*
- *Replicate over time the hours due to D&D which were over-and-above those that were contracted, but were required to carry out the project.*

In achieving these purposes, a modeller is likely to want to firstly gain an understanding of how D&D is created in the project. When discussing the nature of D&D in section 2.4, it was noted that one of the most significant circumstances for D&D arose from a feedback phenomenon. Therefore an understanding of the feedback structure created from managerial actions is an important step in modelling D&D. It could therefore be concluded that the Castle project did involve feedback loops, and that their explicit modelling is appropriate when considering the purposes of the Castle modelling process.

Having discussed the disruptive triggers experienced in the Castle project, the following section will now discuss those that occurred during the Pirate project. This will be done so that a comparison can be made between the disruptive triggers experienced in the two projects.

3.4.2 Pirate Project

Based upon the information gathered on the Pirate project, the main disruptive triggers experienced during the project can be summarised in the following 8 categories:

1. Underestimation of the scope of work when estimating the number of hours required to complete the project.
2. Delayed corporate learning.
3. The unpredictable nature of the project meant that the management of the project was unable to gain a normal level of reduction in the time to produce the first product through planning actions carried out before the commencement of the project.
4. Lost learning between products in production due to a higher than expected level of staff turnover.
5. Increased individuality between products. This caused a reduction in the learning that could be achieved between products in production.
6. When corporate learning was achieved between products in production, it was at a rate that was slower than had been expected.
7. An increased number of non-routines (NRs) than had been expected. A NR represents additional work that is required on a product in order for it to meet the specifications laid down in the project contract. NRs are normally identified by the plaintiff when either stripping or rebuilding the product. These were caused by the condition of the product being of a poorer quality than had been expected at the beginning of the project.
8. Delays in approval of NRs.

Category 1: The first trigger category meant that the project involved:

- Tasks taking longer to complete than had been expected at the beginning of the project. As mentioned in the discussion of the first disruptive trigger category in the Castle project, the type of delays that occur when tasks take longer to complete are often not immediately noticed. This is due to the uninterrupted progression of work. However, as work is taking longer than expected, the project will gradually fall behind schedule. The only reason that this type of delay would create a more complex disruption to the project, and indeed cause feedback to arise, would be if managerial actions were taken in order to compress the project in an attempt to avoid the schedule overrun.

- Additional tasks over those that had been expected at the beginning of the project. When these tasks are noticed, although the flow of work would generally not halt, management attention would be captured as the new tasks need to be fitted into the project schedule and therefore any delays in the overall project would be highlighted.

Categories 2 – 6: Trigger categories 2 to 6 resulted in tasks taking longer than originally expected due to a reduction in the learning that was expected to be gained between products in production. The only reason that this type of delay would create feedback in the project would be if managerial actions were taken in order to compress the project in an attempt to avoid a schedule overrun.

Category 7: Category 7 disruptions meant that additional work was being found throughout the life of the project from NRs. Management would therefore need to schedule these tasks. These could potentially result in a lengthening of the project and potential managerial actions in reaction to this.

Category 8: Category 8 disruptions meant that by the time the work for an NR was expected to be carried out, the plaintiff may not have obtained the approval to carry out the work from the defendant. The flow of work would therefore halt unless managerial actions were taken to work around this delay.

Based on the three groups of disruptions discussed for the Castle project in section 3.4.1, the disruptions experienced during the Pirate project are summarised in table 3.2.

When comparing table 3.2 and table 3.1, it can be seen that the disruptions occurring during the Pirate project were less likely than those during the Castle project to cause an immediate managerial action to be taken when they occurred.

Disruption	Likelihood of immediate managerial action being taken in response to disruption	Disruptive trigger categories for the Pirate project
Halts the flow of work	MOST LIKELY	8
Additional tasks		1,7
Tasks take longer than expected	LEAST LIKELY	1,2,3,4,5,6

Table 3.2 - Categorisation of the Managerial Actions Taken in Response to the Disruptions from D&D Triggers in the Pirate Project

When considering the overall cause map for the Pirate project, all but one of the 299 loops (rework causing cross-impact, causing rework) contained the variable *compression*. The feedback loops underlying the Pirate project were therefore totally dependent upon the use of managerial actions to compress the project. This leads to the conclusion that it is only appropriate to model the feedback structure in the Pirate project if the use of managerial actions to compress the project were a significant disruption to the project. However, table 3.2 suggests that managerial actions were less likely to be used in response to disruptions when compared to the Castle project. The feedback loops that were created from managerial actions were less likely to have been reinforced during the course of the project. In fact, when modelling the Pirate project, it was revealed that a large majority of the overrun of the project could be accounted for without taking account of project compression. A result of this was that D&D from the feedback triggered by the use of various managerial actions did not play a sufficiently significant role in the overrun of the project for the modelling process to take a detailed account of it.

The above explorations demonstrated that the main disruptive triggers in a project could be examined to determine whether or not modelling the feedback structure is appropriate to the purposes of modelling D&D for litigation. When exploring the Castle and Pirate projects it was concluded that it was more appropriate to model the feedback structure in this project if managerial actions were likely to be immediately taken in response to disruptions in order to compress the project. An assessment therefore needs to be made of the type of disruptions a project has experienced. However, it was also noted that this alone was not sufficiently conclusive in order to respond to whether or not the feedback structure was an important element to consider when attempting to meet the modelling purposes of the Pirate project. It was only at the quantification stage that a decision was finally reached. However, this is not the most efficient way in which to carry out a modelling process. An ability to decide this earlier in the modelling process would be beneficial.

3.4.3 Summary of the Explorations of Criterion 2

Section 3.4 began with the second of Forrester's criteria. In order that it could be of more use to a modeller it was altered to:

- *SD not only applies to situations which contain feedback loops, but where an explicit modelling of the feedback structure is appropriate when considering the purpose of the modelling process.*

When considering this criterion in the light of the Castle and Pirate projects, it has been concluded that if D&D is a significant cause of the overrun of a project, then the project will contain feedback loops and the feedback structure will be appropriate when considering the purposes of the modelling process. However, a modeller needs to be aware of situations such as the Pirate project where D&D did not appear to play a significant role in the overrun experienced in the project, even although it played a significant part in the cause map constructed as a part of the analysis of the project.

3.5 Criterion 3: Changes Through Time are of Interest

We are told by Forrester that if variables of a situation change over time then the situation is referred to as exhibiting dynamic behaviour. He states that this dynamic behaviour is generated by feedback (1968a). Forrester also states that "every decision is made within a feedback loop" (p4-4, 1968a). It can therefore be concluded that a situation that contains *decisions* creates *feedback* and thus generates *dynamic behaviour (i.e changes through time).* Virtually *any* organisational situation involves decisions being made. Therefore, it would be difficult to describe an organisational situation that would not have changes through time occurring.

Since it has been concluded that changes through time will occur in an organisational setting, the question then becomes *are these changes through time of interest?* To assess whether or not such changes through time would be of interest to a situation, it is worthwhile considering the *purpose* of the modelling process.

When considering the relevance of changes through time to the purposes of modelling D&D for litigation, purpose 2 stands out as being particularly relevant i.e.

- *Replicate over time the hours due to D&D which were over-and-above those that were contracted, but were required to carry out the project.*

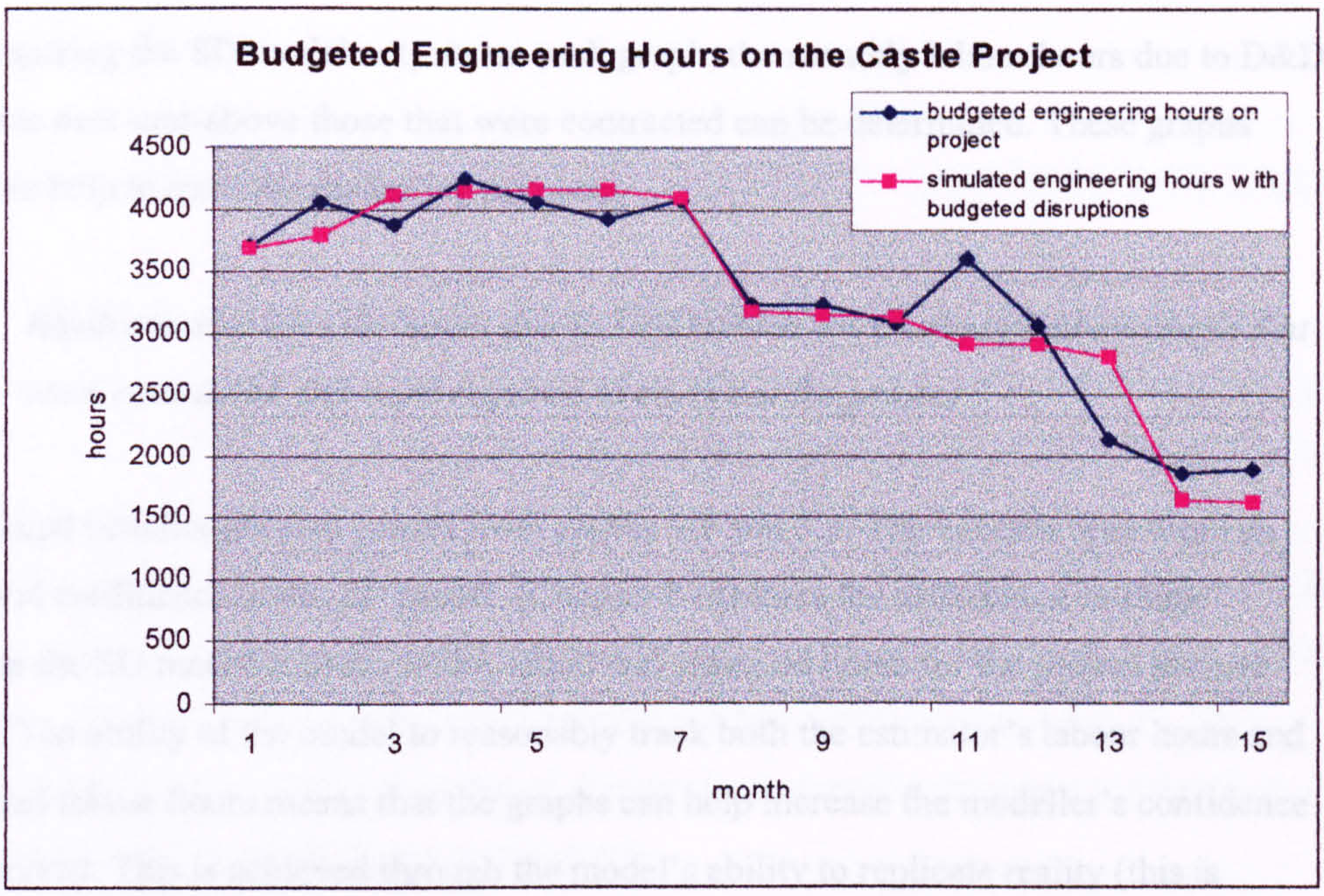
The need to replicate hours over time means that the hours worked by the labour force will need to be tracked as they change from one time period to the next. The Castle and Pirate projects will now be used to demonstrate the role of tracking labour hours over time when modelling D&D for litigation.

3.5.1 Castle Project

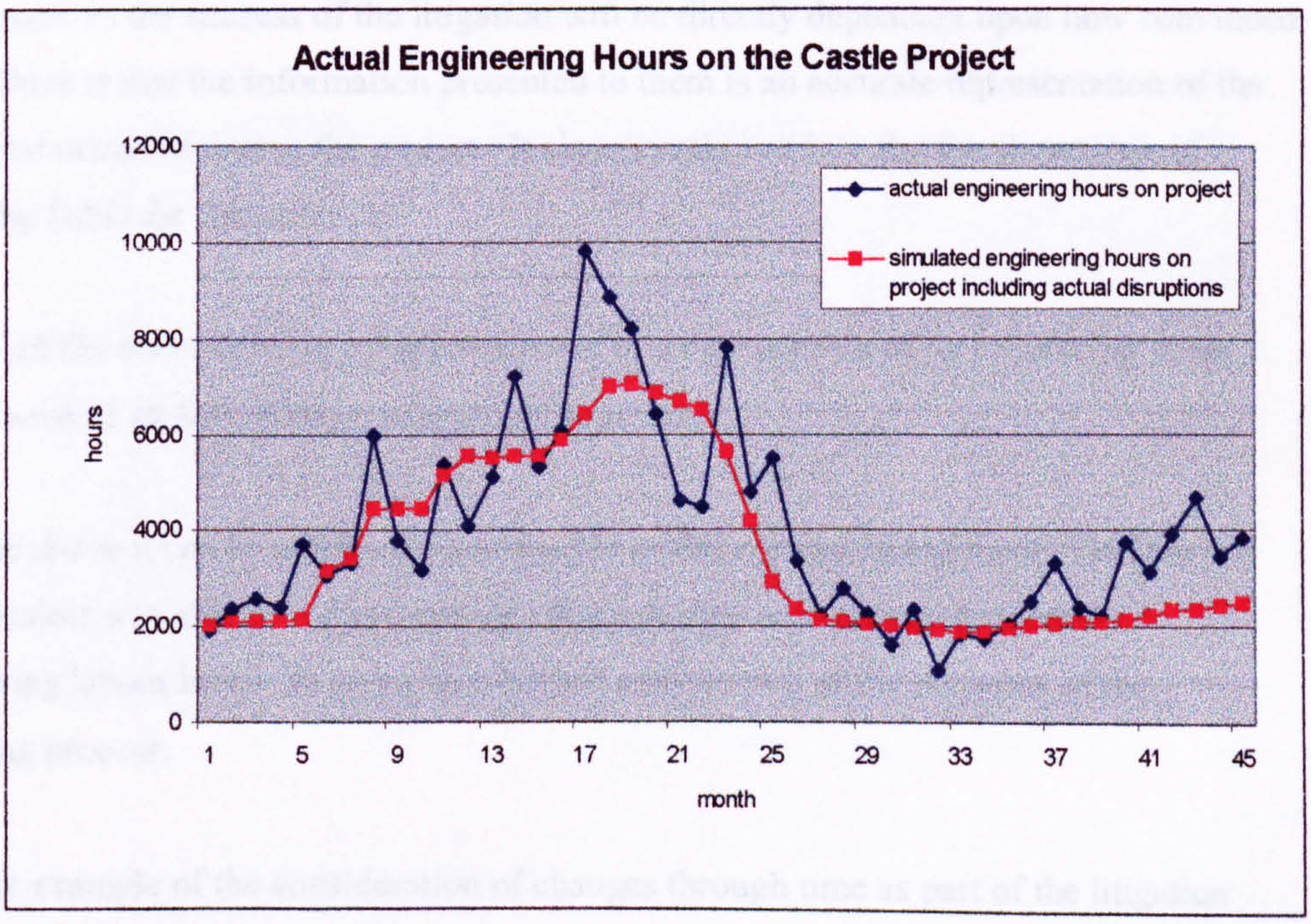
The following two graphs show:

(i) Graph 3.1 - the expected monthly engineering hours calculated by the estimators prior to the commencement of the project compared to the SD model output of monthly engineering hours when the only disruptive triggers included are those that were anticipated at the commencement of the project.

(ii) Graph 3.2 - the actual engineering hours experienced on the project compared to the SD model output of monthly engineering hours when the disruptive triggers included were those that were actually experienced during the project.



Graph 3.1 – Budgeted Engineering Hours for the Castle Project



Graph 3.2 - Actual Engineering Hours for the Castle Project

By comparing the SD model outputs on each graph, the monthly labour hours due to D&D that were over-and-above those that were contracted can be determined. These graphs therefore help to meet the modelling purpose:

- *Replicate over time the hours due to D&D which were over-and-above those that were contracted, but were required to carry out the project.*

Additional benefits are also gained from graphs 3.1 and 3.2. The benefits arise from an increased confidence in the SD model. (Chapter 8 explores the differences in shape between the SD model output and the actual and expected hours for the project in more detail). The ability of the model to reasonably track both the estimator's labour hours and the actual labour hours means that the graphs can help increase the modeller's confidence in the model. This is achieved through the model's ability to replicate reality (this is discussed fuller in chapter 8). This replication of reality also means that the graphs may help in the process of gaining the various stakeholders' confidence in the model (e.g. the plaintiff, the defendant, a judge or a jury). This is an important measure in a litigation environment as the success of the litigation will be directly dependant upon how convinced the audience is that the information presented to them is an accurate representation of the events that occurred during the project. It also directly ties into the fourth purpose of modelling D&D for litigation i.e.

- *All the above have to be demonstrated in a way which will be convincing to the several stakeholders in a litigation audience.*

From the above it can be concluded that the SD model constructed to model D&D in the Castle project was able to capture changes through time with respect to monthly engineering labour hours. In doing so it helped achieve two of the purposes of the modelling process.

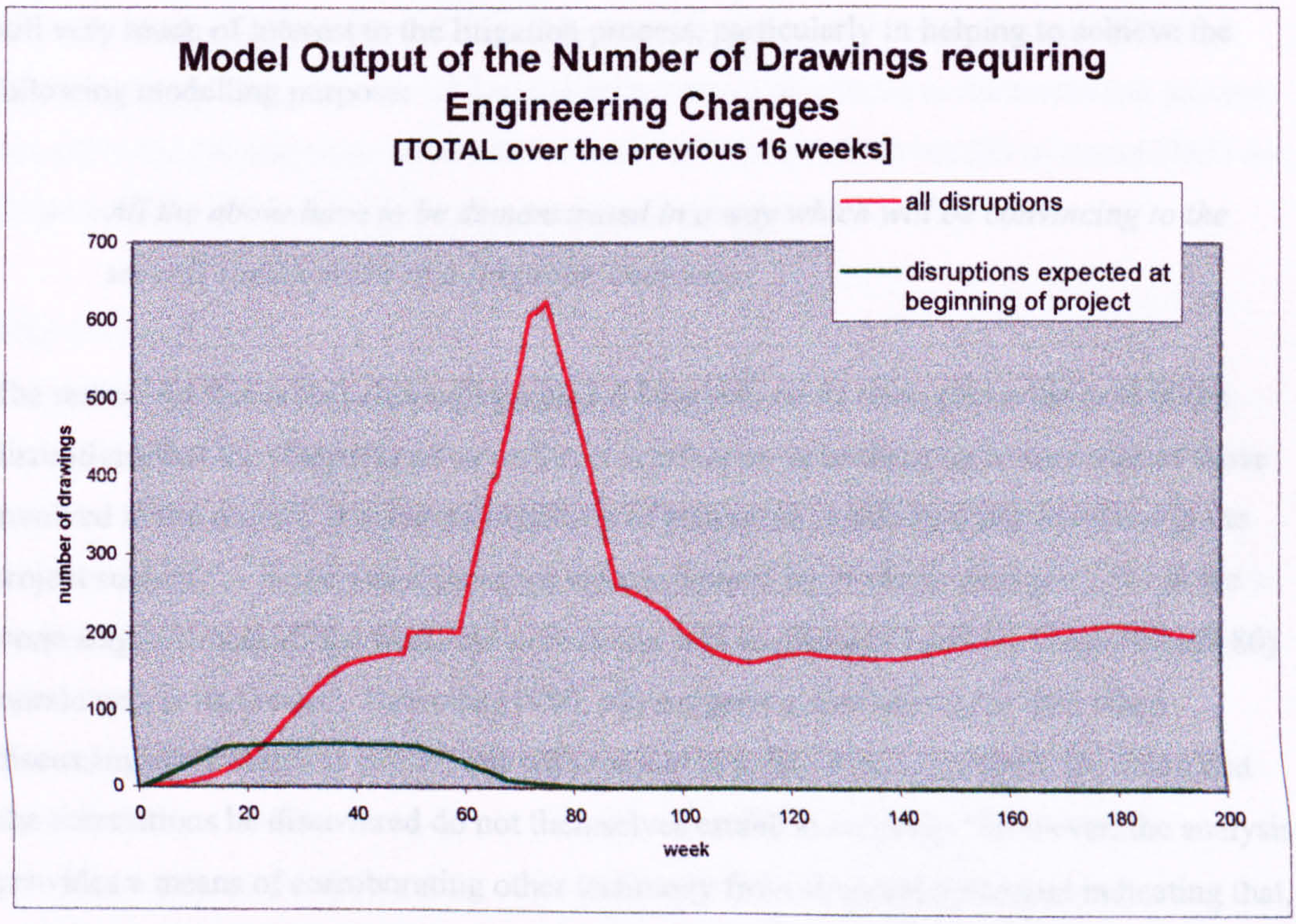
A further example of the consideration of changes through time as part of the litigation process is demonstrated by graph 3.3. This graph shows the comparison between the number of drawings requiring change based on:

(i) the original budget calculated by the estimators (the lower line) and

(ii) the demands on engineering from all of the disruptions experienced on the project (the upper line).

In both cases the curves were generated by the SD model constructed for the Castle project.

When considering the reasons for the shape of the data displayed on the graph, the data replicating what actually occurred on the project was split up into several time periods. For example, time 0 to 20 weeks, time 20 to 60 weeks and time 60 to 90 weeks each display a different shape on the upper line. Reasons for the shape of the output during each time period can be explained through events that were observed during the project. For example, the large peak between time 60 to 90 weeks can be explained as being due to a significant body of additional work arriving during the period a few weeks prior to week 60. This additional work caused a large number of changes to engineering drawings to take place during the time of the peak shown on the graph.



Graph 3.3 – Number of Drawings Requiring Engineering Changes in the Castle Project

An important difference in the behaviour of the variable displayed in graph 3.3, compared to the variables displayed in graphs 3.1 and 3.2, is that the changes over time are believed to be mainly caused by exogenous inputs to the system. This is in contrast to graphs 3.1 and 3.2 where the changes over time were believed to be mainly due to endogenous system behaviour. (The main causes of the changes in labour hours between graphs 3.1 and 3.2 were discussed as triggers of D&D in section 3.4). When Forrester stated that "SD does *not* apply.... to situations where *changes through time are not of interest*" (1968b, p605), although he does not qualify exactly what is meant by this, his other work indicates that he is concerned with changes through time which are caused endogenously by the system structure. Such endogenous system behaviour can be *triggered* by an exogenous input. However, the changes through time should not be *entirely due to* the behaviour of an exogenous trigger. This means that the type of changes over time represented in graph 3.3 were not the type of changes over time that would call for SD to be used as a modelling approach. However, it is interesting to note that the changes displayed on this graph were still very much of interest to the litigation process, particularly in helping to achieve the following modelling purpose:

- *All the above have to be demonstrated in a way which will be convincing to the several stakeholders in a litigation audience.*

The reason for this is that although graph 3.3 does not, on its own, prove the case of the disruptions that the plaintiff had to suffer, it is effective in backing up testimonies of those involved in the project. It acted as a back-up of statements made by those involved in the project such as "...at the worst point we were swamped by work on changes", "...at the worst stages almost all the work we were doing was on changes" and "...around (week 80) morale was at its lowest". Nahmias (1980, p8) suggests a similar use for data when discussing work that was carried out with respect to a shipbuilding project. He noted that the correlations he discovered do not themselves establish causality "However, the analysis provides a means of corroborating other testimony from shipyard personnel indicating that, in fact, this was the case". The most senior member of the plaintiff's claim team for the Castle project also found graph 3.3 very powerful as it tied in "...exactly with a 'summing up' statement that I made at one point that this subcontract...actually ceased to be a project to manage the (Castle) installation and became a contract to manage the changes that were coming from (the defendant)". This helped to reinforce one of the key messages that the plaintiff wanted to get over to the defendant that "the management of change became the prime activity of each of the key departments". Graph 3.3 therefore proved effective in aiding the explanation of the result of various disruptions on the progress of the Castle project. It helped to demonstrate the effect of disruptions on the project to various stakeholders in the litigation audience. Since the data is shown as a graph it also means that this process becomes easier as it is very visual and tends to be easier for a layperson to follow rather than simply presenting them with data.

An exploration of graphs 3.1, 3.2 and 3.3 has shown that the tracking of variables over time proved useful in helping to demonstrate the effect of disruptions during the Castle project. However, when considering whether or not changes through time for a particular variable are of interest to the modelling process (and therefore to the consideration of whether or not it is appropriate to use SD to model D&D for litigation) consideration has to be given to whether or not the behaviour of the variable of interest is created through exogenous or endogenous system behaviour. Only those variables whose behaviour is created through endogenous system behaviour and whose behaviour is of interest to the modelling process can aid in the consideration of whether or not it is appropriate to use SD to model D&D for litigation. Graph 3.1 and 3.2 displayed such variables for the Castle project.

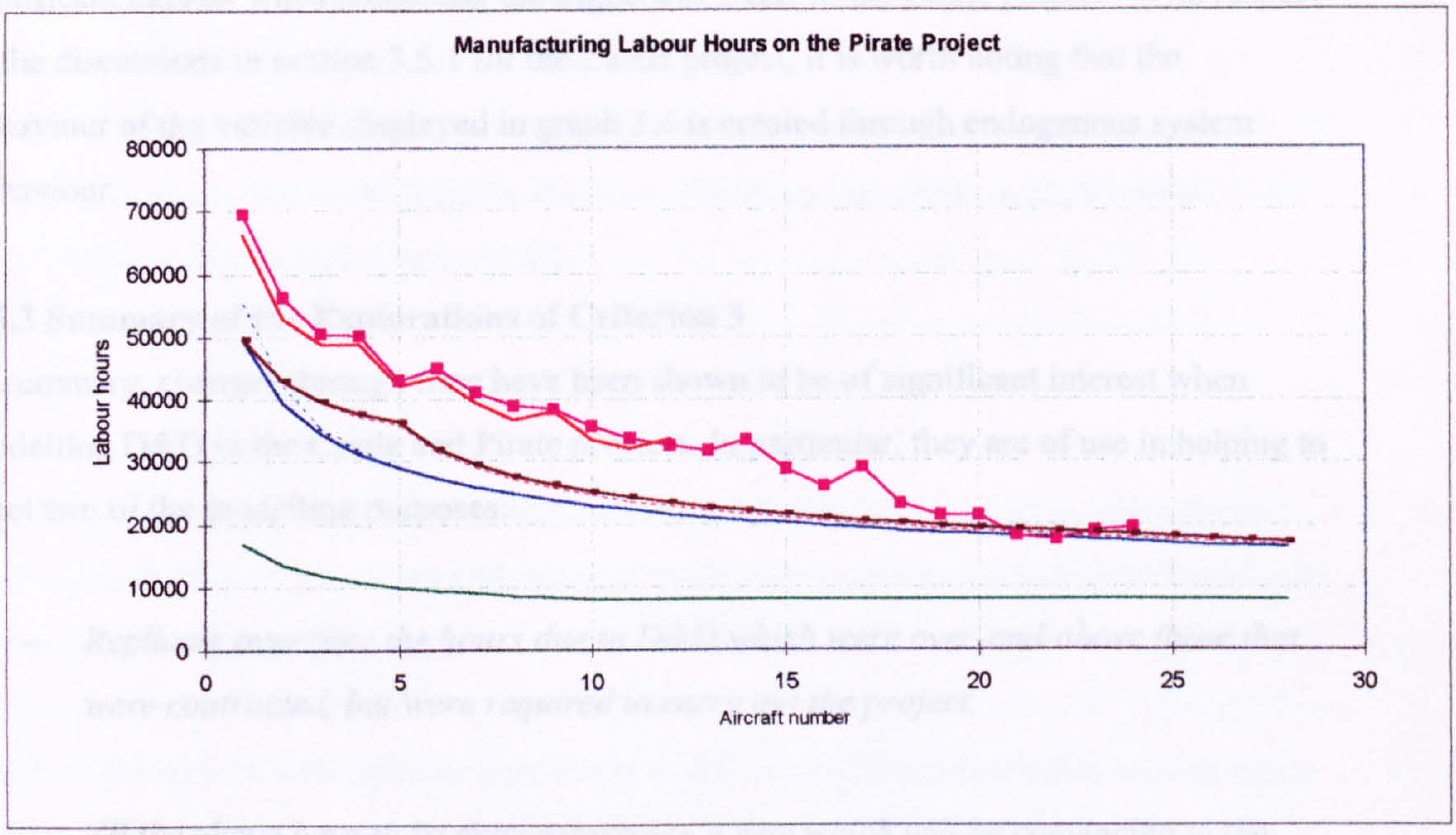
3.5.2 Pirate Project

Graph 3.4, which is shown below, represents the build-up of the hours spent modifying each aircraft in the Pirate project. As detailed in section 3.3.4, the hours displayed on this graph were determined from a spreadsheet analysis of the major causes of the overrun in the Pirate project. The graph shows the changes in the labour hours from one aircraft to the next. Each curve represents the increase in labour hours due to various disruptive triggers.

The curves shown on the graph represent the following (moving from the bottom curve upwards):

1. Green line: The hours estimated in the contract as those that would be required to modify each aircraft.
2. Blue line: The hours estimated in the model that would be required to modify each aircraft, allowing for increased labour turnover.

3. Blue dotted line: As for the previous curve, but also allowing for the fact that the project was a developmental project.
4. Red dotted line: As for the previous curve, but also allowing for delayed corporate learning.
5. Brown line: As for the previous curve, but also allowing for increased individuality between the aircraft.
6. Red line: Total actual hours experienced on the project minus an estimate for the number of hours due to self-inflicted D&D.
7. Pink line: Total actual hours experienced on the project.



Graph 3.4 – Build-up of Labour Hours on the Pirate Project

By plotting the labour hours required in the production of each aircraft, a learning curve for the project will be formed. The estimation of the learning curve for a repetitive manufacturing project plays an important role in the estimation of the labour hours required to complete the project (Yelle 1979, Argote and Epple 1990). Many of the disruptive triggers listed in section 3.4.2 cause a direct disruption to the learning curve for the Pirate project. For example, the effect of delayed corporate learning or the effect of increased individuality between products causing a reduction in the learning that could be gained. As the disruption of the learning curve was a major driver to the overall project overrun, it was deemed important to be able to model the effects of the various disruptions on the learning curve for the project. Graph 3.4 was therefore seen to be an important tool to aid the explanation of why the labour hours increased as they did. Graph 3.4 aided the replication of the labour hours over-and-above those that were contracted but were required to carry out the project due to the various disruptions that occurred during the project.

Changes through time due to disruptions to the learning curve therefore proved of significant interest when modelling the triggers of D&D in the Pirate project. With respect to the discussions in section 3.5.1 for the Castle project, it is worth noting that the behaviour of the variable displayed in graph 3.4 is created through endogenous system behaviour.

3.5.3 Summary of the Explorations of Criterion 3

In summary, changes through time have been shown to be of significant interest when modelling D&D in the Castle and Pirate projects. In particular, they are of use in helping to meet two of the modelling purposes:

- *Replicate over time the hours due to D&D which were over-and-above those that were contracted, but were required to carry out the project.*

- *All the above have to be demonstrated in a way which will be convincing to the several stakeholders in a litigation audience.*

However, when considering whether or not changes through time for a particular variable are of interest to the modelling process (and therefore to the consideration of whether or not it is appropriate to use SD to model D&D for litigation) consideration has to be given to whether or not the behaviour of the variable of interest is created through exogenous or endogenous system behaviour. Only those variables whose behaviour is created through endogenous system behaviour and whose behaviour is of interest to the modelling process can aid in the consideration of whether or not it is appropriate to use SD to analyse D&D

for litigation. In both the Castle and Pirate projects, variables exist whose behaviour is created through endogenous system behaviour and is of interest to the modelling process.

3.6 Conclusions from Forrester's Criteria

If the three criteria introduced in this chapter were considered only in the light of Forrester's explanations of their meaning, then it would be difficult to imagine organisational situations where any of these criteria would not hold. From this there are two conclusions that could be drawn:

Either:

(i) Forrester appears to believe that SD can be used to model all organisational situations. However, if this is the case, then it was not made explicit when discussing the limitations of SD.

Or:

(ii) If Forrester does believe there are situations for which SD is not an appropriate modelling approach, then his attempt to identify these situations has not been made clear.

Much of Forrester's writing leads a reader towards the conclusion that the former of these conclusions holds. When asked for limitations of SD in order that it could be labelled a theory, the generic nature of the responses indicates the difficulty that Forrester had in answering this question.

In order to take the explorations further, the three criteria were expanded by reflecting on the experiences of modelling D&D for litigation. In doing so, a number of conclusions were made.

This chapter has highlighted the importance of considering the purposes of the modelling process when forming a decision on whether or not SD is an appropriate method to use to model a given situation. Based on this, Forrester's criteria were altered to provide more useful criteria to aid the decision whether or not SD should be used to model a situation. The second and third criteria were altered to the following:

- *SD not only applies to situations which contain feedback loops, but where an explicit modelling of the feedback structure is appropriate when considering the purposes of the modelling process.*

And

- *SD does not apply to situations where changes through time are not of interest to the purposes of the modelling process.*

Both of these criteria invite the modeller to make an assessment in the light of the purposes of the modelling process. When using these criteria it is unlikely that the conclusion would be a definite 'yes' or 'no'. It is more likely that the use of such criteria will lead to a response that would indicate the *extent* of the interest to the purposes of the modelling process. This will undoubtedly cause difficulties unless the conclusions are at either extremes of the scale i.e. 'is of significant interest' or 'is of virtually no interest' to the purposes of the modelling process. This may cause many outcomes that conclude that SD *could* be used to model a situation, but are inconclusive to whether it *should* be used. Although this means that a modeller will still face difficulties using the criteria, their improved comprehensiveness enables them to form part of a revised set of criteria set out at the end of chapter 5. Following an exploration of existing criteria in the SD literature, the aim of this revised set of criteria is to provide a modeller with a comprehensive, practical method with which to assess the appropriateness of SD to analyse D&D in a given situation for litigation.

When exploring the use of the revised criteria when modelling D&D in a litigation environment, the following were concluded:

- It was suggested that systemic interrelationships would be highlighted if a systems approach were taken when analysing a situation. The focus of the explorations therefore turned to whether or not a systems approach is *necessary* when modelling D&D for litigation. From the explorations, it was concluded that non-systems analysis modelling approaches used in the Castle and Pirate projects did not fully explain and capture all of the effects of D&D. However, in using systems analysis approaches there were some concerns over how easily litigation audiences were able to understand and thereby be convinced by them. Both of these conclusions produced hurdles that would need to be overcome when either type of approach was used in practice.

- If D&D is a significant cause of the overrun of a project, then the project will contain feedback loops and the feedback structure will be appropriate when considering the purposes of the modelling process. However, a modeller needs to be aware of situations such as the Pirate project where D&D from the feedback triggered by the use of various managerial actions did not appear to play a significant role in the overrun experienced in the project. However, it was noted that it was not possible to reach a final conclusion on this until late in the modelling process. As this is not the most efficient way in which to carry out a modelling process, an ability to decide this earlier in the modelling process would be beneficial.

- Changes through time are of significant interest to the purposes of modelling D&D in the Castle and Pirate projects. It was noted that when considering whether or not changes through time for a particular variable are of interest to the modelling process, consideration has to be given to whether or not the behaviour of the variable of interest is created through exogenous or endogenous system behaviour. Only those variables whose behaviour is created through endogenous system behaviour and is of interest to the modelling process can assist the consideration of whether or not it is appropriate to use SD to analyse D&D in a given situation for litigation.

In the light of these conclusions, there is certainly no evidence to suggest that SD *should not* be used to model D&D for litigation. However, as expected, it is inconclusive on whether or not it *should* be used.

Having now fully explored Forrester's criteria, the next chapter turns to the exploration of a second set of criteria taken from the SD literature. These criteria were laid down by Coyle in 1977. They represent a more thorough and detailed set of criteria when compared to those laid down by Forrester.

CHAPTER 4: EXPLORING CRITERIA DETAILED BY COYLE IN THE SD LITERATURE

4.1 Introduction

In chapter 3 it was stated that the SD literature would be used to explore what criteria, if any, have been defined as those that should be used to identify whether or not a situation can be modelled using SD. This would enable an exploration into the types of situations the modelling approach was intended for and would enable an investigation into whether or not SD should be used to model D&D for litigation. Also, it was stated that data gathered from the author's involvement in the modelling of D&D for two litigation cases would be used to explore the criteria in an attempt to aid the process of discovering whether or not SD should be used to model D&D for litigation.

Based on the above, three different sets of criteria were chosen from the literature. The first set of criteria, put forward by Forrester in 1968, was explored in chapter 3. This chapter will now explore the second set of criteria set out by Coyle in 1977.

4.2 Coyle's Criteria

Coyle (1977, p357) puts forward his criteria as he addresses *Criteria for Project Selection.* In 1996, Coyle (p348) noted that there is an "enormous range of the potential applicability of system dynamics, but the fact that a methodology *can* be applied to a particular problem does not mean that the circumstances are necessarily propitious to a *successful* application." To help form an opinion on "...whether or not a particular managerial problem is likely to be a suitable candidate for dynamic modelling", Coyle (1977) sets down 14 criteria. Coyle (1996) discusses a summary of this same list of criteria when he revisits the issue of the appropriateness of SD to model a given situation in *System Dynamics Modelling: A Practical Approach.* The criteria are intended to be taken as *guidelines* and not as a set of formal tests. This point is important, as many of the criteria are practical tests for a modeller to consider and should not be regarded as a reason to fully accept or reject a particular situation as being appropriate to be modelled by SD. However, Coyle's discussion of this set of criteria provides more detail than the set of criteria explored in chapter 3 and may therefore help to progress the explorations in this thesis. Since Coyle sets down the criteria quite explicitly as 14 different points, each of these points will now be considered in turn.

4.3 Criterion 1: "Is there any dynamic behaviour?"

It has been shown that one of the most significant causes of D&D arises from the effects of feedback loops (see section 2.4). If this view is accepted, then by Forrester's principle number 4.2-2. (1968a, p4-5), which states that dynamic behaviour is generated by feedback (a definition supported by Coyle when he discusses dynamic behaviour in systems (1977, p23)), *a system with feedback D&D will create dynamic behaviour.* This issue was also covered in section 3.5 when discussing Forrester's criterion, *changes through time are of interest.* This section noted that it would be difficult to describe any organisational situation that would not have changes through time (i.e. dynamic behaviour) occurring. Therefore, the only part of this criterion that prompted any exploration was *...are of interest.* For this part, the purposes of the modelling process were considered in order that an evaluation could be made about whether or not the dynamic behaviour of D&D was of interest to the purposes of the modelling process. These explorations are relevant to Coyle's second criterion detailed in section 4.4. Before looking at this criterion it can be concluded that the first of Coyle's criteria adds nothing new to the investigations in this thesis.

4.4 Criterion 2: "Do the dynamics matter, and why?"

As mentioned in section 4.3, this criterion is considering the same issues as Forrester's criterion changes through time are of interest considered in section 3.5. This section concluded that changes through time are of significant interest to the purposes of modelling D&D for litigation. This means that the dynamics do matter when considering them in the light of the purposes of the modelling process. The reason given for them mattering was that their consideration was particularly helpful in meeting two of the modelling purposes. These are:

- *Replicate over time the hours due to D&D which were over-and-above those that were contracted, but were required to carry out the project.*

- *All the above have to be demonstrated in a way which will be convincing to the several stakeholders in a litigation audience.*

Again, this criterion mirrors one of Forrester's that has already been explored and therefore adds nothing new to the investigations in this thesis.

4.5 Criterion 3: "Are there any loops?"

Coyle continues (1977, p357) by stating that "... a dynamic model has to contain loops...". Therefore, this criterion links directly with Coyle's first criterion *is there any dynamic behaviour?* discussed in section 4.3. If the answer to criterion 1 is 'yes', then the answer to this criterion will also be 'yes'. This question therefore does not create any additional criterion to aid the explorations in this thesis. The interest in *loops* when assessing criteria was captured in the revised criterion discussed in section 3.4. The revised criterion is as follows:

- *SD not only applies to situations which contain feedback loops, but where an explicit modelling of the feedback structure is appropriate when considering the purpose of the modelling process.*

When exploring this criterion in section 3.4 it was concluded in section 3.4.3 that, based on the experiences of the Castle and Pirate projects, if D&D is a significant cause of the overrun of a project, then the project will contain feedback loops and the feedback structure will be appropriate when considering the purposes of the modelling process.

Although the criterion in this section has already been covered by discussions in section 3.4, it is worth noting that Coyle provides a modeller with practical advice on the consideration of this criterion. In the discussions following the stating of the criterion, Coyle suggests the use of a simple influence diagram (Coyle 1977, 1996) to convince the modeller of the existence of loops. As discussed in section 2.4, cause maps were constructed as one of the first stages of modelling each of the projects discussed in this thesis. As for influence diagrams, these maps can be used to discover what loops, if any, are surfaced from interviews with those involved in the project.

Apart from the more practical discussion on the use of influence diagrams to aid the decision on whether or not loops exist, this criterion adds nothing new to the investigations in this thesis.

4.6 Criterion 4: "Are there any alternative system structures or control policies?"

Coyle (1977, p358) suggests that if there are no alternatives, then "the system is either perfect or paralysed" therefore "the study is pointless because ... nothing can be changed". This is obviously important if the purpose of the study is to create a change in the system. However, as noted by Richardson and Pugh (1981, p38, Italics added), "The purpose of a system dynamics model is *understanding*". This may mean that a model may be created to improve peoples' understanding of the structure of the system.

For example, consider the Production-Distribution Simulation often referred to as the Beer Game (Jarmain 1963). This can be played as a manual simulation, but has also been created into a quantitative SD model (Goodman et al 1993). The intention of building the model is not to find alternative system structures or control policies for the system, but for players to improve their understanding of the structure of the system.

Coyle (1996, p348-349) states that "...the purpose of any management science investigation is to change a situation...This factor of change is still true even when the investigation is into, say, the processes which led to the collapse of the Maya. The object now becomes to influence the ways in which scholars have interpreted and understood that problem or to help them to organize their insights, and hence generate new ones ... but the net effect will still be change." The *change* described in this situation is focussed upon the way in which the users of the model *think*, rather than on changing the specific system being modelled. This does not mean that alternative system structures or control policies are *not possible* in such studies, but simply that they do not form the focus of the modelling purpose.

Coyle (1977, p5, Italics added), describes SD as "... that part of management science which deals with the *controllability* of managed systems over time, usually in the face of external shocks". Based on this objective, it would be reasonable to place an emphasis on the possibility of alternative system structures or control policies. However, the example mentioned above demonstrates that SD modelling can and *has* been used beyond the objective of controlling managed systems. Rather than this criterion being used as a reason to *reject* SD as a modelling approach, it may still be used as a reason to *accept* SD as a modelling tool since SD is suited to the controllability of managed systems. However SD can also be used to meet other objectives.

Before progressing on to the next criterion, the next section will consider the current criterion with respect to the projects discussed in this thesis.

4.6.1 Castle and Pirate Projects

A major difference between the use of a SD model in a litigation environment and other more common uses of SD (for example, the use of SD for policy analysis) is that the focus is on a *post-mortem* study and the outcome is not expected to have a direct input into future actions. The main focus of such a study is to provide a plausible *explanation* of what occurred in a project. This is to enable those involved in the claim situation to gain a better understanding of why a particular project resulted in both time and cost overruns. Of course, by gaining such an understanding of a project, any managers involved in the process may make use of their improved understanding when managing future projects. In this way the SD modelling process may indirectly influence future managerial actions.

During the Castle and Pirate modelling processes, alternative system structures and control policies were *not* considered. The only changes made to the models were to consider the effect on the system due to the occurrence of different exogenous triggers. This was done so that the SD model could be used to compare what was contractually expected to occur in the project against what actually occurred. For example, the output from the SD model constructed as part of the Castle project provided the output shown in graphs 3.1 and 3.2 in section 3.5.1. These graphs demonstrate the SD model output when the exogenous disruptive triggers are set to what was expected in the contract and to what actually occurred in the project. Simulating both of these situations was an important part of the modelling process as it helped to achieve the following modelling purpose:

- *Replicate over time the hours due to D&D which were over-and-above those that were contracted, but were required to carry out the project.*

Although alternative system structures and control policies are not the focus of the modelling considered in this thesis, possible alternatives do exist in the systems. Each of the projects discussed in this thesis could have been managed in many different ways. For example, instead of hiring more people in the Pirate project, management could have placed more emphasis on working shifts or overtime. Also, instead of using overtime in the Castle project, management could have used pressure on workers in an attempt to accelerate the project. There are many different ways in which management can attempt to control a project. This is an area that the defendant may want to focus on in an attempt to demonstrate that the project could have been managed more efficiently. Therefore, the defendant may ask for alternative portfolios of managerial decisions to be simulated in the model and hence alternative control policies for the system need to be considered. However, in doing so, the interest lies in the ability of SD to enable an improved *understanding* of the system rather than to support *change* in the system.

In summary, alternative system structures and control policies exist in the systems being discussed in this thesis. Although they are not the main focus of the modeller's attention, the modeller may need to consider the impact of different managerial decisions on the outcome of the project.

When considering this criterion it has been concluded that rather than it being used as a reason to *reject* SD as a modelling approach, it could be used as a reason to *accept* SD as a modelling tool. This is because SD is suited to the controllability of managed systems. However it has been shown that this is not the only reason to use SD as it has been used to meet other types of modelling objectives.

4.7 Criterion 5: "Can it be done?"

As Coyle indicates, this is a rather weak criterion in the early stages of an investigation. It is weak in the sense that only an 'intuitive' answer can be given as a response. Practically, however, it is very strong. If a modeller has no sense on how a problem could be tackled using a modelling approach, then that is a good enough reason for the modeller not to use the particular modelling approach. Of course, the answer to the question could depend upon the modeller's experience and skill base. Such a situation complicates the use of this criterion and it is assumed that this discussion is outside the scope of this thesis. It is therefore assumed that the modeller in question has sufficient SD modelling experience to be able to make an informed decision regarding this criterion.

Coyle (1977, p358) intends the modeller to stand back and seek "an intuitive feel for whether system dynamics is likely to be the best approach". It may prove difficult to be able to judge if SD is the *best* approach to use. However, at least forming an opinion upon whether it is a *viable* approach is a useful test to carry out.

At this stage, to be able to form an opinion on the suitability of SD, Coyle suggests that the modeller looks for likely technical and behavioural snags and problems and asks "how should it be done, if at all?". If, even at an early stage, the modeller can form a mental picture of, say, the main flows and levels of the model, or which areas are going to prove the most difficult to model, then being able to think about the problem as practically as this, probably means that the use of SD is possible. As mentioned above, if the opinion is formed that 'It *cannot* be done', then this is a good enough reason that SD *should not* be used as the modelling approach. It should be noted that the modeller's experience will influence what the modeller can and cannot visualise. However, any model is influenced by the modeller's experience (Ackoff and Sasieni 1968, Rivett 1972, Checkland 1981, Mitchell 1993, Pidd 1996) and this needs to be considered before SD is rejected as a modelling approach in any situation.

The actual use of the SD modelling process in D&D claim situations (Cooper 1980, Weil and Etherton 1990, Ackermann et al 1997) proves that SD *can* and *has* been used to analyse specific cases of D&D for litigation. However, in this thesis the projects specifically being considered are large aircraft modification projects. Therefore the only question that can be explored is whether or not SD can practically be used to explore D&D in litigation cases for this type of project.

For the projects discussed in this thesis the author was involved from the beginning of the projects and so went through an initial process of considering whether or not SD could be used to model the D&D in the projects. This process involved a consideration of whether or not it could practically be done.

4.7.1 Castle Project

In the case of the Castle project, the first contact the modellers had with the client took the form of a two-hour meeting with a group of managers from the project. At the beginning of the meeting, the group was asked to summarise the events that occurred that they believed contributed towards the time and cost overrun in the project. After asking this initial question, the modellers allowed the group of managers to freely carry out a brainstorming session. The modellers only interrupted the group when points of clarification were required. This was deliberately done so that the modellers allowed the managers to form their own opinion of what they believed to be the important factors in the project overrun. Also, this would ensure that the modellers' experiences of analysing other project overruns were not allowed to influence the beliefs expressed by the managers.

After the initial meeting the author individually reflected on the material gained from the meeting and considered whether or not a SD model *could be done*. Using only the material gained from the two-hour meeting, an initial cause map was constructed to highlight any feedback loops that existed in the system. Following this, the structure of a SD model that could potentially be used to model the information gathered from the initial meeting was constructed. This approach was taken to enable an initial conclusion to be reached on whether or not a SD model would be an appropriate way in which to model the system based on the initial information gained. This process meant that the modeller was able to consider what she perceived to be the main levels and flows of the system. This process was carried out relatively easily. The initial SD model structure can be seen in figure 4.1 below.

At this stage, the modeller was able to form the opinion that a SD structure *could* be visualised and constructed for the project. Any potential problems with the population of the model with data were not specifically considered at this stage. A discussion on possible issues with data and information for the model is covered in the next criterion in section 4.8.

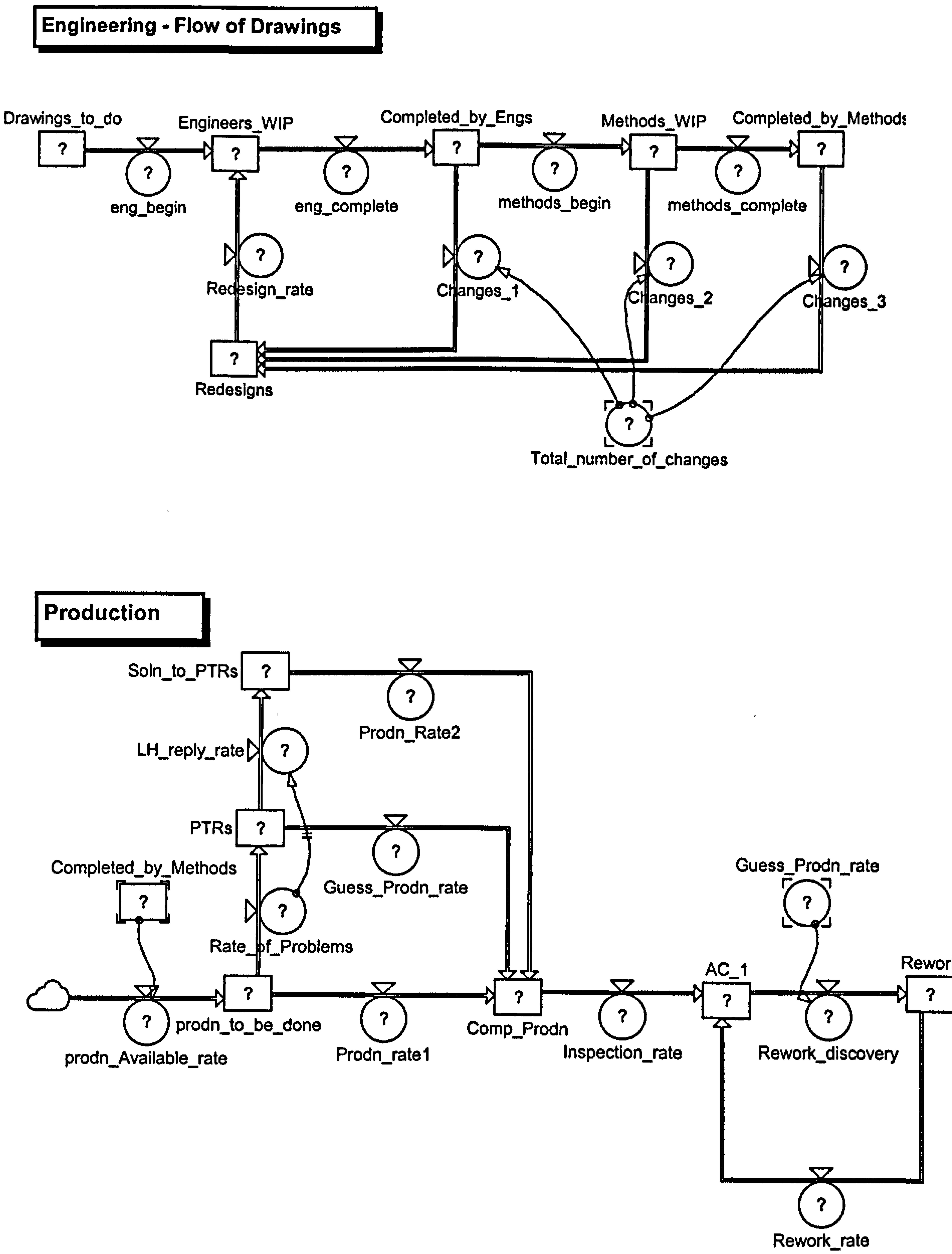


Figure 4.1 – The Initial SD Model Structure for the Castle Project

The main issue with the process described above is that, when summarising the events of the project, the client group would have been focussing on issues that were the main drivers of the project overrun in their opinion as the plaintiff. As no data was involved at this early stage, there was no evidence to support the overview being given by the plaintiff. A biased view of the events of the project would certainly have been gained. Also, since the initial meeting was relatively short, the information gained may have only focussed on a subset of the events that occurred during the course of the project. Therefore, although an initial SD model structure was constructed, the author had to bear in mind that this model may have ended up being non-representative of the full information that would be gathered on the project. The consideration of the SD model structure therefore became an iterative process. After each round of information gathering the author returned to this issue and asked herself the initial question of how she would now visualise the main levels and flows to ensure that the use of SD was still appropriate to model the information that had been gathered to date. In the case of the Castle project, each time new information was gathered by the modellers, the change in the way in which the author perceived the main flows and levels of the SD model was minimal. This can be observed in the comparison between the initial and final SD structures. The final SD model structure can be seen in figure 4.2.

Of course it may well be the case that, when considering any new information, the author was heavily influenced by the initial SD model structure when revisiting how to structure the system. However, this is not of great importance when investigating the criterion *can it be done?* The important issue here is that it was still possible to perceive how it could be done.

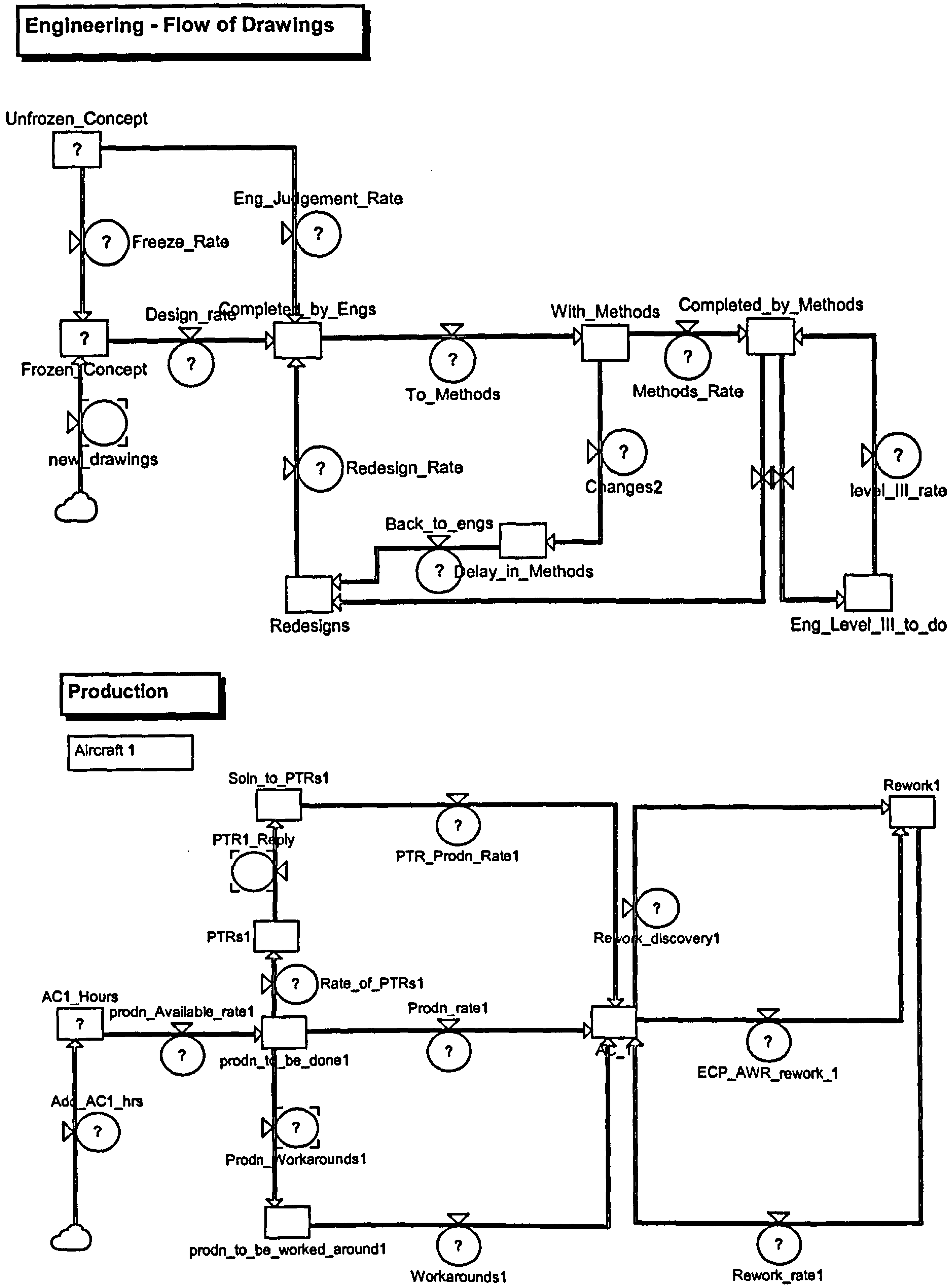


Figure 4.2 – The Final SD Model Structure for the Castle Project

4.7.2 Pirate Project

In the case of the Pirate project, an initial SD model structure was also constructed early in the modelling process. This can be seen in figure 4.3 below.

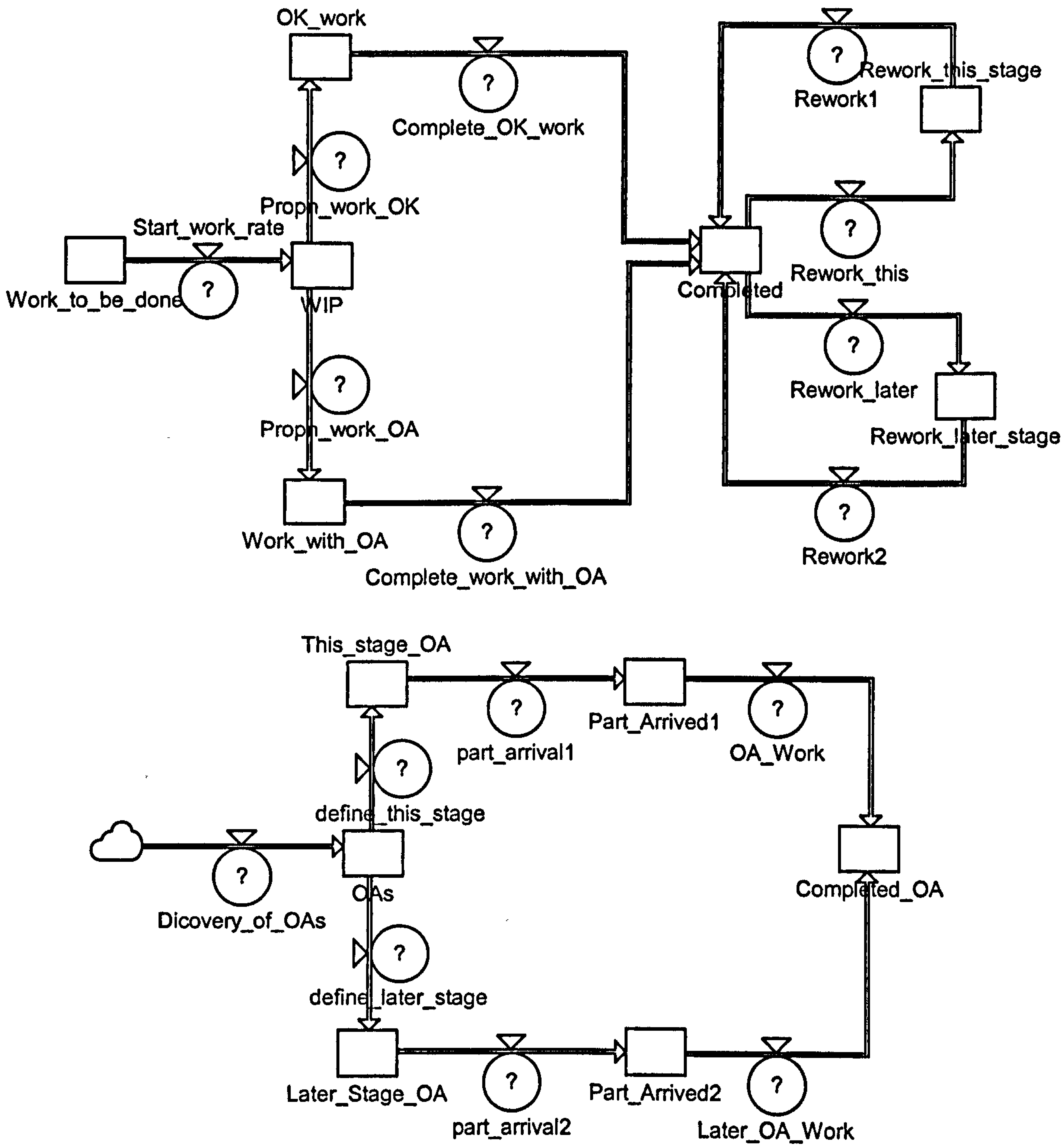


Figure 4.3 – Initial SD Model Structure for the Pirate Project

The initial structure demonstrated the modellers' initial thoughts on the main stocks and flows required to model D&D. However, in progressing through the modelling process, additional information and an improved understanding of the project meant that the modelling approach was altered. The main reason for this was that, as discussed in section 3.3.4, it was concluded that a large proportion of the cost overrun could be taken account of without any analysis of the effects of D&D. As it appeared that the amount of hours due to the effects of D&D may have been minimal, the effort involved in constructing a model to fully explain the effects of D&D outweighed the potential benefits.

Although a SD model was not constructed to fully explain the effects of D&D in the Pirate project, the question that still remains is *could it be done?* In considering this question, the main disruptive triggers that occurred in the Pirate project would need to be modelled. As noted in section 3.4.2, a majority of the D&D triggers led to additional hours being required which could be dealt with as an input to a SD model. This would therefore be relatively simple to model in a SD model. Section 3.4.2 also noted that the feedback loops underlying the Pirate project were dependent upon the use of managerial actions to compress the project. Therefore, any model of D&D would need to capture the compression that took place during the project. The managerial actions that were taken during the project were for example, hiring new staff and the use of overtime. Neither of these are seen to create any unmanageable problems when constructing a SD model. Further managerial actions undertaken were the rescheduling of individual production activities. Capturing this type of change is not so straightforward in a SD model. However, this is not unique to the Pirate project and indeed occurs in many projects that are compressed. The general problems of capturing this in a SD model are discussed in more detail in section 6.4 when discussing the quantification of the outcome of D&D.

This above discussion suggests that if a SD model of D&D had been required in the Pirate project, it would have been possible. However, this conclusion has only been based on the possibility of constructing a SD model structure, it does not take the population of the model into account.

The conclusion for both the Castle and Pirate project is that they could be visualised, at an early stage, in terms of stocks and flows and this did not prove to be a difficult process. The author also found this criterion a useful criterion to use during the modelling process. It provided a good practical test that can be used as a continual check as new information is gained throughout the modelling process. For this reason it is a criterion that will be used to form part of the revised set of criteria.

In this section, the criterion *can it be done?* has been used to consider the layout and visual structure of a SD model. However, ensuring a model can be *structured* is not a modeller's only concern. The model also needs to be populated to be successful. Coyle states that the criterion *can it be done?* is clarified by the more detailed and practical responses required to criteria 6 – 9. These criteria will now be discussed in turn.

4.8 Criterion 6: "What data and information are available?"

If it is not possible to populate a model through lack of data or information, then depending upon the purpose of the model, this may render the modelling *unusable*. The projects described in this work have been built as part of litigation claims. During such projects, large amounts of explicitly valid and verifiable data and information have been collected in preparation for a legal case. If further data is required as part of the modelling process, it is often given high priority within the organisation. The reason for this is due to the value of the litigation claim to the organisation. Any claim settlement can be viewed as a source of a large amount of direct profit to the organisation. For example, the smallest claim made as part of the projects discussed in this thesis was for £17million. Therefore, any help the organisation can provide in the construction of a robust model to aid their case is certainly worth their while. The value of the model to the organisation can mean that data collection in this type of environment is given a higher priority than other types of modelling that may be carried out for an organisation. Other forms of modelling may be viewed as beneficial to the organisation, but not as a priority with a direct financial benefit. However, in litigation claims it is desirable to gain as much of the data required as possible, since the more data collected, the more potential evidence there is to back-up the plaintiff's case. A lot of resources and attention are therefore normally placed on data collection. In fact, if the plaintiff is not willing to commit the resources required to enable a thorough data collection, then this could be extremely detrimental to the model construction. In particular, it could be a direct cause of the process not being sufficiently thorough to support the following modelling purpose:

- *All the above have to be demonstrated in a way which will be convincing to the several stakeholders in a litigation audience.*

Although it is argued that a lot of resources are often placed on data collection in a litigation environment, it may be the case that the data required simply does not exist within the organisation, is too subjective to enable accurate collection or proves extremely difficult or time consuming to gather. The question which the modeller and client need to ask themselves in this situation is whether the model, without the required data, puts the plaintiff's case in a stronger position than not having the model at all. If the answer to this question is 'yes' then it is still worthwhile to construct the model using, for example, expert opinions or judgmental views from those involved in the project in place of the data that is unavailable. However, eliciting expert judgements can prove difficult, especially for very subjective variables. The modellers need to be confident that the judgements that are gathered from experts are reasonable, non-biased estimates that improve the information available for the model.

When considering the projects discussed in this thesis the question that will now be asked is whether or not data and information were available in the Castle and Pirate projects which enabled models to be constructed that put the plaintiff's case in a stronger position when compared to not having the models.

4.8.1 Castle Project – Cross-Impact Data

One of the feedback loops observed in the cause map for the Castle project is *rework* leading to *cross-impact* (on other work) leading back to *rework*. This is a feedback loop that arises in many large engineering projects. Due to the interrelationships between parts of a product, when rework occurs on one part of a product there is often a knock-on effect on other parts of the product. The result of this is that the other parts of the product may also require rework. This rework, in turn, may cause other parts of the product to be affected and so on. Although cross-impact occurs in many projects, attempting to gather data to be able to quantify the effect can prove difficult. Attempting to gather data to quantify the cross-impact effect during the engineering phase of the Castle project was a task that proved very difficult for the modellers.

The data that the modellers were attempting to gather in order to be able to populate the engineering cross-impact effect in the SD model were:

(i) The average number of engineer hours required to rework a drawing.

(ii) For each hour of rework that was required on an initial drawing, how many drawings on average would be affected due to cross-impact and how many hours of rework on average would each of these require.

However, this data was not directly recorded during the life of the project. Indeed, due to the complex nature of cross-impact, it would be very difficult to distinguish, and therefore record, which drawings were directly affected by changes to a specific drawing. It was therefore decided that a sample of reworked drawings would be randomly chosen. From this sample of drawings, the average hours reworked and the average number of drawings affected by cross-impact would be determined. Unfortunately the modellers and the plaintiff's organisation were based in different countries. Therefore, due to the difficulties of the modellers getting quick access to the project's information database, a member of the plaintiff's organisation carried out this analysis for the modellers.

The major problem with attempting to carry out such an analysis on reworked drawings was that the plaintiff's analyst was unable to adequately trace which drawings impacted which other drawings. Some of the problems in gaining the estimates required were reported by the plaintiff's analyst as the following:

- The designer's times were not recorded in a manner that would permit the estimates required by the modellers.
- The database did not track drawing revisions prior to a particular type of revision. This would make it very difficult to assess the events leading up to certain revisions.
- Most of the designers who worked on the project were no longer working in the same part of the organisation and in fact many had left the organisation.
- If the estimates were to be calculated, it would have been necessary to assign a senior designer to the work who can study a number of changes and in each case access the necessary drawing database to review each drawing impacted and estimate the time necessary to make the change.

- However, the way in which a designer would tackle the problem in retrospect may bear little relation to how it was originally approached. The conditions could not be duplicated.

Due to these problems, the information requested by the modellers was not available. However, the plaintiff's analyst was able to suggest the following data as the best surrogate:

(i) The number of drawings that were impacted in total from a number of engineering changes.

(ii) The average time it took a designer to rework a drawing by discovering the total hours reworked for all the drawings affected by all of the engineering changes being analysed.

This meant that rather than being able to discover how many drawings were impacted each time around the rework /cross-impact loop, only the total number of drawings impacted by this effect could be obtained. There was therefore insufficient information for the SD model to be able to exactly simulate the rework/cross-impact loop. Instead, when an engineering change occurred, the model would trigger a one-off effect where the total amount of drawings impacted by cross-impact would be immediately identified as requiring rework. This meant that the feedback loop between rework and cross-impact was effectively being *erased* from the SD model and the effect of the loop was being quantified externally from the model. The modellers initially felt very uneasy as feedback was being erased from the model. However, it must be noted that this was only done due to a lack of data and the modellers found this to be an acceptable way forward as they believed that the plaintiff's case was still in a stronger position compared to the plaintiff not having the model at all.

When reflecting on the reasons for the inability to get the data required, the following points can be made:

- The modellers found it difficult to communicate exactly what they required from the plaintiff's analyst. Various attempts were made to ask the question in different formats. However, the plaintiff's analyst misinterpreted many of these. This was not helped by the fact that the modellers and the analyst were in different countries and so contact was often limited to e-mails and telephone calls. The modellers felt that if they had been able to sit down with the analyst and the database of information, a quicker resolution could have been arrived at. However, it is not surprising that the modellers had such difficulty in explaining their needs as people find feedback a difficult concept to understand (Sterman 1989a, 1989b; Paich and Sterman 1993, Diehl and Sterman 1995). Modellers should be aware of such issues and take them into consideration when formulating data requests.

- There was a lack of availability of engineering personnel who had worked on the project. The majority of those involved in the project had either left the organisation or had been transferred to other divisions. This meant that when going through the data, the plaintiff's analyst was unable to gain help by those who had been involved in the creation of the data and were therefore familiar with the data. This was bound to have an impact on the interpretation of the data.

Although the outcome of the above work was that the rework/cross-impact loop was erased from the cause map, in reflection this was not a particularly negative outcome. By representing the cross-impact effect through a feedback loop this assumes that the cross-impact effect from any initial piece of rework will have an infinite knock-on effect to the reworking of drawings. However, in reality, each piece of rework would only have a finite number of knock-on effects.

Modelling the cross-impact effect as close to reality as possible is an important part of the modelling of D&D in a large engineering project. If there are an infinite number of levels of knock-on from an initial piece of rework, then this will delay the overall completion of drawings beyond what would actually occur in reality. When the completion of drawings is delayed, this will have a knock-on effect to manufacturing, where the commencement of work will be delayed. This means that the modelling of cross-impact as an infinite feedback loop can cause an overestimation in the impact that cross-impact has on the overall progress of the project. However, since the rework/cross-impact feedback loop was not modelled in the Castle project, this potential overestimation was avoided. Instead, all the drawings requiring rework due to cross-impact were identified immediately. However, the effect that this does not capture, due to the lack of data, is that in reality different levels of drawings will be reworked in turn. This means that there will be some delay between the initial piece of rework and the final drawing affected by cross-impact. Therefore, in reality, the cross-impact effect lies somewhere between an infinite and a one-time knock-on effect to drawings.

If the data had been available in the Castle project, then an improved way of considering cross-impact in such a modelling situation would be as follows:

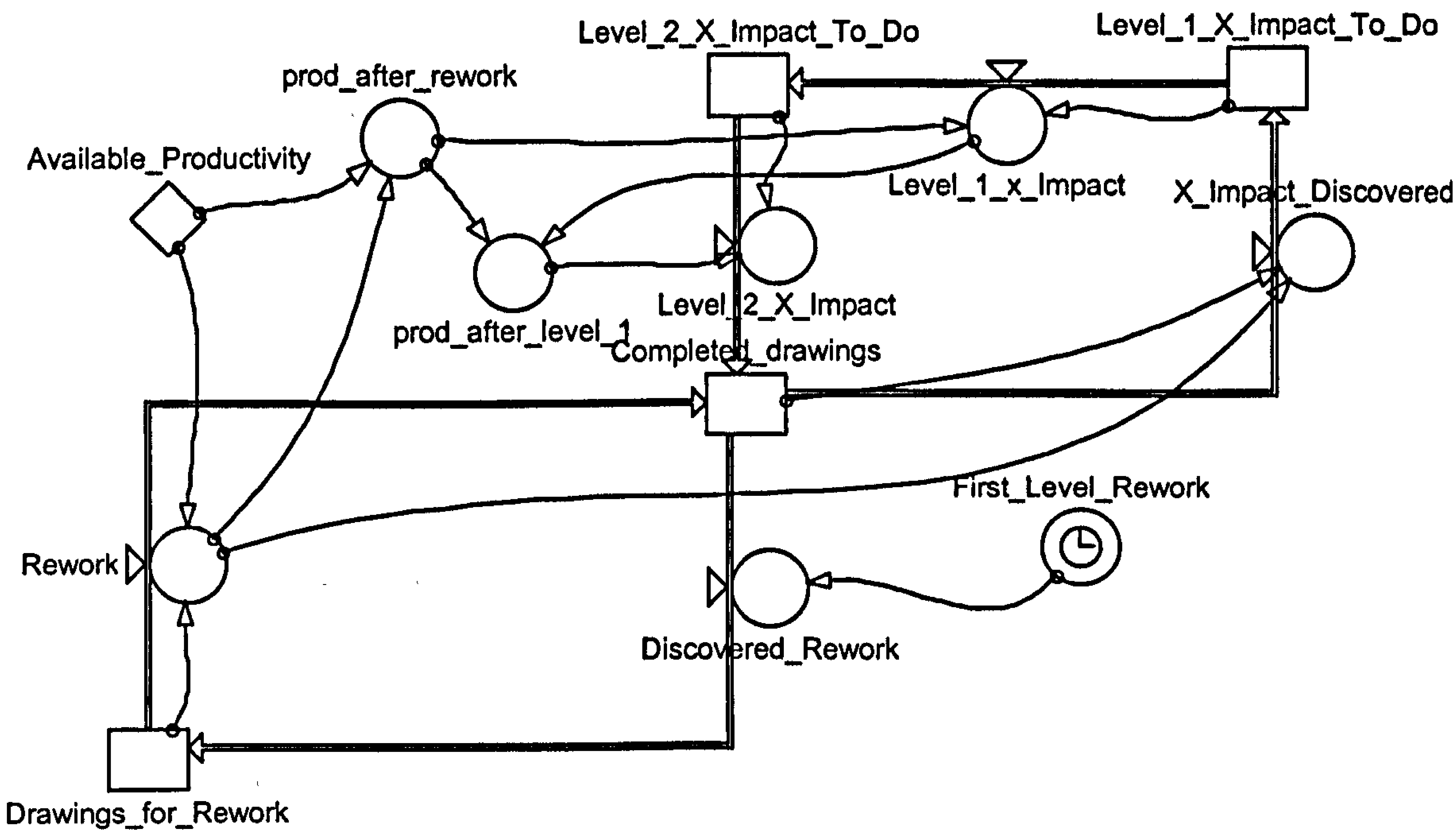


Figure 4.4 – Improved Structure for Modelling Cross-Impact in the SD model for the Castle Project

Figure 4.4 assumes that, on average, the cross-impact effect impacts two levels of drawings. This number would obviously be project specific.

The following figure represents the way in which the rework/cross-impact loop would be modelled if the infinite feedback loop was to be fully translated into the SD model:

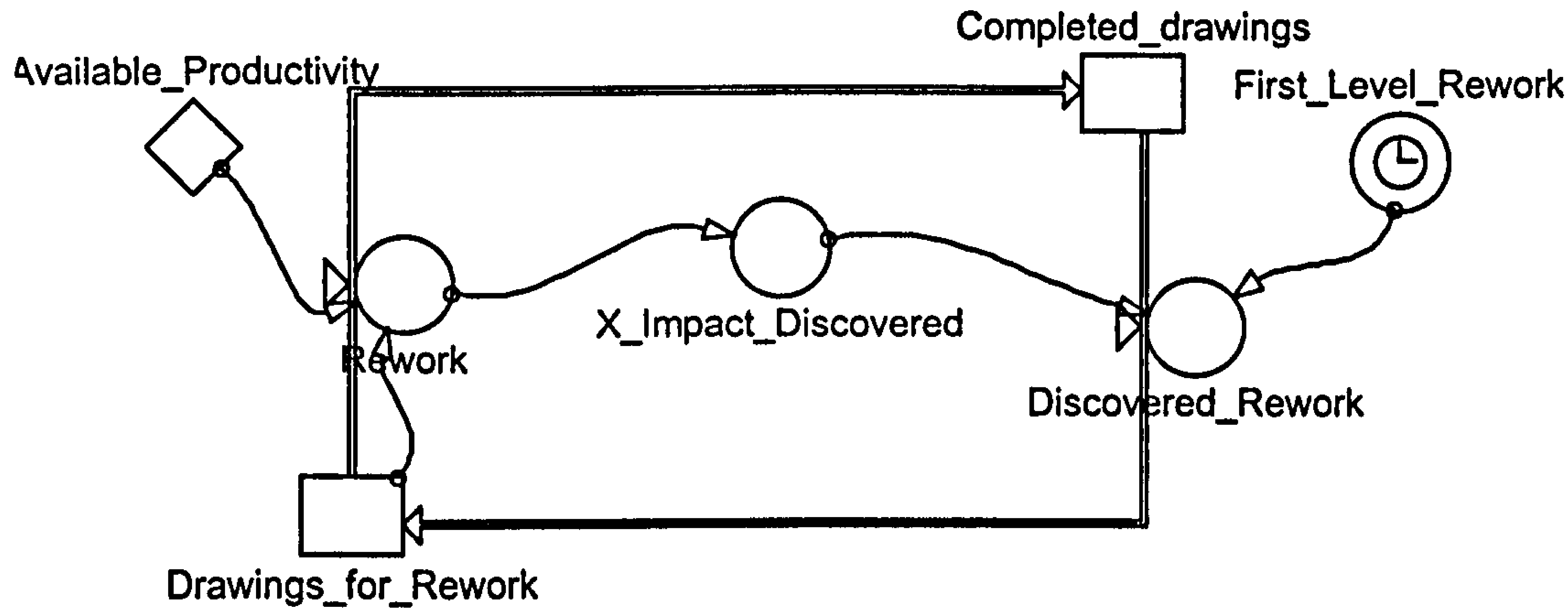
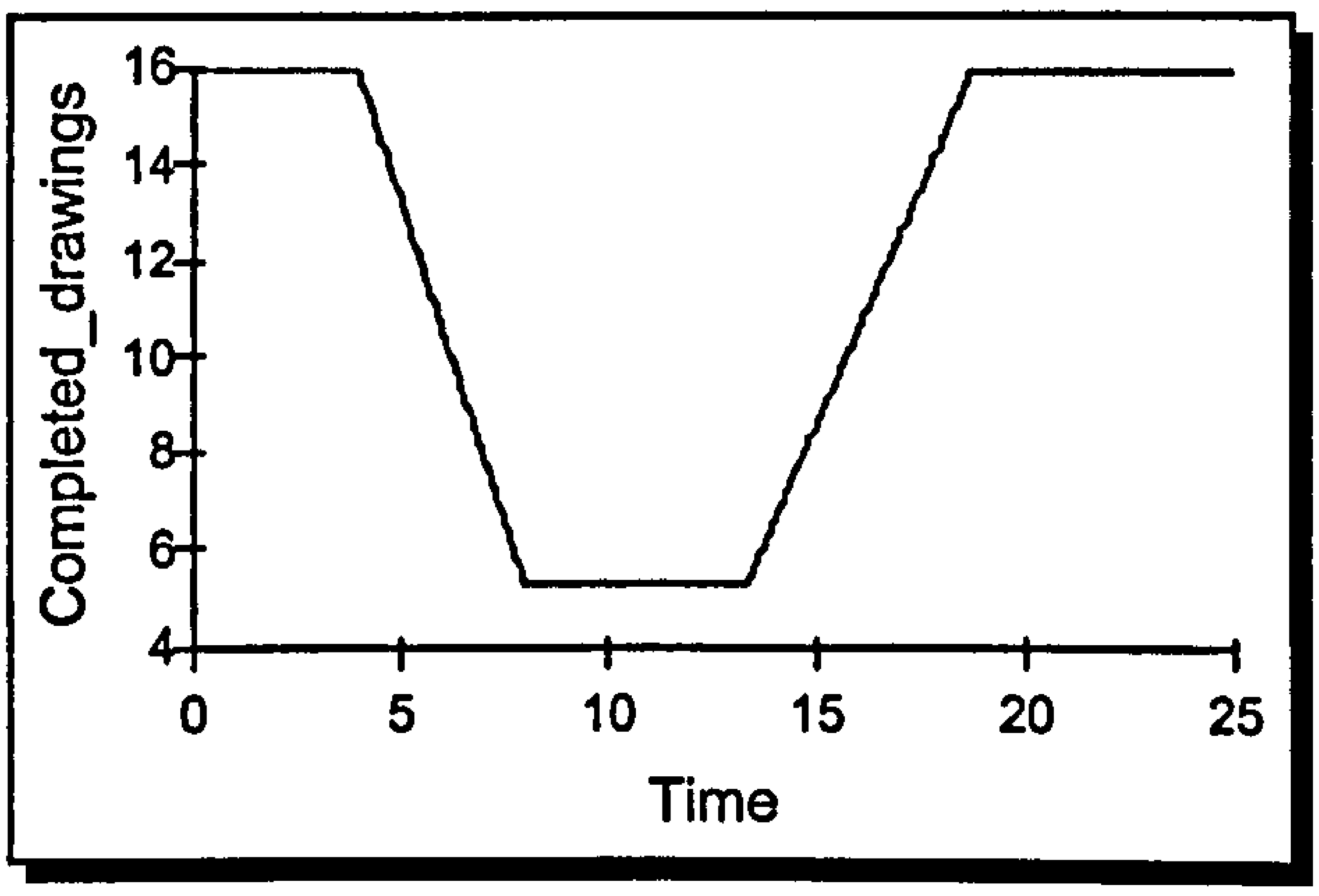
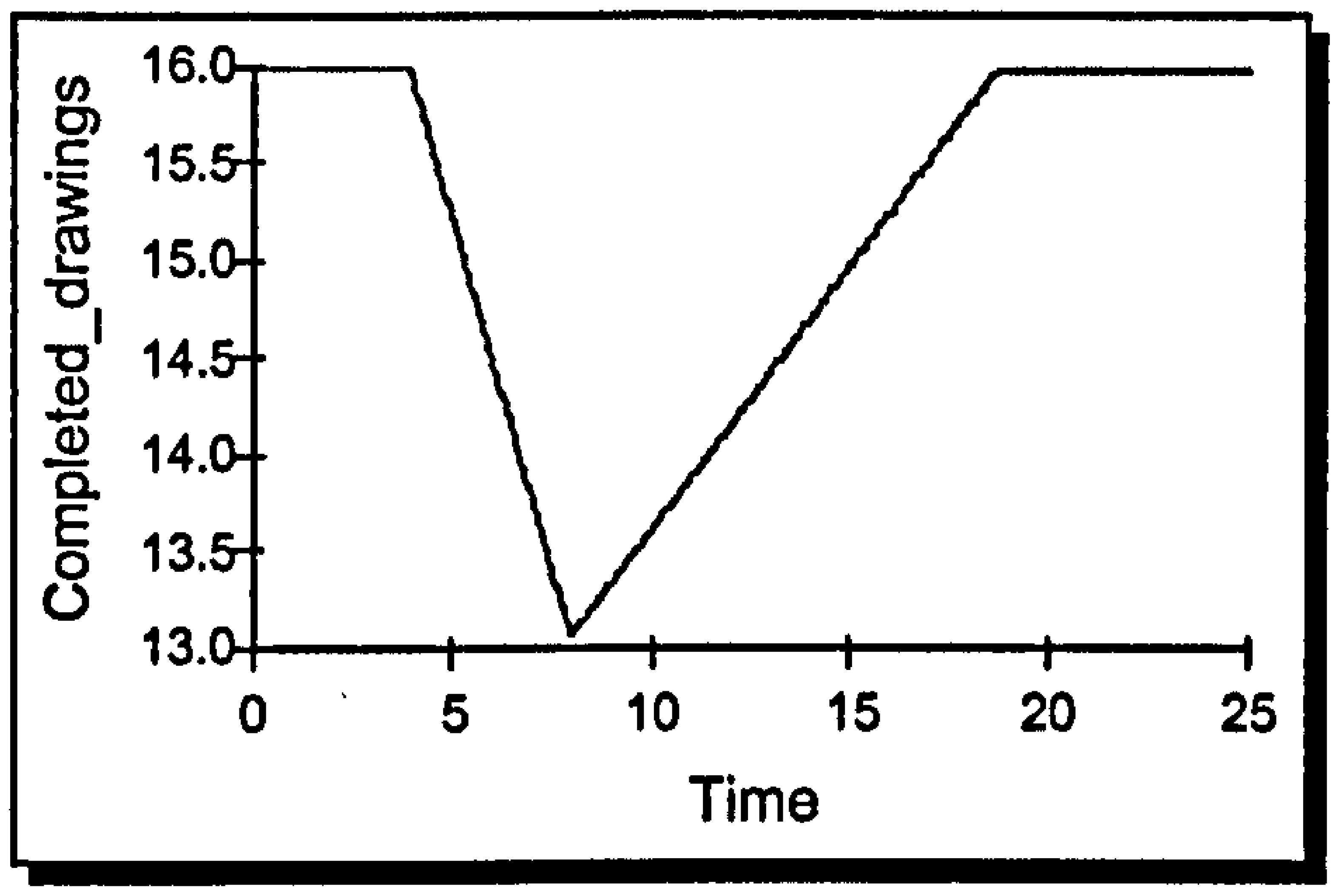


Figure 4.5 – Full translation of the Cross-Impact Feedback loop into the SD model for the Castle Project

If the models contained in figures 4.4 and 4.5 are populated with example data (there are 16 completed drawings at time 0, 1 drawing per week requires rework between times 4 and 8 and rework can be completed at the rate of 1 drawing per week), then the following results can be achieved when the level 'Completed_Drawings' is observed over time:



Graph 4.1 – Output from the Model in Figure 4.4



Graph 4.2 – Output from the Model in Figure 4.5

The completed number of drawings shown in graph 4.1 reduces more dramatically between time 4 and 8 when compared to graph 4.2. This illustrates that the rework caused by the cross-impact effect is identified earlier in the improved model structure shown in figure 4.4 when compared to the model structure shown in figure 4.5. The earlier identification of the rework also has an important effect on the time taken to complete all the required rework on the drawings. The rework appears to be completed by approximately time 18 in both graph 4.1 and 4.2. When examining the actual numbers of completed drawings for each model, all drawings are indeed completed by approximately 18 time units for the model shown in figure 4.4. However, this is not the case for the model shown in figure 4.5 since a very small fraction of the drawings continues to be sent for rework throughout the simulation of the model. As previously mentioned, the importance of this outcome is the implications that this has for the availability of drawings for the manufacturing function. If the completion of drawings is delayed in the engineering function, then this may cause a delay in the commencement of the work to be carried out in the manufacturing function. By modelling the rework/cross-impact loop as a feedback loop in the SD model may therefore cause an overall delay in the model that did not exist in the real system.

The author has found that in the projects discussed in this thesis, the information recorded about projects is generally insufficient to be able to gain all the data ideally needed to construct the types of models required as part of a litigation case. However at the beginning of a project, organisations do not normally plan to be in a litigation environment and therefore do not usually record the level of detail required as part of a litigation claim. This was noted by the plaintiff's analyst in the Castle project when he stated "we just do not hold the information you require ... we would never normally need it ... it's only because it's a claim that it's required". However, this is a reasonable choice for an organisation to make. In order that data can be recorded to the level of detail most useful to a litigation process, assuming this can actually be done, a lot of resources would need to be allocated to this process. However, this is not normally a priority task in the management of a project since detailed recording of data is not required in order to complete a project and would therefore be seen as a wasted effort.

When dealing with the issue of a lack of data when modelling cross-impact in the Castle project, it can be concluded that if it is not feasible to gather the data ideally required to populate a model, then it is important to discover what data *can* be gathered. An assessment can then be made on whether or not the data is sufficient to place the client in a stronger position when compared to the client not having the model at all. If it was the case that the data could not support the use of the model, a consideration then needs to be made on whether another form of modelling may make better use of the data that is available and be able to place the client in a stronger position.

4.8.2 Castle Project – Reliability of Subjective Data

The criterion explored in this section should not only ask if data and information are available but if the data and information available is also *reliable*. This was an issue that needed to be considered when carrying out interviews with those involved in the Castle project. Interviewing was an important part of the data gathering for the Castle project. Since the data gathering was carried out as part of a litigation case, those involved in the project were potential witnesses and therefore any information that they could provide on the project may result in being used as evidence in the litigation case. Many of the staff involved in the project had been working on large engineering projects for many years. However many had never before experienced the amount of time and cost overrun experienced on the Castle project. Due to this, staff were left very demoralised from the experience.

When D&D is rampant, managers will report that the resulting complexity is akin to chaos (Eden et al 2000, p296). At one point in the Castle project, there were so many changes occurring on the project that staff on the Castle project stated that they had felt that they had almost lost control of the project, that no matter what they did, large amounts of rework simply kept on occurring. This was a statement made by managers with many years of experience who were beginning to lose confidence in their work, beginning to almost believe that they were to blame. The reasons for this were that the large amount of changes occurring meant that it was impossible to track the original trigger of these changes. Therefore, the project almost seemed to be in chaos. This meant that whilst gathering information from staff, they had to be continually reminded that they were not the ones to blame and that they should not feel discouraged by the experience. Cause maps were useful in explaining to the staff how many of the events that they had experienced could be traced back to identifiable causes attributable to the defendant.

When gathering data from those involved in the project the reliability of the data needs to be considered. Influences such as demoralisation can affect the way in which those involved in the project recall the events of the project. It is important to attempt to get as honest an opinion as possible from staff on the events that occurred during the project. However, demoralisation meant that they began the process often unfairly feeling as though they were to blame.

4.8.3 Pirate Project – Reliability of Recorded Data

The reliability of data also needed to be questioned during the Pirate project. The worst situation occurred when the modellers requested information on the labour force during the life of the project. The modellers were told that this information was held with both the personnel department as well as wages department. On receiving information from both the departments, the modellers discovered significant discrepancies between them. Also, both sets of information contained data that was clearly incorrect. For example, there were cases where the leaving date recorded for a member of staff was prior to their commencement date.

The type of data discrepancies described above are certainly not unique to the Pirate project. Gathering data for any modelling process is never usually a straightforward exercise. In particular, when organisations hold a mass of data on a large number of people, errors are bound to occur in the recording of the data. However, when collecting data to support a model for a litigation case, particular emphasis should be placed on the reliability of any data for the following reasons:

(i) A model may be put under detailed scrutiny from the defendant. This means that any data used to populate the model will also be placed under detailed scrutiny.

(ii) Those people providing data about the project are likely to have to produce witness statements as part of the litigation case or may even end up in the witness box as part of a court case. Any unreliable information that is discovered from witnesses can have a damaging effect on the plaintiff's case.

(iii) Contradictions between data should be avoided. If any contradictions are spotted by the defendant's and brought to an audience's attention, then this can create doubt in the plaintiff's case. If a number of such situations are found, then this can lead to a build-up of doubt in the validity of the plaintiff's case.

The most important issue is that, if discrepancies are discovered when gathering data for a model, it is possible to be able to resolve the discrepancies. This was the case in the Pirate project.

From the experiences of the Castle and Pirate projects, when considering the data and information available on a project a modeller should ask themselves the following question:

- *Is there sufficient, reliable information and data available which will enable a model to be populated to such a level that the plaintiff's case is in a stronger position when compared to the plaintiff not having the model?*

Although there were issues about the availability and reliability of data on both the Castle and Pirate projects, the ability to resolve many of the issues meant that a model was able to be constructed to such a level that the plaintiff's claim team believed that their case was in a stronger position when compared to not having the model.

4.9 Criterion 7: "Can we define the variables?"

In stating this criterion, Coyle refers to the difficulties in ensuring that everyone involved in the modelling process understands and agrees with each of the variables used. The importance of this criterion is summarised by stating that "it is essential to pay careful attention to the problem of defining what is meant by each of the variables, or the people who have to be convinced of the virtue of the recommendations will either not understand them, or not believe them, and the project will be futile" (Coyle 1977, p358). This suggests that those involved with the modelling process need to fully understand the variable definitions. However, when modelling D&D for litigation there is also a need to ensure that the wider model audience understands the model. The model's wider audience refers not only to the plaintiff, but potentially also the defendant, any experts hired by the plaintiff or defendant, a judge and a jury.

When Coyle discusses this criterion, he is referring to agreement on (i) whether a variable should be included in the model i.e. within the system boundary and (ii) agreement on how a variable is to be defined or modelled. However, in considering this criterion, Coyle was reflecting on the importance of getting buy-in from a number of people who would potentially be basing future decisions on the model. This means that they need to believe in the structure of the model so that they can believe in the model's output. However, when modelling D&D in a litigation environment the importance shifts to convincing an audience that was not part of the modelling process that the model adequately represents reality. This audience will need to be convinced that the system boundary includes all the variables that should be there and that they understand the definition of the variables that are included in the model and agree with their definitions.

Although the wider audience has been defined as the defendant, any experts hired by the plaintiff or defendant, a judge and a jury, in reality any experts hired by the defendant can play a significant role when attempting to convince the audience that the model adequately represents reality. These experts are normally hired by the defendant to make an expert judgement on whether or not the model can actually prove what the plaintiff states it does. If the expert is convinced then this will generally feed down to the wider model audience.

Coyle introduced the criterion in this section as a way of ensuring that those who are involved in the model building will believe in the model so that they believe in the model output and base future decisions on it. However, in a D&D litigation environment, this is not required as managers of the plaintiff's organisation will not be basing future decisions on the model and therefore the same buy-in is not required. Instead, this criterion is most useful as a check to ensure that an expert would be convinced by the validity of the model.

Although it is stated that the importance in modelling D&D for litigation switches to convincing those not involved in the modelling process, it is obviously still worth ensuring that all those involved in the modelling process can agree on the variables included in the model, otherwise the model may never reach a wider litigation audience. From the criterion explored in this section, the following question needs to be asked:

- *Are there any variables which would be so difficult to agree (i) if the variable should be included in the system boundary or (ii) how the variable should be defined/modelled such that those involved in the modelling process lose faith in the model?*

4.9.1 Including a Variable in the System Boundary

The use of cause maps in the projects discussed in this thesis aided the process of agreement on the system boundary. For example a cause map was produced as one of the first steps for the Castle project in the modelling process. This was initially created from a draft claim document created by the client. The client was given the cause map and asked to comment on whether or not it captured all the issues that they believed to be attributable to the project overrun. In following this process, the cause map became a tool that could be used to check the coherency of the claim document. For the Castle project, the client did request additional variables to be added to the cause map. This not only helped to confirm which variables should be included in the system boundary, but also highlighted areas that were missing from the claim document. In other projects where the modellers have been involved at an earlier stage in the project, a draft claim document did not exist. In these cases interviews would be carried out in order that all the opinions of those involved in the project could be taken account of and brought together in a cause map (Williams et al 1995, Ackermann et al 1997).

Once a cause map has been constructed, the triggers to D&D are highlighted and the map reduced to those concepts required for a quantitative model of D&D (Williams et al 2000).

At this stage, any disagreement on the variables that should or should not be included in the quantitative model could be aired and discussed. The use of the cause map in this way provides a relatively transparent model to facilitate discussion. Also, it enables a certain level of agreement to be gained before the SD modelling commences.

The experience of the author has been that there has been little disagreement on the variables that should be included in the system boundary when modelling D&D in a litigation environment. This is likely to be due to only a small number of people actually being involved in the modelling process. For example, in both the Castle and Pirate projects the modellers maintained continuous contact with only one member of the plaintiff's team. The other team members (for example the senior manager and the lawyers) would be involved in less frequent meetings. However, the agenda for these meetings normally focussed on the way in which the overall model could be best presented to the defendant's to convince them that the model was a reasonable representation of reality, rather than being concerned with the detail of individual variables.

In summary, the cause map process used allowed the system boundary to be elicited and to enable a clear agreement to be reached on the variables that should be included in the litigation model. However, it should be noted that this process only involved a small number of people.

4.9.2 Defining a Variable

When considering the agreement of the definition of a variable it is worth noting that projects modelled in a litigation environment have been defined by a contractual agreement. Some of the baseline variables, especially those modelled in the original budget run, will be defined and quantified within this contract. This should therefore help to reduce any potential disagreement during the definition of the variables. Of course, the contract will not cover variables such as expected levels of fatigue and morale that are very difficult to quantify.

At the early stage of deciding whether or not SD should be used, this criterion could be used to attempt to identify any major disagreements that could arise during the course of the modelling process. For example, if variables such as fatigue and morale need to be modelled, a discussion, early in the modelling process, may be required with those involved to gain an agreement on the causes and effects of such variables.

Examples of the type of considerations that are required when defining a subjective variable are discussed next in the context of the Castle project.

Castle Project – Defining Productivity as a Variable

In the Castle project it was believed that low productivity levels played a significant part in the outcome of the project and therefore an agreement on the definition of this variable was important.

The level of productivity of workers on a project can have a huge impact on the progress of a project. When modelling productivity it is therefore important to gain agreement on what the variable actually represents. However, literature that discusses issues with productivity informs us that "While there is agreement that productivity is important, there is little agreement on what the term *productivity* means" (Pritchard 1990, p8). Pritchard (1990, p8) goes on to say that the term *productivity* has been used to cover different concepts such as efficiency, output, motivation and work quality. This means that when people discuss the term *productivity* they may have different understandings of what it means.

The difficulties in defining the term productivity also mean that it can be a very difficult variable to quantify. In an attempt to determine the impact that low productivity had on the project, a group session was held with a number of managers involved in the project on the plaintiff's side. Cause maps were used as part of this group session to aid the definition of productivity in the Castle project. In particular, those variables that caused productivity to change over time were highlighted. These variables can be seen in figure 4.6 below.

One way to help get a group of people to think about a variable in the same way is to define the dimensions of the variable and discuss anchor points. For the Castle project, productivity was modelled as a factor within the range 0 to 1.The important anchor point to define with this term was what 1 represented. In this case, 1 was taken to be the level of productivity that could be achieved if the project had progressed as had been expected in the budget. That is, the level of productivity that could have been achieved if no unexpected disruptions or delays had occurred.

Using a group session meant that any differences in the way in which members of the claim team defined productivity could be discussed and they could move towards reaching an agreement on how the variable should be modelled so that those involved in the modelling process would not lose faith in the model. However, this type of session only involves two of the model's audiences; the modeller and the plaintiff. It does not move towards getting buy-in for the model from the defendant, or any potential judge or jury.

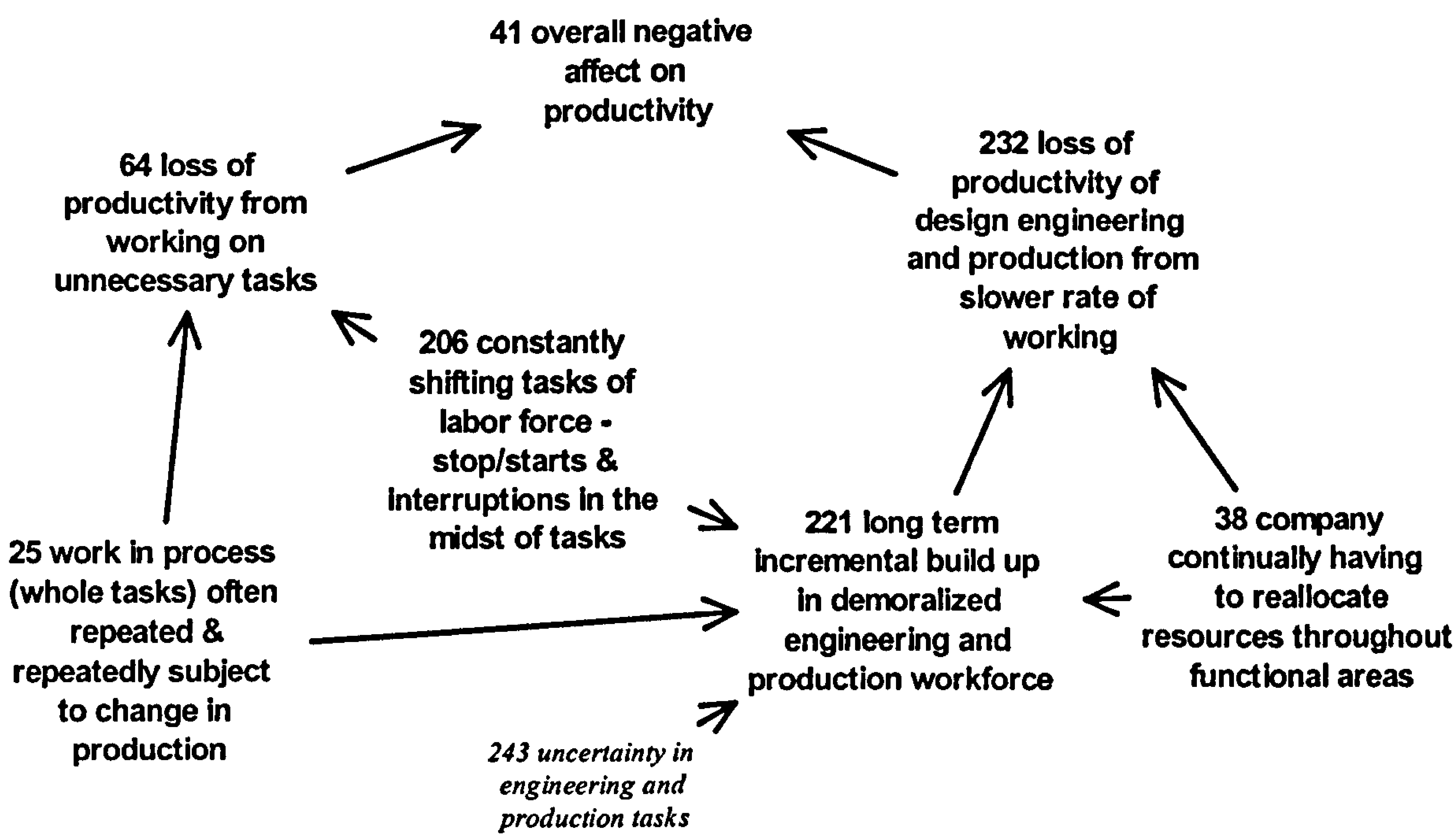


Figure 4.6 - The Variables that Affected Productivity in the Castle Project

In conclusion, the criterion considered in this section focuses on getting agreement between those involved in the modelling process. When modelling D&D in a litigation environment, agreement on variable definitions may only be needed between a small number of people from the plaintiff's claim team. Gaining agreement from a small number of people means that the potential level of conflict between the group should also be kept at a low level. Also, the focus when modelling D&D for litigation is not only to convince people on the plaintiff's side of the validity of the model, but also to convince those included in the wider audience i.e. the defendant, any expert modellers, a judge or potentially a jury. The criterion covered in this section therefore does not play as an important role as it may do in other types of SD modelling studies. The focus of model validity is directed more towards convincing the defendant of the validity of the model. This is an important consideration of the modelling process and will be covered in chapter 6. For these reasons, the criterion will not be used as part of the revised set of criteria being constructed as part of this thesis.

4.10 Criterion 8: "Where are the dangers of over simplification?"

Coyle (1977, p359) explains this criterion by noting that "A common cause of failure... is the oversimplification of the model, either because reality is too complicated to model, or because a realistic model is too expensive or too difficult to analyse." In order to address this issue a modeller needs to consider the purposes of the modelling process. In the modelling of D&D for litigation, one of the purposes is to:

- *Replicate over time the hours due to D&D which were over-and-above those that were contracted, but were required to carry out the project.*

This purpose means that all the causes of the need for labour hours on the actual project need to be replicated as closely as possible in the simulation model. (Note that the variations between actual and simulated labour hours are discussed in chapter 8). By requesting the replication of hours, this means that the modelling purpose is asking for specific quantitative results to be produced by the model. This type of detail is not a standard use of SD modelling. For example, Veit (1978, p535) states that SD models "...are basically strategy models... (their) purpose ... is not to make precise quantitative predictions of the future, but rather to indicate the trends of the key variables...".

One of the main reasons for SD focussing on modelling the qualitative behaviour of systems is due to the presence of noise in social and economic systems. Forrester (1961, p124) notes this point by stating that "even though stable organizational structure, policies, and human reactions exist and these determine the principal dynamic characteristics of a system, we cannot assume a *perfect* model in which *every* relationship is known *exactly*. Therefore, we are committed to models in which *every* decision function has, at least in principle, a noise or uncertainty component. By definition, the exact time pattern of this noise is unknown, and we have not discovered its generating causes.... Not knowing the instantaneous values of the noise, we can still study the *kind* of behaviour exhibited by the system". Following from this (Forrester 1961, p125) "A dynamic system model should therefore be expected to represent and to predict the behavior characteristics ... of the actual system. It should not be expected to predict future system state except to the extent that the system has continuity and momentum characteristics that will cause present conditions and trends to persist for a time in spite of noise disturbances." With these statements, Forrester is warning against using SD models to gain precise *future* values of a system. However, this has not been the intention of the use of SD models in modelling D&D in a litigation case the author has been involved in. The purpose is to be able to explain what actually *happened.* Therefore, there is no consideration of the *future* occurrence of noise, only the *actual* noise that has occurred in the past. This discussion is now taking the reader towards another important area in this thesis, that of the level of detail required to fulfil the purposes of the modelling process. This will be discussed in chapter 6.

Since the model only requires to replicate noise that has already occurred, more precise, detailed SD models may potentially be feasible. When modelling D&D in a litigation environment the detail is important, as not only do labour hours need to be replicated by the model, but this also has to be done in a manner that will convince the audience that the model is an adequate representation of reality. This is so that another of the modelling purposes will be met:

- *All the above have to be demonstrated in a way which will be convincing to the several stakeholders in a litigation audience.*

Therefore, when modelling D&D for litigation, the dangers of oversimplification are that at least two important modelling purposes may not be met. However it should be remembered that any model is a simplification of reality, otherwise it would not be a model, but a full replication of reality. Therefore, this means a balance needs to be obtained between too much and too little simplification of the system.

When modelling D&D for litigation, the projects discussed in this thesis have not indicated that oversimplification is an issue due to the complexity of reality. However, a modeller may need to spend time and effort understanding a complex part of the real system to ensure that when aggregating elements of the system the SD model still correctly captures the nature of the project. The resources required to model the detail are also not normally regarded as expensive since priority is usually given to the model construction. Therefore, the criterion detailed in this section becomes more of a consideration during the modelling process rather than as a part of whether or not SD should be used. It is not something that has been seen to prevent the use of SD in modelling D&D as part of the projects discussed in this thesis.

4.11 Criterion 9: "What level of aggregation is needed?"

Coyle (1986, p 404) notes that SD is quite capable of providing as much or as little detail as the modeller requires and that this links with the previous criterion as "the other risk of oversimplification is that insufficient detail has been modelled" (Coyle 1977, p359). Also the danger here is that "an aggregated model will produce results which are completely different from a disaggregated equivalent" (Coyle 1977, p360). Thus, when building a SD model, the level of aggregation has to be monitored. However, as with criterion 8, this is more of a consideration during the modelling process rather than as part of whether or not SD should be used.

4.12 Criterion 10: "What facilities are needed?"

This criterion touches on the specific resources that will be needed in order that the modelling process can take place. The examples given by Coyle (1977, p360) consider the computer time available and manpower required by the modeller. This is obviously an issue for the modeller to consider in the planning stages of the process. However this is not of particular importance in a general discussion on the suitability of modelling D&D with the SD modelling approach.

4.13 Criterion 11: "What training will be required?"

Coyle (1977, p360) discusses this criterion in terms of training managers involved in the modelling process to enable them to fully appreciate what the SD methodology can and cannot be expected to do. In an ideal world this would certainly be an appropriate approach for members of the plaintiff's claim team. However in practice, members of the claim teams for the cases discussed in this thesis were senior members of organisations. Each of them did not rate this training sufficiently important to find the available time in their busy schedules. The training which they perceived to be appropriate were short presentations given by the modellers to give them a general understanding of the basics of SD. They perceived these to be sufficient as they perceived the modeller to be the expert on their side and any expert hired by the defendant to be the expert on the other side. Any technical dialogue would then take place between these two parties.

Again, in an ideal world, training in SD would be appropriate for the lawyers involved in the claim. The lawyers are the ones that need to be able to argue the use of SD in any negotiations or court environment. However, in practice, reluctance has been encountered due to the time involved in this and hence, in the cases discussed in this thesis, only minimal training was given to the lawyers by the modellers.

Training in SD is something that, ideally, the various audiences of the model should all gain. However, practicalities mean that this was minimal for the audiences in the cases discussed in this thesis.

4.14 Criterion 12: "How much can we afford?"

Section 4.8 noted the priority given by an organisation to collect data for a model for litigation due to the high value of the litigation claim. For this reason it concluded that a lot of resources and attention are normally placed on supporting the modelling process. Also, if the plaintiff is not willing to commit the resources required to enable a thorough modelling process, then this could be extremely detrimental to the model construction. If it was the case that a limited amount of resources was available, a modeller would need to question if a sufficiently thorough piece of work could be carried out to enable a model to be constructed that would put the plaintiff's case in a stronger position than if the model did not exist. One question the modeller may ask is if the available resources would support the construction of a cause map, but not a SD model. One of the purposes of a SD model is to help achieve the following modelling purpose:

- *Replicate over time the hours due to D&D which were over-and-above those that were contracted, but were required to carry out the project.*

In doing so, a lot of detail is required in order to adequately replicate the real system (see chapter 7). By only concentrating on the qualitative model, this would reduce the resources required. Of course, in making this decision the modeller needs to consider the requirements of the plaintiff. These requirements will include whether or not a qualitative plausible explanation of what occurred in the project is helpful to the plaintiff, or whether the plaintiff believes that having a model that produces a quantitative result is important.

As previously mentioned, the plaintiff is normally willing to spend resources on the modelling process, since any return on the claim is equivalent to direct profit to the organisation. If a problem does arise, it is more likely to be linked to the time available for the modelling process. This leads onto the next criterion.

4.15 Criterion 13: "How long have we got?"

Litigation can be a very long process. If the plaintiff expects to go to court then, assuming the modeller is brought into the process in the early stages, the modeller will have a reasonable amount of time available to carry out the modelling process. However, it is more likely that the plaintiff would want to attempt to negotiate with the defendant, to attempt to avoid the costs of a court case. In this case, the modeller may be asked to produce a model in a short time scale. The exact time scale will determine the type of modelling that will be carried out. For example, cause mapping may be possible, but advancing the model to a SD model may not. It will also determine the detail of the model. Even if the modeller only focuses on a cause map, a tight time scale may mean that the model captures a relatively small number of peoples' perceptions. In such a case, the model could be extended to make it more robust if additional time was made available to the modeller. The time spent creating the model will also determine the level of confidence the modeller has in the model and its results. The modeller will need to make a judgement on whether there has been adequate time to capture a sufficiently large number of perceptions to ensure that the events of the project have been reasonably reflected in the model. This type of judgement is subjective and based upon the experience of the modeller.

4.16 Criterion 14: "How about implementation?"

This criterion is of a very specific nature to the environment in which the modelling process is expected to take place. The criterion aims to get the modeller to consider the more practical issues such as knowing the decision-maker and understanding the specific politics of the situation.

4.17 Conclusions from Coyle's Criteria

The first three of Coyle's criteria do not add anything new to the investigations in this thesis beyond those criteria discussed in chapter 3. However, it should be noted that there was more detail included in Coyle's criteria than those introduced by Forrester. In particular, criterion 3 provides a modeller with practical advice on how to identify feedback loops in a system.

When considering the fourth criterion, i.e. are there any alternative systems structures or control policies?, it was concluded that although the criterion could not be used as a reason to *reject* SD as a modelling approach, it could be used as a reason to *accept* SD as a modelling tool. This is because SD is suited to the controllability of managed systems, but it has been shown that this is not the only possible objective when constructing a SD model. SD models have also been used to attempt to change the way that users of the model think.

The fifth criterion was considered to be very practical in nature and is based on the judgement, experience and skills of the modeller. This criterion has explored new material beyond the criteria discussed in chapter 3. Criteria 6 – 9 were also seen to be an extension of criterion 5.

When investigating criteria 5 – 9, revised criteria to add to those summarised in section 3.6 are:

- *Can the projects be visualised at an early stage in terms of stocks and flows?*

- *Is there sufficient, reliable information and data available which will enable a model to be populated to such a level that the plaintiff's case is in a stronger position when compared to the plaintiff not having the model?*

These two criteria summarise the additional contribution by Coyle to the revised set of criteria that will be brought together at the end of chapter 5. They can be summarised as very practical in nature and therefore very helpful criteria for a modeller to take into consideration.

When investigating the Castle and Pirate projects in light of criteria 5 – 9, the following conclusions were:

- Both the Castle and Pirate projects could be visualised at an early stage in terms of stocks and flows and this did not prove to be a difficult process. The author found this check to be very useful. It provided a good, practical test that could be used as a continual check as new information is gathered throughout the modelling process.

- Although there were issues with a lack of certain data and unreliable data in the Castle and Pirate projects, a model could still be constructed which the modellers believed placed the plaintiff's case is in a stronger position when compared to the plaintiff not having the model.

- From the experiences of the Castle and Pirate projects it was concluded that the consideration of gaining agreement on the definition of individual variables was not as important when modelling D&D for litigation as in other types of SD modelling studies. The focus of model validity is directed more towards convincing the various audiences of the validity of the model.

- It was noted that when modelling D&D for litigation, the dangers of oversimplification are that at least two important modelling purposes may not be met. However it should be remembered that any model must be a simplification of reality, otherwise it would not be a model, but a replication of reality. Therefore, this means a balance needs to be obtained between too much and too little simplification of the system. However, it was concluded that the consideration of the dangers of oversimplification and the level of aggregation in a model were more of a consideration during the modelling process rather than as part of whether or not SD should actually be used. These criteria had not been seen to prevent the use of SD in modelling D&D in the Castle or Pirate projects.

Criteria 10 – 14 highlighted more practical issues that need consideration. When exploring these criteria in terms of modelling D&D for litigation, the following conclusions were made:

- Training in SD is something that, ideally, the various audiences of the model should all gain. However, practicalities mean that this was minimal for the audiences in the cases discussed in this thesis.

- The plaintiff is normally willing to spend resources on the modelling process, since any return on the claim is the equivalent of direct profit to the organisation. Any resource problems are more likely to be linked to the time available for the modelling process. In such a case the modeller will need to make a judgement on whether there has been adequate time to take the modelling process sufficiently far to enable a model to be constructed that will mean that the plaintiff's case is in a stronger position than not having the model at all. This type of judgement is subjective and based upon the experience of the modeller.

In light of these conclusions, there is certainly no evidence to suggest that SD *should not* be used to model D&D for litigation.

The explorations in this chapter have progressed the investigations into whether or not SD can be used to model D&D in a litigation environment. Additional criteria have been identified which will be used to form part of a revised set of criteria detailed at the end of chapter 5.

As mentioned in section 3.1, three different views on potential criteria for the use of SD as a modelling process were to be explored in this thesis. Therefore, to conclude the explorations of these criteria, the third set, which takes on more of a multi-methodological point of view than the previous two sets of criteria, will be discussed in the next chapter.

CHAPTER 5: EXPLORING CRITERIA DETAILED BY FLOOD AND JACKSON IN THE SD LITERATURE

5.1 Introduction

In chapter 3 it was stated that the SD literature would be used to explore what criteria, if any, have been defined as those that should be used to identify whether or not a situation can be modelled using SD. It was commented that this would enable an exploration into the types of situations the modelling approach was intended for and would enable an exploration of the issues involved when using it to analyse D&D for litigation. Also, it was stated that data gathered from the author's involvement in the modelling of D&D for litigation cases would be used to aid the explorations.

Three different sets of criteria were taken from the literature. The first, put forward by Forrester in 1968, was explored in chapter 3. The second, laid down by Coyle in 1977, was discussed in chapter 4. This chapter will now explore the third and final set of criteria which can be inferred from Flood and Jackson's (1991) 'system of systems methodologies'. This set of criteria has been chosen as a contrast to Forrester and Coyle's criteria as Flood and Jackson adopt a more multi-methodological approach when discussing their criteria.

5.2 Flood and Jackson's Criteria

Although the criteria in this chapter will be taken from Flood and Jackson's (1991) 'system of systems methodologies', it should be noted that the idea for such a 'system' was first discussed by Jackson and Keys (1984). The idea behind the 'system of systems methodologies' is that the appropriateness of using a particular systems approach to model a problem can be assessed through the problem context being explored. SD is viewed as only one type of systems approach. Flood and Jackson's 'system' aims to aid a modeller in distinguishing how to choose between the available systems approaches.

One of the reasons for exploring Flood and Jackson's set of criteria is due to the multi-methodological view it adopts. This is a very different view from that taken in the sets of criteria discussed in chapters 3 and 4. It could be argued that Flood and Jackson's criteria have an additional credibility compared to Forrester and Coyle's criteria when testing the validity of the SD modelling approach. Both Forrester and Coyle have devoted the majority of their working lives to the use of SD rather than any other modelling approach. Flood and Jackson, on the other hand, are not committed to SD in the same way and their views may therefore be less biased towards the use of the modelling approach. However, an alternative argument could be that with their commitment to SD, Forrester and Coyle may have more experience in using the SD modelling approach and thus be better informed of when it should, and should not, be used. Whatever the reader's opinion is on these discussions, the fact that Flood and Jackson take a different stance to Forrester and Coyle, makes their criteria worthy of exploration.

When choosing between modelling approaches, Flood and Jackson explore two dimensions:

(i) the relative complexity of the system
- this can be categorised as being either complex or simple (definitions for both of these are discussed in section 5.4 below).

(ii) the relationship between the participants who stand to gain (or lose) from the modelling process
- this is classified as unitary, pluralist or coercive (explored in full in section 5.3 below).

Each of the three participants' categories can be combined with either of the two systems classifications. This yields six possible *ideal type* categories with which to define a problem context.

Although the above categories deal with a relatively over-simplified structure, they can still prove useful. Section 3.2 discussed the lack of published work in providing detailed criteria with which to determine the suitability of SD as a modelling approach to explore a particular situation. Therefore, any substantiated view considering the suitability of a problem to be modelled by a particular approach has the potential of playing a significant role in this thesis. Jackson and Keys (1984) note that the practical implications of this approach are that, if it is to be taken seriously, it could provide a means for problem solvers to determine what systems approach is appropriate to each specific problem context. This is in preference to problem solvers simply going ahead and applying their preferred methodology to each circumstance. Lane (1994) warns against this by stating that "attaching oneself to a single method only, especially if the underlying assumptions of the method are not clear and apparent, is a dangerous enterprise" (p119). Lane goes on to quote Burrell and Morgan (1979, p168) who state that "the selection of a particular type of analogy to represent a system in advance of a detailed analysis of its structure and mode of operation is akin to prescription in advance of diagnosis".

In chapters 3 and 4, Forrester's and Coyle's criteria were used to explore whether or not SD is a suitable approach to model D&D for litigation. Both processes involved the consideration of a set of criteria that could be used to determine whether or not the modeller believed SD to be a suitable modelling approach to be used for the situation. However, Flood and Jackson's criteria create a different process. In their case there is a *choice* of modelling approaches and the investigations aim to discover which group of approaches is the *most* suitable for a situation. The differing approaches are compared against one another with a group of approaches being chosen as the most suitable *relative* to the other groups of possible approaches.

A possible conclusion that can be formed when considering Forrester's and Coyle's criteria is that SD is *not* an appropriate modelling approach for the situation being explored. If this were the case, then the modeller would be no better informed of what modelling approach would be suitable. On the other hand, Flood and Jackson's criteria would give the modeller an indication of the type of modelling approach to use. However, the ambiguity that exists with Flood and Jackson's criteria is that the conclusions only lead towards a potential *group* of approaches, with SD being one possibility, rather than being SD specific as is the case of Forrester's and Coyle's criteria.

Based on the 'system of systems methodologies' approach, Flood and Jackson have concluded that the assumptions underpinning SD reflect a *unitary-simple* problem context. Lane (1999, p505) criticises this labelling of SD by suggesting that "these impoverished views would appear to persist in seeing SD as, "an attempt to apply the ideas of control engineering to socio-economic problems" (Keys, 1988, p8), ignoring or being unaware of the importance to the personal experience of model building and the associated process of experiential learning". However, Lane does not attempt to provide an alternative label for SD using Flood and Jackson's terminology, instead he explores the social theories implicit in SD practice. The possibility that Flood and Jackson are in effect 'type-casting' SD models should be considered. When considering Lane's comments in the light of the modelling of D&D for litigation, it should be noted that the models discussed in this thesis are not used as the basis of future decisions. The type of buy-in from managers is therefore different to that required when using SD to carry out, for example, policy analysis. Also, the audience for the model goes beyond the managers in the plaintiff's organisation. For example, convincing the defendant that the SD model adequately represents reality is an important step in the process. However, the defendant will not be part of the modelling process, therefore "the personal experience of model building" is not possible with all the model's audiences when modelling D&D in a litigation environment.

Further comments on Flood and Jackson's approach are made by Dash (1994, p 96). Dash's opinion on Flood and Jackson's approach is that it is "a very constructive suggestion to the issue of applicability of SD". However, Dash (1994) goes on to comment that their view of SD "appears to hold more strongly for the *system simulation* phase of SD... With respect to the *system representation* phase, i.e., QSD (Qualitative System Dynamics), the literature suggests that nonunitary situations have also been addressed... This implies that the quest for identifying the area of the most effective deployment of SD will remain elusive until some degree of stability is achieved in the theoretical foundations of the discipline itself".

Taking the different comments on Flood and Jackson's approach into account, the fact still remains that in searching through the literature they are one of the few authors who actually make an attempt at distinguishing between the suitability of a problem to be modelled by a particular modelling approach. Their work therefore plays an important role in this thesis.

In section 5.1 it was stated that data gathered from the author's involvement in the modelling of D&D for litigation cases would be used to explore the criteria in an attempt to aid the explorations of the issues involved in using SD to analyse D&D for litigation. This chapter will therefore consider Flood and Jackson's criteria with respect to the data gathered as part of the Castle and Pirate projects. When considering this data, if it can be concluded that these projects have a *unitary-simple* problem context, then this would help in gathering evidence to suggest that SD *might* be a suitable modelling approach for the explorations of D&D for litigation.

5.3 The Participants Dimension

Flood and Jackson (1991) define the participants as those individuals or parties who stand to gain (or lose) from a systems intervention (p33). Following this definition, individuals on both sides of the litigation process should be included as participants. However, a litigation process differs from other forms of intervention when considering those that should be considered as participants. As the plaintiff has initiated the modelling process, the defendants would intentionally not be included as participants in the process (assuming the purpose of the modelling process is not for negotiation between the two sides of the litigation case). The participants would only include individuals on the side of the plaintiff. When modelling D&D in a litigation environment, a claim team represents the plaintiff's organisation. In the projects that the author has been involved in, the individuals that have been a part of the claim team have been lawyers, a claim manager, a senior manager from the organisation's corporate office and the head of the group where the project was based. Other members of the plaintiff's organisation who have also been involved in the litigation process have been managers from the project who act as witnesses.

The unitary, pluralist and coercive relationship between the participants are summarised in table 5.1 (p34).

When making a selection between Flood and Jackson's three participants' categories to represent the participants of a D&D litigation claim, unitary stands out as the most *likely* category. This is mainly due to the fact that the participants are all from the same organisation and would all be seeking the same outcome of the modelling process, that is the plaintiff succeeding in gaining compensation from the defendant for cost overruns relating to the project under investigation. This would hopefully mean that there is a lack of conflict between the decision-makers, therefore resulting in cohesiveness. This also occurs since members of the team are *chosen* by senior management from the plaintiff organisation. Team members who may potentially cause conflict in the team are therefore unlikely to be chosen. Of course, the conclusion would be very different if all those individuals or parties who stand to gain (or lose) from the systems intervention were included as participants, for example including the defendants as participants.

	Unitary	Pluralist	Coercive
1	Share common interests	Have a basic compatibility of interest	Do not share common interests
2	Hold values and beliefs which are highly compatible	Values and beliefs diverge to some extent	Values and beliefs are likely to conflict
3	Largely agree upon ends and means	Do not necessarily agree upon ends and means, but compromise is possible	Do not agree upon ends and means and "genuine" compromise is not possible
4	All participate in the decision making	All participate in the decision making	Some coerce others to accept decisions
5	Act in accordance with agreed objectives	Act in accordance with agreed objectives	No agreement over objectives is possible given present systemic arrangements

Table 5.1 – Characteristics of the "ideal type" Participants Dimension – Taken from Flood and Jackson 1991, p34

In order to explore whether or not the participants of the Castle or Pirate project could be labelled as unitary, the following will consider each of the criteria shown in table 5.1 with respect to the two projects.

5.3.1 Interests

The participants can be seen to share some common interests. Since the participants are only taken to be from the plaintiff's claim team, then each of them will be interested in presenting an analysis of the project which results in persuading the defendants that some of the overruns resulted from actions initiated by themselves. They are all also likely to want to maximise the level of compensation.

An area where the attitudes to the objective of the process may differ lies with the amount of compensation to claim. To illustrate this we can consider the objectives of two of the participants in the Castle project:

The President of the Group – The group's accounts bore the loss incurred due to the project being disrupted and delayed. Therefore, it was the President of the group that had to defend this loss to the organisation's corporate office. In an attempt to improve the position of the group, the President wanted the claim to request maximum monetary compensation.

Senior Manager from Corporate Office – The interests of the senior manager lay with the organisation as a whole. The manager believed that it was important for the plaintiff's organisation to do business with the defendant's organisation in the future. For this reason, the senior manager was attempting to strike a deal with the defendant which was partly monetary compensation plus a deal on future work. The senior manager believed that this was in the best interest of the organisation.

This example illustrates individual participant's specific interests. However, whatever the individuals would prefer as a result of the claim, each of the interests are still in line with the overall common interest of gaining compensation from the defendant.

Of course, the conclusion under this section would be very different if all the parties who stood to gain (or lose) from the modelling process were included as participants. This would include any parties who are involved in the litigation process and would therefore cover all the potential audiences for the model. These are the defendant's claim team, expert modellers, a judge and potentially a jury. Each of these audiences have very differing objectives. These objectives will be discussed in chapter 6. If all these parties were to be included as participants, then the conclusions would need to be that the participants do not share common interests. This is particularly emphasised when considering the plaintiff's and the defendant's objectives.

5.3.2 Values and Beliefs

Through their participation in the claim team, the participants of the team will normally all believe that the defendant's organisation was to some extent at fault during the project and that the plaintiff's organisation should be compensated. Although this was the case in both the Castle and Pirate projects, difficulties did arise with witnesses that were being used from the plaintiff's organisation.

Pirate Project - In the Pirate project, one of the arguments made by the plaintiff was that the defendant had been involved in previous work that would have given the organisation sufficient knowledge at the beginning of the contract to know that the plaintiff had under bid the contract. In gathering statements from participants of the project, the lawyers found that some potential witnesses were not in agreement with the blame being placed on the defendant. Instead, their belief was that the blame lay with the plaintiff's estimators who constructed the bid before the project commenced. Ensuring compatibility in statements and beliefs is obviously an important part of the litigation process for an organisation. Therefore any difference in beliefs needs to be carefully managed by a claim team.

Castle Project - In the Castle project, when the plaintiff's claim team began interviewing those who were a part of the project, they found that some of the participants were inclined to blame themselves for some of the project overrun. This situation was explored in section 4.8.2 when discussing the reliability of data. It was noted in this section that there were so many changes occurring on the Castle project that staff felt that they had almost lost control of the project. No matter what they did, large amounts of rework simply kept on occurring. Managers with many years of experience had lost confidence in their work and were beginning to believe that they were to blame. However, once the chain of arguments describing the occurrence of D&D were explained to them and they could see plausible reasons for the outcome of the project, their morale was boosted. This meant that they were able to consider the possibility that the defendant was at fault and make a judgement on whether or not they agreed with the explanations. Of course a potential outcome of this is that witnesses may only show a belief that the defendants is to blame in an attempt to free themselves from blame.

If differences in participant's values or beliefs do arise, then the plaintiff's claim team need to consider the causes of these differences and whether each of the differences need further research or whether the differences are due to a lack of appreciation of the system by the participant.

5.3.3 Agreement upon Ends and Means

The agreement upon "ends" comes from the shared objectives described in criterion 1. The agreement on "means" can be discussed at different levels of detail. The participants are in agreement that litigation is the procedure that should be used to enable them to meet their objectives. This shows agreement of the process that should be undertaken. However, individuals may have their own thoughts and ideas on how that litigation process should be carried out.

Examples of the type of "means" the claim team may have discussions on are as follows:

- Whether to attempt negotiation or mediation with the defendant or to focus on taking the defendant to court. The senior manager from corporate office will have an opinion on the direction that the organisation's corporate office wants the claim process to proceed, however the lawyers will also provide advice in the area.
- Whether to use particular documents or witnesses as part of their evidence. The claim manager usually has the most knowledge about the data and witnesses. Therefore, the claim team is likely to place particular importance on his opinions on how to deal with this.

These examples illustrate that each member is a part of the team because they have a particular input to make to the progression of the claim. Therefore in many cases, when a decision has to be made on how to move forward, the member of the team with the most knowledge in the area of discussion is likely to direct the way forward. Of course, this does not mean that there are no differences of opinions when, for example, decisions overlap with different individual's knowledge. However, the reduced number of members of the plaintiff team does help in reducing the number of differences that occur. Also, the common objectives of the team mean that it is reasonable to assume that the participants *largely* agree upon "ends" and "means".

5.3.4 Decision-Making

When considering the process of decision-making during the litigation process, the managers used as witnesses on the project are generally not involved. However, each of the other participants are involved in this process. If the power of each of the individuals is considered, the senior manager from corporate office has been the individual who has made the final decisions and therefore the opinions of each of the individuals in the claim team are not equally weighted when decisions are made.

Flood and Jackson's criteria distinguish between the participants (i) all participating in the decision making and (ii) some coercing others to accept decisions. Although not explicitly mentioned, Flood and Jackson's criteria appear to be focussed more towards the consideration of the opinions of different organisations or different divisions within an organisation. This is in contrast to the opinions of individual people within the same organisation or division. That is, Flood and Jackson are attempting to capture the differing opinions of cohorts of people rather than a consideration of individuals' personal opinions. If this is true, then there is really only one cohort of people included in the participants – the plaintiff's organisation and therefore it could be concluded that all the participants participate in decision making. However if individuals' opinions are taken into consideration, then evidence from both the Castle and Pirate projects would conclude that the corporate office representative has the overriding power when decisions are made.

5.3.5 Act in Accordance with Objectives

In each of the projects discussed in this thesis the objectives of the litigation process were agreed upon by those in the plaintiff's claim team. Once litigation was chosen as the process to follow, those who were a part of the plaintiff's claim team each worked towards the objectives that were agreed upon.

When discussing the beliefs of participants, the managers who were witnesses from the project were highlighted as a group of people where differences had existed between their views and the views of the plaintiff's claim team. For example, in the Pirate project it was mentioned that some potential witnesses were unhappy with the blame being placed on the defendants for the under-bidding of the project. They believed that the plaintiff should take some of the blame. Acting in accordance with the objectives of the Pirate project means that participants should support all the arguments for claiming compensation from the defendant. However, some of the witnesses were not happy in doing so for this particular claim. The plaintiff's claim team therefore needed to be aware of which witnesses would act in this manner and fully understand the arguments for them acting this way. Indeed, in this case, the claim team agreed with the witnesses' argument that the estimators on the project under-bid the contract. However, in support of their claim, they were arguing that this was due to inexperience in the type of work covered in the project. On the other hand, the defendant did have experience of this type of work and they did not share that experience with the plaintiff. Due to their previous experience they should not have accepted the estimate that the plaintiff submitted. However, some of the witnesses did not agree with this argument and still believed that the blame lay with the plaintiff.

Although the plaintiff's claim team appear to have acted in accordance with the objectives in the Pirate project, some of the witnesses, who may be perceived as participants in the process, have not acted in accordance with all of the objectives. However, the plaintiff's claim team was aware of this situation and had planned ways of dealing with the differences.

5.3.6 Summary of the Participants Dimension

In considering each of the criteria discussed in this section, it is unlikely that any of the responses would be a definite 'yes' or 'no'. Instead, judgement needs to be made about the scale of unitary, pluralist or coercive characteristics that the participants display. However, in the cases of modelling D&D for the Castle and Pirate projects, it appears that a *unitary* label could be placed on the participants of the process. However, it should be noted that this is mainly due to the omission of many of the parties who stand to gain (or lose) from the process. Also, members of the claim team are chosen by senior management from the plaintiff organisation. Therefore, members who may potentially cause conflict in the team are unlikely to form part of the team.

Those participants who are a part of the plaintiff's claim team certainly appear to act as unitary participants. The group of people that have had the largest divergence from this are managers acting as witnesses. Their beliefs and understanding of the events of the project need to be considered carefully to ensure that these beliefs and the plaintiff's claim team's beliefs are in line with one another. If they are not, then the plaintiff's claim team need to find the best way of incorporating the witnesses beliefs into their case. This is important since in a litigation process, cohesiveness between the participants in the plaintiff organisation is desirable to enable their case to have the best possible chance of success. A unitary relationship is therefore important for a successful litigation process.

5.4 The Systems Dimension

When we turn to the classification of the system, it is difficult to reach a conclusion for some of the criteria. This means that reaching a classification on whether or not the system is simple or complex is not straightforward. Table 5.2 compares the characteristics of the two classifications.

	Simple	Complex
1	There are a small number of variables	There are a large number of variables
2	There are few interactions between the elements	There are many interactions between the elements
3	Attributes of the elements are predetermined	Attributes of the elements are not predetermined
4	Any interaction between elements is highly organised	Any interaction between elements is loosely organised
5	Well-defined laws govern behaviour	They are probabilistic in their behaviour
6	The "system" does not evolve over time	The "system" evolves over time
7	"Sub-systems" do not pursue their own goals	"Sub-systems" are purposeful and generate their own goals
8	The "system" is unaffected by behavioural influences	The "system" is subject to behavioural influences
9	The "system" is largely closed to the environment	The "system" is largely open to the environment

Table 5.2 – Characteristics of the "ideal type" Systems Dimension - Taken from Flood and Jackson 1991, p33

Flood and Carson (1992, p20) state that complexity is associated with people and things. This arises from the definition that "systems are situations as perceived by people". Also, Jackson and Keys (1984, p475) state that the conclusion on whether or not a system is simple or complex depends upon the *purpose* the observer of the system has for

considering it. Other authors also agree with this statement (for example Beer 1979, Espejo 1987). These *subjective* elements are not explicitly reflected in Flood and Jackson's criteria displayed in table 5.2. The perceptions of the observers of the system have to be considered and thus the answers to the nine criteria could differ depending upon the people who are observing the system. Beer (1979) identifies with the difficulty of defining measures of complexity. He states that "...complexity is the result of the way that systems behave and interact. And if systems are subjective phenomena, then we are going to have trouble in determining a measure. The whole idea of measures is to be objective..."(p32).
The observers of the system are taken to be the individuals who are involved in the modelling process. The observers therefore include the modellers and some of the plaintiff's claim team. It should be noted that each of the plaintiff's claim team have been involved in the modelling process to varying degrees. For example, the claim manager has been involved to the greatest extent with the model, whereas other members of the claim team may only see the model a couple of times throughout the modelling process. Although the modellers are external to the organisation, they use information gathered from the participants to formulate their own perceptions of the system.

Each of Flood and Jackson's nine criteria will now be explored in turn.

5.4.1 Number of Variables

Keys (1988, p69) suggests that "a suitable measure of complexity might be one which indicates the amount of understanding which can be gained from the models of the problem-situation ... influence diagrams and digraphs have the advantage of rigour, allow for a degree of quantification in the measures and are also easily accessible by non-specialists". In the case of the Castle project a cause map was constructed to gain an understanding of the problem-situation. This cause map may therefore be used to gain an indication of the number of variables that exist in the system.

Eden et al (1992) discuss measures that can be used to indicate the complexity of a cause map. They suggest that "...analyses that depend upon the number of nodes should be treated with great care" (p312) since this could be dependant upon such factors as the experience of the interviewer and the length of the interview. This may suggest that a modeller should be careful when basing a judgement of the complexity of a system on the number of variables in a representation of a system. However, in the case of Flood and

Jackson's criteria, it is only one of nine measures used and therefore a judgement is based upon a portfolio of measures, as suggested by Eden et al (1992).

If the cause map of the system is used to aid an exploration of this criterion, then it could be suggested that a count of the concepts in the map provides a guide to the number of variables contained in the system. A count of the variables in the Castle cause map tells us that there are 127 variables. It could be argued that the arrows representing relationships within the map are also variables as they provide information beyond that represented by the concepts. In this case the number of variables would then total 296. However, no matter which of these measures of the number of variables is taken, the criterion is not particularly useful as no guide is given to what number of variables is required to be able to conclude that a system has a *small* or *large* number. The observers of the system may individually perceive the system to involve a large number of variables and thus believe it to be complex. However, what one person may consider being a large number of variables, another person may not. Therefore, no explicit conclusions can be made for this criterion.

5.4.2 Number of Interactions

This criterion is again unhelpful since there is no indication given as to what is taken to represent *few* interactions and what represents *many* interactions. Using the cause map constructed for the Castle project, the *number of links* given in this map can be taken to represent the degree of interaction between the variables of the system. In this map there are 169 links. However, although this number provides us with a measure of the number of interactions in the system, the lack of detail on what can be concluded to be *few* and *many* interactions prevents any conclusions from being reached.

Rather than simply considering the number of links in the cause map, the number of feedback loops may be another way of measuring complexity. The greater the number of feedback loops, the more dynamic the system is and therefore the more complex the interactions between the variables of the system. This is in line with Senge's (1990, p71) description of *dynamic complexity* rather than *detail complexity* which only considers the number of variables in a system.

Five hundred and thirteen loops were found to exist in the cause map for the Castle project. The author would consider this to be a large number of loops leading to a complex situation. However, although this is the conclusion made by the author, there is no clear evidence on the conclusion that Flood and Jackson would reach based on their criteria.

5.4.3 Attributes of the Elements

In seeking further explanation about Flood and Jackson's meaning of whether or not the attributes of the elements are *predetermined*, Jackson and Keys (1984, p475) explain that for complex systems "not all the attributes of the parts of the system will be directly observable. As a result it is difficult to understand the nature of the system completely". Vemuri (1978, p2), from whom Jackson and Keys take their definitions of a complex system, comments that if the attributes are not necessarily observable, then this is normally because "the structure, or configuration of the system is ... (not) self-evident". However, this does not lead to a full appreciation of how an assessment of this criterion can be made.

When the modelling team initially got involved with the Castle project, they were given a copy of a draft claim document. This document summarised the plaintiff's view of the system. From this document the modellers were able to gain a picture of how the system was structured and thereby constructed a cause map. This illustrates that the plaintiff was able to highlight the important elements of the project that played a part in the overrun of the project. Some of the interactions between elements of the system were included in the claim document. However, others had to be drawn from later interviews with those people who were involved in the project. Some attributes of the elements of the system were also included in this claim document. For example, the claim document included data on the lengths of delays in the defendant responding to questions arising from production. It also included the number of changes requested by the defendant and the hours of engineering and production work estimated for each of these changes.

Since the modellers were constructing a SD model that represented the D&D that occurred on the Castle project, the modellers had to gather a lot of data to enable the model to be populated. The need for quantification meant that the modellers had to seek information about the attributes of various elements of the system. The data for some of the attributes was straightforward to request and gather, whereas others proved to be more difficult. This is captured in section 7.3.2 where parameters of the SD model are classified by how easy or difficult it is to (i) define the parameter value being requested and (ii) obtain data on the numerical values of the parameter. Although many of the parameters were very difficult to define and gain data for, none were impossible. However, the question being considered in this section is would the observers of the system define all the attributes of the elements as "directly observable"?

In the Castle project, the difficulties associated with gathering data regarding the effect of cross-impact were discussed in section 4.8.1. This section concluded that the difficulties were due to:

- The modellers finding it difficult to communicate exactly what they required from the plaintiff's analyst. Data on cross-impact was not directly available, therefore the modellers had to attempt to explain what they required from the plaintiff's analyst so that he could help in obtaining data that may be relevant to the quantification of cross-impact.
- The lack of availability of engineering personnel that had worked on the project. This caused difficulties in the interpretation of any of the engineering data that was available.

Due to the method that was used to record data and the delays that were a part of cross-impact (e.g. delays between a change in a design and the ramifications of the change), observing all the implications of cross-impact was not a straightforward process. Although it was agreed that cross-impact occurred in the system, no observer directly observed its occurrence. Returning to Jackson and Keys' (1984) definition of the criterion discussed in this section, the above indicates that not all attributes of parts of the system are *directly* observable.

In returning to the original characteristic, Flood and Jackson (1991, p33) asked whether or not the attributes of the elements of the system were or were not *predetermined*. Without any further explanation from them, in the case of modelling D&D in large engineering projects, it is assumed that Flood and Jackson are asking whether or not the attributes of the elements were set before the project began, or if they occurred unexpectedly during the project.

When modelling D&D for litigation, one of the model outputs is to simulate what would have occurred if the project had run as expected in the contract. This model run will therefore consist of elements of the system that are predetermined by the contract. When modelling the actual outcome of the project, the only difference from the budget run is that the unexpected D&D triggers are included. Therefore, by their definition, they were not expected before the project began. These triggers may be completely unexpected elements; for example an underestimation in the number of hours required to complete the project. Otherwise the triggers may be elements whose attributes were not as expected; for example an excessive number of changes in engineering, or the occurrence of excessive delays by the defendant in response to queries. Therefore, due to the nature of D&D triggers, the attributes for some elements of the systems explored in this thesis were not predetermined.

5.4.4 Interaction between Elements

Schoderbek et al (1980, p79) describe this criterion as being due to "the existence or lack of predetermined rules and regulations which guide the interactions of the elements and/or specify the attributes of the system's elements". This therefore ties in with the previous criterion, which concluded that the attributes of some elements in the projects being explored in this thesis are not predetermined. However, rather than focussing on the *attributes* of the elements in the system, this criterion considers the *interactions* of the elements of the system.

When discussing the last criterion, the triggers of D&D were highlighted as elements of the system that occurred unexpectedly during the project. Two different types of D&D triggers were noted; unexpected elements and unexpected attributes of elements.

For the elements whose attributes were unexpected, the relationships between the elements were generally as expected before the project began. However, a situation where this may not be the case is where a *portfolio effect* occurs. This occurs if a number of D&D events lead to an outcome that is not the effect of any individual D&D trigger, but only occurs because the specific portfolio of D&D triggers occurred together. For example, in a project that the author was associated with, a number of different D&D triggers each individually caused a delay to the project. However, when they all occurred together, this total delay was so extensive that the project extended into the winter months. For this particular project, working during the winter was not easy. This either caused huge delays in activities whilst the project waited for spring to arrive, or activities took a far longer period of time to complete. No individual D&D trigger caused the winter working, it was the portfolio of events that was identified as the cause. This means that unexpected interactions can be created from a portfolio of D&D triggers which are seen to have unexpected attributes.

For elements that are themselves unexpected, they too can be a part of a portfolio effect and hence can have unexpected interactions. However, when considering the elements individually, the interactions between these unexpected elements and the other elements of the system were expected. To illustrate this, the elements that fall under this description in the Castle and Pirate projects are as follows:

Castle Project

- The defendant told the plaintiff to use work-arounds throughout the engineering and production processes. The use of work-arounds may have been unexpected, but their outcome was not unexpected. The outcome of work-arounds was some advancement in work, but also additional rework from incorrect judgements made by those people who were advancing the work.

Pirate Project

- Underestimation of the scope of work when estimating the number of hours required to complete the project. Again, the outcome of this event was not unexpected. There were more hours of work required to do the project and hence an immediately compressed project schedule.

- The unpredictable nature of the project meant that the management of the project was unable to gain a normal level of reduction in the time to produce the first product through planning actions carried out before the commencement of the project. Events may have been unexpected, but their outcomes were not. This, again, created more hours of required work on the project and hence an immediately compressed project schedule.

The discussion in this section therefore concludes that interactions between elements in the system are mainly expected and hence highly organised. However, the exception to this is the occurrence of a portfolio effect, where a number of events can cause an unexpected outcome. This means that these interactions are unexpected and hence loosely organised.

5.4.5 Laws Governing Behaviour

In the Castle and Pirate projects the elements governing the behaviour of the system are actions from managerial decisions and exogenous triggers into the system. This conclusion comes from the discussion in section 3.4 that 90% of the feedback loops identified in the cause map for the Castle project and 100% of the feedback loops identified in the cause map for of the Pirate project were produced by decisions taken on the project. Since managers do not normally make decisions in a random manner, but normally plan them with each decision supporting a set of objectives, this supports an argument that defined laws govern the behaviour of the system. However, it should be noted that during the projects discussed in this thesis it proved difficult to fully define these laws since they proved to be very complex.

Although there is a structure to decision-making that is not probabilistic by nature, there will always be some random noise around managerial decisions. This means that, although decisions are generally made in line with specific objectives, random events may cause decisions to differ slightly from expectations. Also, management may use intuition as well as particular decision making rules, otherwise an automated decision-making process could replace the decision-makers. This means that actual decisions may further vary from decisions based on decision rules.

In extreme cases, some managers have reported that when a project is very disrupted and D&D is rampant, the resulting complexity is akin to chaos (Eden et al 2000, p296). When faced with making decisions in such a chaotic situation, managers may become very uncertain of what strategies to follow. There are great difficulties in estimating labour requirements since the manager is unable to predict what is in store for the future of the project. This may cause managers to either base decisions on incorrect measures, or they may base their decisions on intuition. Decisions may then become more based upon gut-feeling rather than well-defined measures. These types of decisions may appear more probabilistic in nature. However, they are not truly probabilistic since they are based upon the manager's experience of managing projects. However, it would be virtually impossible to model this form of decision-making through a defined structure.

In conclusion, unless the decision-makers believed the project to be in such chaos that they abandoned defined decision making rules, the laws governing the system behaviour in a managed project are more likely to be relatively well-defined rather than probabilistic.

5.4.6 System Evolution

Jackson and Keys (1984, p475) explain that system evolution is largely due to "the fact that such systems are in constant interaction with the environment – they are 'open' rather than 'closed'." Thus, this criterion can be considered together with the ninth criterion; *Interaction with Environment.* The consideration of whether the system is 'open' or closed' partly depends upon the decision of what the system boundary should encompass. It could be the case that the system boundary could be expanded to include interactions with the environment that would mean a conclusion of an 'open' system could be altered to a 'closed' system. This is based on the assumption that it is *possible* to model all of the system that would need to be modelled to consider the system 'closed'.

In the cases of the Castle and Pirate projects, all the variables contained within the cause maps were agreed between the modellers and the plaintiff's claim team as the variables that contributed towards D&D. The only interactions the systems have with the environment are the triggers from the environment that cause inputs to the system being modelled.

Of course, the conclusion that the systems are 'closed' is based on information gathered from the plaintiff's organisation and does not include the opinions of those involved from the defendant's organisation. For example, in the Castle project waiting on information that was required from the defendant led to delays in the plaintiff's progress. However, it may be the case that the defendant would argue that they were delayed due to disruptions from the plaintiff. Alternatively, when the defendant made changes to drawings, they may argue that this was caused by actions taken by the plaintiff. Both of these circumstances mean that a part of the system being modelled is affecting an element of the system that has been modelled as part of the environment. The conclusion of whether or not the system being modelled is 'open' or 'closed' therefore depends upon the observers' perceptions of the system. For the plaintiffs of the Castle and Pirate projects, the system being modelled is 'closed'. Also, based upon the information available to the modellers, they believed the systems being modelled to be 'closed'. As illustrated above, the modellers opinion on this may change if they were exposed to the defendant's perceptions of the system. However, if this were to occur it may be possible for the modellers to expand the system boundary to include elements of the environment to enable the system to become 'closed' again. However, in a litigation environment it is unlikely that the defendant would be included in the system. They are normally defined as an element of the environment since the objective of the model is to support the plaintiff's claim for compensation.

5.4.7 Sub-systems

All systems can be seen to have sub-systems and, indeed, each system can be treated as a sub-system of a wider system. This criterion considers whether or not any of the sub-systems pursue their own goals. If any of the sub-systems do pursue their own goals, the distinction that has to be made is whether or not the sub-systems' goals are in line with the overall system goals. Complexities arise when sub-systems' goals are in competition with each other or with the overall system's goals.

In considering both the Castle and Pirate projects, sub-systems pursing their own goals can be seen to occur in both these projects. The goals of the workforce on the projects are one area where this occurred.

Workforce Goals

The goals for most engineering projects are to produce a product that meets all the specifications set by the contract and to complete the work on time and in budget (Turner 1993, Webster 1993, Lock 1994). The labour force is a sub-system of the overall project and management would normally plan and manage labour requirements in an attempt to meet the overall project goals.

In the Pirate project, retaining production staff proved very difficult. This was due to the location of the project. It took place in an area where there was a lack of staff due to competition from other organisations when hiring labour. The location also meant that it was very difficult to attract labour from other parts of the country. However, if the company was successful in attracting labour from other parts of the country, they did not retain them for long periods of time. Once employment was available in other locations, labour would leave the company. For these reasons, the labour turnover for the project was far higher than would normally be expected. The workforce's goal was to seek the 'best' employment, where 'best' to them meant taking location and salary into account. However, this goal led to high turnover rates on the project, this meant that learning losses ensued. Losses in the learning gained by workers meant that work took longer than expected and hence slowed down the overall progress of the project. This hindered the project from being completed on time.

High labour turnover was also seen to occur in another area of the Pirate project. Aircraft tails were being manufactured by a separate division of the plaintiff organisation. In managing the manufacture of the aircraft tails, management of this division did not give the project high priority. The goals of the management for this division was to give the most attention to those projects that were of most importance to it. This meant that the 'best' workers were transferred onto the more important projects. As time progressed, more and more workers were transferred from the plaintiff's project onto other projects. This resulted in a reduction in the quality of tails that were manufactured and delays in the supply of the tails. The reduction in quality of tails meant an increase in the rework required on the tails. This caused even more delays in the project. The seeking of different goals by the division that manufactured the tails meant that there were even more difficulties for the plaintiff in attempting to meet their planned schedule.

Labour was also seen to strive towards separate goals in the Castle project. Other projects that were given higher priority in the organisation meant that management changed a number of times on the project. Management would see the change as a move in the right direction for their careers and hence helped to achieve their goals as individuals. However, different management on a project means different management styles and changing methods of decision-making. This discontinuity can cause disruptions to a project.

The time and budget overruns on the Castle project meant a lot of pressure being placed on workers to attempt to catch up the schedule. Workers from other departments were transferred onto the project at certain times to help deadlines to be met. However, due to the pressures felt on the project, these workers did not normally wish to work on the project for a second time. Therefore, each transfer of workers meant that a large proportion had not previously worked on the project. This meant that there was a limited benefit gained from the previous learning by the workers. In this case, labour was striving towards minimising the pressure they felt as individuals. However, this was in competition with the system's goals of benefiting from the learning gained by workers and hence reducing the time spent to do work and so attempting to meet deadlines.

The above examples show that sub-systems have pursued their own goals in the Castle and Pirate projects and that these goals have been in competition with the overall system's goals. The examples discussed illustrate how workers are systems themselves with their own goals. This is the case in any project. In the Castle and Pirate projects circumstances meant that the goals of the overall project were in competition with the individual workers' goals.

The systems dimension criterion discussed in this section appears to conclude that the systems being discussed have a complex system attribute. However, it is interesting to note that each of the above examples were not modelled in the SD models as sub-systems pursuing their own goals. Instead, each of the high turnover rates were modelled as exogenous variables. Argumentation was used to trace the causes back to an action taken by the defendant. By doing this, the sub-systems were effectively taken as being outside the system boundary. Although this is one method of handling the situation, it may leave the models open to criticism. In using actual turnover rates, the models do not explain that the actual turnover rates are quantitatively reasonable. This argumentation is only carried out through the qualitative models. There could be other causes that have contributed towards the high turnover rates that the defendants can argue were initially caused by the plaintiff.

If a plaintiff's claim team was not happy with this weakness, the modeller may attempt to expand the system boundary to include the sub-systems. However, to do so, the modeller would then need to model the goals of the individual workers. For example, In the Pirate project, this would lead to modelling the attractiveness of other organisations over time. The term *attractiveness* may have many different meanings to different workers, thereby increasing the size and complexity of the model. Whether or not a satisfactory representation could be modelled is a complex issue and one that would need further research. Although an interesting question, it is outside the scope of this thesis.

5.4.8 Behavioural Influences

Jackson and Keys (1984, p476) expand on this criterion by explaining that behavioural problems may arise since decisions are affected by "political, cultural, ethical and similar factors" and thus "makes it difficult for the problem solver to fully understand the 'rationale' behind decisions made by actors in the system". In any organisation it would be difficult to conclude that any decision-making is not in some way affected by such factors as those listed above (Morgan 1997). The only question is *to what extent* they are affected.

If the main driver of individual decisions was taken to be a long-term strategy on how the system was agreed to be managed, then factors such as politics etc would only influence the decisions through random disturbances.

The decision-making that is modelled in the litigation models is decisions regarding the possible acceleration of the project. This is manifested through, for example, the hiring of new workers or getting workers to work overtime. These types of decisions are typical of those that will be modelled when exploring D&D in any large engineering project for litigation.

In the Castle project the plaintiff considered the minimising of the total delay in the project as an overall aim. This was driven by both an organisational way of working as well as being due to the defendant's expectations. Updated schedules were produced on a regular basis (every two weeks at the height of the project) based on individual worker's estimates of the work that was left to do. These schedules provided information on which the program manager could base their decisions. Assuming that the program manager did indeed base his decisions on this information, this illustrates a rational way of making decisions.

However, not all decision-making in the Castle project was as rational as being based on regularly gathered information on the project. An example of this was the decision of the number of methods labour hours to include in the bid for the project. After the methods hours had been estimated, the engineering manager reduced the number of methods hours that was actually included in the estimate. The reason for this is that he perceived a high level of methods hours as an expectation that the engineers were incompetent at their job and hence the methods workers would have to compensate for this by redoing drawings. Since the engineering manager did not wish his engineers to be seen as if they were not going to do a good job, he did not think that the methods workers needed all the hours that they had estimated. This decision was therefore taken based upon friction between different managers involved in producing the estimate for the project, rather than a rational calculation of how many methods workers were actually required. The peculiarity of this particular decision was only highlighted since the graph of estimated labour requirements did not follow a typical shape.

In assessing the extent to which behavioural influences affected the decision-making process that was a part of large engineering projects that overrun in time and money, a general rule cannot be made. Instead each project can only be considered individually. For example, behavioural influences did, to some extent, affect the decision-making processes for the Castle project. However the majority of the decision-making processes that were modelled were based on logical causal effects e.g. the project needed to be accelerated, therefore a decision was made that an additional amount of workforce would be brought to help achieve this.

In any large projects that are undertaken by large organisations, behavioural influences will exist. Decision-makers are individuals with their own experiences, beliefs and goals. Each of these factors is going to affect the decisions they make. The consideration with this criterion is *to what extent* decisions are affected by these factors. Behavioural influences do appear to have influenced some of the Castle decisions. However, a general statement cannot be made when considering the modelling of D&D in a large engineering project. A conclusion can only be reached for an individual project.

5.4.9 Interaction with Environment

This criterion was covered during the discussion for the sixth criterion; *System Evolution.*

5.4.10 Summary of Systems Dimension

It is not surprising that the criteria put forward by Flood and Jackson to define a simple and complex system do not lead to a definite conclusion. Flood and Jackson (1991, p32) acknowledge that their classification is an "... "ideal-type" classification which draws out some key features of different problem contexts, but which does not expect any "real world" appreciation of a problem context to fit exactly any one box". This means that any system can only be judged on a scale of how close or far from an ideal-type it is.

A further difficulty in using the criteria is the lack of explanation Flood and Jackson give about each of them. This means that the interpretation of the criteria is a loose process that may have an influence on the conclusions that are arrived at.

Bearing the above in mind, the conclusions from each of the criteria are summarised in table 5.3 below.

	Criterion	
1	Number of variables	Unable to reach a clear conclusion on this criterion.
2	Number of interactions	Unable to reach a clear conclusion on this criterion.
3	Attributes of elements	For some of the elements, their attributes are not predetermined. Some of the elements therefore have a characteristic of a ***Complex*** system.
4	Interaction between elements	The majority of the elements have predictable interactions. However, unpredictable interactions can occur through a portfolio effect. This means that a few interactions may display ***Complex*** system characteristics.
5	Laws governing behaviour	The laws governing the system behaviour in a managed project are more likely to be relatively well-defined rather than probabilistic. This is the characteristic of a ***Simple*** system
6	System evolution	Based upon the information available to the modellers, they believed the systems being modelled to be 'closed'. This is the characteristic of a ***Simple*** system.
7	Sub-systems	Sub-systems do pursue their own goals and these goals can be in competition with the overall system goals. This is the characteristic of a ***Complex*** system.
8	Behavioural Influences	Each project needs to be considered individually to determine if they are affected by behavioural influences. As an example, the Castle project was effected by behavioural influences and hence had a characteristic of a ***Complex*** system.
9	Interaction with Environment	Following on from criterion 6, this also has the characteristic of a ***Simple*** system.

Table 5.3 - Conclusions from the Explorations of the Systems Dimension Criteria

Having gone through each of the criteria for the projects discussed in this work, no definite conclusions have been made. No clear conclusions were arrived at for two of the criteria. Although they are less subjective than the other criteria, no guidelines are given on how to objectively distinguish between a simple and complex system. One criterion needs to be assessed on an individual system basis. One criterion tends towards complex system characteristics. Two criteria conclude that a minor number of elements display complex characteristics. The remaining three criteria display simple system characteristics.

It is difficult to come to an overall conclusion from this section, as some of the criteria remain unanswered. Also there is no indication if it is simply a matter of numbers i.e. if more criteria indicate simple, does this mean that the system tends more towards a simple system. If this is so, then excluding the unclear criteria, the system classification appears to lie closer to the simple categorisation with the potential for a few elements to display strong complex characteristics.

5.5 Conclusions from Flood and Jackson's Criteria

Flood and Jackson warn us that the classifications they provide for both the participants and systems dimensions are "ideal-type" classifications and that they should only be used to inform a debate when considering which of the six problem contexts a situation falls into. Based on explorations into the two dimensions the following was concluded:

Participants Dimension - this led to the conclusion that in the cases of modelling D&D for the Castle and Pirate projects, it appears that a *unitary* label could be placed on the participants of the process. However, it should be noted that this is mainly due to the omission of many of the parties who stand to gain (or lose) from the process. It was noted that a unitary relationship is important for a successful litigation process since cohesiveness between the participants in the plaintiff organisation is desirable to enable their case to have the best possible chance of success.

From the explorations of the participants dimension the following issues can be highlighted as areas where cohesiveness between the participants of the plaintiff team may not occur. These are issues that need consideration when modelling D&D for litigation.

- The interests of each individual of the plaintiff's claim team are likely to be in line with the overall common interest of gaining compensation from the defendant. However, individuals may have different opinions about what they would prefer as a detailed outcome of the claim.
- In both the Castle and Pirate projects, difficulties arose with witnesses that were being used from the plaintiff's organisation. The beliefs of some witnesses about the reasons for the failure of the projects were different to other members of the claim team. If this occurs, the plaintiff's claim team needs to consider the causes of these differences. The claim team then needs to consider whether each of the differences need further research or whether they are due to a lack of appreciation of the system by the witnesses.
- Generally, when a decision has to be made on how the plaintiff's claim team should move forward, the member of the team with the most knowledge in the area of discussion is likely to direct the way forward. However, differences of opinions can occur when, for example, decisions overlap with different individuals' knowledge. Of course, the reduced number of members of the plaintiff team does help in reducing the number of differences that occur. However, if we consider the power of each of the individuals in the claim team, in both the Castle and Pirate projects the senior manager from corporate office was the individual who had ultimate responsibility for the final decisions. Therefore, the opinions of each of the individuals in the claim team are not equally weighted when making decisions.

Systems Dimension – explorations were less conclusive than the participants dimension. Evidence from the projects discussed in this thesis suggests that the systems discussed do not lie at either end of the scale. They lie somewhere between a purely simple or purely complex system. However, there is evidence to suggest that the system lies closer to the simple categorisation with the potential for a few elements to display strong complex characteristics.

From the explorations of the systems dimension the following issues can be highlighted as areas where there is potential for increased levels of complexity in the system. These are issues that need consideration when modelling D&D for litigation.

- The triggers of D&D. Complexity arises in the system since the occurrence of the triggers are either unexpected or the way in which they occur are unexpected. Also, a number of D&D triggers may cause a portfolio effect where the triggers lead to an outcome that is not the effect of any individual trigger, but is caused due to the occurrence of the specific portfolio of triggers.
- Sub-systems of the overall system can cause complexity by pursuing their own goals. For example, workers on a project may leave the project to increase their salary or as part of promotion. Although this will be in line with their own goals and ambitions, it may prove detrimental to the goals of the project. This occurs since the project is losing the learning that the individual has gained about the project.
- Behavioural influences will exist in any large project. Decision-makers are individuals with their own experiences, beliefs and goals. Each of these factors will affect the decisions they make. For any individual project, a consideration has to be made as to *the extent* that the decisions are affected by these factors.

The labelling of unitary and possibly tending more towards simple, means that the modelling of D&D for litigation may be possible using SD when considering Flood and Jackson's criteria. Of course it should be noted that if the problem context we are dealing with is not strictly unitary-simple, we do not necessarily rule out the use of SD as suggested by Flood and Jackson. Lane's (1999) earlier quote implied that SD modelling does not need to be restricted completely to this problem-context. Keys (1990, p487-8) also notes that it has been suggested that SD can be used in complex situations. He states that "the recent attempts to make a methodology grounded in hard systems thinking appropriate to situations characterized by complexity and plurality have meant that SD has become increasingly open to criticism from the soft systems perspective." The potential reasons for the capability of SD to handle such complex situations can further be identified when Keys (1990, p487) goes on to comment that "... the nature of the situations being addressed in the early stages of decision-making is typically characterized by high complexity and plurality.... (they) require soft systems thinking if they are to be adequately confronted by analysts... In making this progress (using a qualitative model for the general investigation of the early stages of the decision-making) SD has begun to move across the boundaries of applicability of hard systems thinking which are imposed by complexity and plurality". This infers that a qualitative modelling process used prior to the SD modelling can be used to initially handle the complexity of the problem.

Causal mapping was used as part of the modelling process for the projects discussed in this thesis. A qualitative tool, such as causal mapping, can help reduce the level of complexity by enabling the observers of the system to gain an improved understanding of the situation (Coyle 1977, Wolstenholme 1982, Wolstenholme and Coyle 1983, Wolstenholme 1990, Eden 1994, Coyle 1996). It has also already been noted that complexity depends upon the perception of the observers of the system. An observer's perception of the structure of the system will partly arise from their purpose for observing the system. The purpose therefore influences how much complexity is required to be observed.

The above explorations have considered a number of criteria based upon the views of the observers of the system. By considering observers' views, this process is *subjective.* Jackson (1990, p661) picks up on this point by asking "whether it is possible to arrive at an 'objective' account of the nature of any problem situation. There would seem to be considerable room for disagreement about whether a system should be viewed, say, as systemic-unitary or systemic-pluralist". Conclusions to the problem-context for the system may therefore be contested. The above explorations are based on the views of one set of observers for the systems discussed.

5.6 Lack of Detail in the SD Criteria

In chapter 3, three sets of criteria were chosen from the SD literature in an attempt to explore the appropriateness of using SD to analyse D&D for litigation. These sets of criteria have now been explored and a summary of the conclusions of these explorations can be given. The summary will be given in section 5.7. However, before moving on to the summary, this section will reflect on the criteria that has been taken from the SD literature and used for the explorations.

In the cases of Forrester and Flood and Jackson some of the criteria used are introduced with minimal explanation. For example, Forrester's *Systemic Interrelationship* or Flood and Jackson's *Number of variables in the system.* Assumptions have been made regarding what it is believed the authors mean by the criteria. However, there is no indication, for example, what Flood and Jackson mean by a *large* number of variables. It is disappointing that there is so little in the SD literature regarding criteria to assess the suitability of a situation to be modelled using SD. However, it is even more disappointing that when an attempt has been made at listing such criteria it has been done with a minimal amount of detail. Of course, this may be a reflection of the difficulty involved in arriving at such criteria.

The lack of detail in the criteria discussed can leave the reader with the impression that the set of situations which SD could be used as a modelling approach is very large. As mentioned in section 3.2, it could be the case that virtually any situation *can* be modelled using SD. However, this view cannot be accepted unless *some* form of verification of the view can be made.

The possible reason for the apparent lack of detail in setting criteria to determine a situation's suitability to be modelled by SD may be seen in Yaman Barlas' opening speech in the 1998 International SD conference. Barlas stated that he no longer approaches a problem with a bag of tools and attempts to fit the tool to the problem. Rather, he chooses the problem to fit a specific tool, that is SD. His preference is to use SD and so he looks about for a problem to fit his tool. Jackson and Keys (1984) note that a problem solver's *world-view* will very largely determine the way he sees and approaches problem contexts. Barlas' world-view means that he has a preference to view problems in terms of SD.

Although Barlas has stated that he tends to use one particular modelling process, Lane (1994) and Burrell and Morgan (1979) warn that some form of rigour should be applied when deciding upon the approach that we are taking when analysing a situation. This is one reason for the first aim of this work which attempts to explore the appropriateness of SD as a modelling approach to analyse D&D for litigation.

Vennix (1996) identifies with both sides of the argument. He understands that somebody may only be familiar with SD, or at least more familiar with this tool, than other types of modelling tools and thus be inclined to automatically use it. However, there can then be a danger of translating any problem into a SD problem (p104). This can occur because we may hear what we would like to hear, rather than what a client means (p105). As alternatively put by Maslow (1966) "If the only tool you have is a hammer you tend to treat everything as if it were a nail".

SD is not the only modelling approach where a lack of detail applies when defining criteria regarding the approach's suitability in modelling a situation. In any form of modelling the choice of one approach over another is not a straightforward process. As discussed above in terms of SD, the method used can often be chosen because the individual has a particular modelling interest. When discussing general modelling methods in the context of Operational Research, Rivett (1972, p6) notes that, "... at its very first point operational research is a subjective personal science. From this first subjective personal approach we are each of us influenced by forms of modelling which are attractive to us...". Therefore it is not simply in the field of SD that we find modellers demonstrating a bias towards one modelling tool, it is a general modelling issue within the Operational Research field.

5.7 A Summary of the Most Informative Criteria

This section attempts to bring together the most informative criteria from each of the three sets discussed in chapters 3, 4 and 5. A revised set of criteria can be formed by summarising the most useful criteria with which to assess the suitability of using SD to model D&D for litigation.

Before addressing the revised set of criteria, it is of interest to consider the existing criteria in the SD literature. In addition to Forrester's, Coyle's and Flood and Jackson's criteria, other criteria were highlighted in section 3.2 from authors in the SD literature. These criteria mainly refer to systems that are dynamic and/or contain feedback and were therefore seen to be equivalent to Forrester's criteria. When considering each of the criteria, the element that seems to be missing is a consideration of the *purposes* of the modelling process. As discussed in the last three chapters, authors in the SD field do identify the importance of the purposes of the modelling process. However, this is not captured in the criteria specifically extracted from the SD literature. This will be reflected in the summary of the most informative criteria below.

A second element that is omitted from any of the existing criteria is whether or not patterns of behaviour or detailed quantitative data is required as part of the purpose of the modelling process. Although the SD literature highlights the use of SD in the consideration of the behaviour of a system, the criteria do not explicitly rule out the use of SD to gain detailed quantitative data. Therefore, SD has not been ruled out from playing a role in the more detailed analysis of D&D for litigation.

5.7.1 Flood and Jackson

Although Flood and Jackson's attempt to take a more multi-methodological view of assessing the suitability of SD as a modelling approach for a situation is admirable, their process contains some problems. These problems have been identified as follows:

- SD need not fit into one classification – there has been argument to suggest that complex situations can be tackled by the approach, as well as simple situations.
- The participants dimension focuses on the environment in which the model is built. It is the author's belief that even if you were dealing with pluralists, this cannot completely rule out the potential use of SD. It may still be useful in, say, the construction of two different models to represent two separate sets of beliefs about the situation and thus aid negotiation between two sets of participants. It is agreed that it may not be useful in some forms of pluralist environments, but not all situations should be ruled out.
- It was concluded that measuring complexity is both subjective and difficult.
- It was difficult to reach clear conclusions for some criteria.

Due to the above problems, Flood and Jackson's criteria have not been included in the revised set of *practical* criteria to help define whether or not SD is an appropriate modelling approach to use to model D&D for litigation.

5.7.2 Forrester and Coyle

From explorations in chapters 3 and 4, the following criteria where highlighted as the most important to aid consideration of whether or not the SD modelling approach should be used to model D&D for litigation:

1. *Does the situation contain feedback loops and are they of importance to the <u>purposes</u> of the study?*

2. *Does the situation exhibit changes through time and are they of importance to the <u>purposes</u> of the study?*

3. *Practically, can it be done?*
   In particular:

(a) *Can the project be visualised at an early stage in terms of stocks and flows?*

(b) *Is there sufficient, reliable information and data available which will enable a model to be populated to such a level that the plaintiff's case is in a stronger position when compared to the plaintiff not having the model?*

As pointed out by Coyle, it should be noted that these criteria are not intended to be taken as formal tests. Instead, they should be used as guidelines to assist the discussion between those who are involved in deciding whether or not SD should be used to model a situation.

The above criteria are chosen as the most powerful questions to be asked at the beginning of a study. If any of these questions are answered negatively, then it is likely that SD would not be an appropriate modelling approach for the situation. On the other hand, if they can all be answered positively, then it is likely that SD is a potential modelling approach for the situation. However, it should be noted that it does not answer the question whether SD would be the *most* appropriate modelling approach. It is interesting to note that the majority of these criteria are taken from work dating back to 1977 and that the explorations in this chapter have failed to conclude that any more up-to-date work has proven more useful than this.

It should be noted that the above criteria are still very subjective and dependent upon the worldview of the modeller. An experienced SD modeller, who has a preference to view the world in terms of feedback and SD, will more easily be satisfied that the criteria hold compared to modellers with a different world-view. It is unlikely that any set of criteria could be defined which did not include the need for a certain amount of subjectivity and thus separate the decision from the modeller's personal view. As noted above, "... at its very first point operational research is a subjective personal science. From this first subjective personal approach we are each of us influenced by forms of modelling which are attractive to us..." (Rivett 1972, p6).

If the criteria above were used as criteria defined by SD to assess whether or not a SD approach can be used to model D&D in the projects discussed in this thesis, then they would suggest that SD could be used. However, it does not inform us whether SD is the *most* appropriate modelling approach.

5.8 Alternative Terminology used to Describe Criteria in the SD Literature

To arrive at a summary of the most useful criteria to use to assess the suitability of using SD to model D&D for litigation, three separate views were chosen from the SD literature. Each of these views used particular terminology when describing criteria. However, it should be noted that alternative terminology is commonly used in the SD literature to describe the system characteristics that help to identify the suitability of SD as a modelling approach. To enable an appreciation of the different terminology that is used, the following lists the most commonly used phrases in the SD literature other than those already discussed in the thesis:

- *Structure causes behaviour* – The existence of feedback loops in a system was discussed in chapter 3. The dynamic behaviour that is produced by a feedback system is created from its internal structure (Forrester 1968a, p4-2). A SD study therefore focuses on the relationship between structure and behaviour. An investigation of the structure of the system therefore needs to be an appropriate approach to take in order that a SD modelling approach can be used to model the system.
- *Accumulation of resources* – SD may be discussed in terms of the dynamic consequences of accumulation and depletion. If the accumulation and depletion of resources over time are important to the purpose of a study, then SD can be used to address the issue (Warren and Langley 1999).
- *System non-linearities* – A SD study will model the non-linear interactions between basic feedback structures in an attempt to reproduce the overall behaviour of a system. A focus on non-linearity in SD is highlighted by Richardson (1991) in his investigation into the history of feedback. He states that "... nonlinearities are viewed by system dynamics practitioners as vital determinants of the interesting or problematic behavior of a dynamic social system" (p308). Therefore, SD can be used to model non-linearities in a system in an attempt to gain an understanding of the behaviour of the system.
- *Endogeneous system behaviour* – "System dynamics seeks endogeneous explanations for phenomena... An endogeneous theory generates the dynamics of system through the interaction of the variables and agents represented in the model... In contrast, a theory relying on exogeneous variables (... that is, from outside the boundary of the model) explains the dynamics of variables you care about in terms of other variables whose behavior you've assumed" (Sterman 2000, p95). This statement highlights the fact that SD should only be used if the behaviour of the system can be explained

through endogeneous behaviour i.e. the behaviour can be fully explained through the interactions of variables *inside* the system boundary.

5.9 Summary of Chapters 3, 4 & 5

The last three chapters have explored:

- *the appropriateness of SD as a modelling approach in the analysis of D&D for litigation.*

Criteria laid down by the SD literature have been used to aid the explorations. Although this process has highlighted a lack of detail in the SD literature, a set of the most informative criteria has been detailed. These criteria conclude that it is *possible* to explore the projects discussed in this thesis using the SD modelling approach. However, there is no indication whether or nor this is the *most appropriate* approach to use.

Potential reasons for the lack of detail in criteria in the SD literature have been discussed. One reason that was highlighted for this was due to the subjective nature of the choice of any modelling approach. An individual's worldview influences the choices they make and hence two different decision makers may result in taking two different decisions when choosing a modelling approach to use for a problem situation.

Coyle (1973, p400, Italics added) points out that "... it is impossible to think of even the simplest operation such as opening a door, which is not a feedback process. Whether it is worth explicitly modelling the feedback structure instead of merely treating the dynamic behaviour as a stochastic process as in stock control is quite another matter. The key lies in the *purpose* of the model". The *purpose* of a study plays an important role in considering an appropriate modelling approach to use for a situation. Indeed, two of the criteria summarised above explicitly require knowledge of the purpose of the study to be able to assess the suitability of SD to model a situation. For this reason the purposes of modelling D&D for litigation will form the focus of the next chapter. This chapter will consider whether or not the SD modelling approach can meet the purposes of the modelling process.

CHAPTER 6: THE PURPOSES OF MODELLING D&D FOR LITIGATION

6.1 Introduction

The previous three chapters have considered the first aim of this thesis:

- *to explore the appropriateness of SD as a modelling approach in the analysis of D&D for litigation.*

Explorations using the SD literature were supported by the use of data gathered as a part of the author's involvement as both researcher and consultant in modelling D&D in the litigation projects detailed in chapter 2. The aim of this chapter is to consider the second aim of this thesis:

- *to explore the issues involved in using SD to analyse D&D for litigation.*

The explorations will be carried out through consideration of *the purposes of the modelling process*. These explorations will again be supported by the use of data gathered during the modelling of D&D in two litigation cases.

From section 2.4.3, the purposes of modelling D&D for litigation were summarised as the following:

1. Demonstrate that a part of the time and cost overruns were caused by D&D through particular triggers of D&D.
2. Replicate over time the hours due to D&D which were over-and-above those that were contracted, but were required to carry out the project.
3. Demonstrate the extent to which the client management of the project was reasonable and the extent that overruns could not have been reasonably avoided.
4. All the above have to be demonstrated in a way which will be convincing to the several stakeholders in a litigation audience.

When considering the above modelling purposes, a possible set of criteria that can be used to assess the appropriateness of a modelling approach to model D&D for litigation can be summarised as follows:

The approach must, at least, be able to:

From purpose 1:

(a) Model exogenous triggers and their outcomes as D&D.

(b) Model the paths of argument from an action to an eventual outcome (so that particular triggers can be seen to have caused D&D).

From purpose 2:

(c) Quantify the outcome of D&D.

From purpose 3:

Again, this gives a need to:

Model the paths of argument from an action to an eventual outcome (so that management actions can be evaluated and audited).

From purpose 4:

(d) Replicate reality convincingly, including:

 (e) Be transparent to lay people from a variety of backgrounds in addition to ...

 (f) ...being sufficiently rigorous to stand up to expert critique.

The feasibility of SD being able to fulfil each of the above six criteria (a) to (f) will be discussed in turn.

6.2 Model Exogenous Triggers and their Outcomes as D&D

In section 2.4, D&D was explored and defined. It was noted that one of the most significant circumstances for D&D arises from a feedback phenomenon. For this reason, D&D analyses often focus on feedback. Therefore, to be able to model D&D, it is vital that a modelling approach is capable of modelling feedback. SD is a modelling approach that was specifically designed to model and explore feedback. Since SD has been designed with these properties, it *should* be capable of modelling D&D.

To explore whether or not SD has actually been capable of modelling triggers and their outcomes as D&D in the projects discussed in this thesis, the main exogenous triggers that occurred during the projects need to be considered. From section 3.4, three groups of disruptive triggers emerged from the triggers that occurred in the Castle and Pirate projects.

These were summarised as the following:

- Disruptions that halt the flow of work.
- Additional tasks that do not interrupt the workflow, but require scheduling.
- Tasks that take longer than expected to complete.

The consequences of each of these groups of triggers are that managerial actions are normally taken in response to them. The modelling method used therefore needs to be able to model the triggers and managerial actions as D&D. The following diagram summarises the three groups of disruptive triggers listed above and their main outcomes. Note that the two lines between concepts 3 and 4 represent a delay.

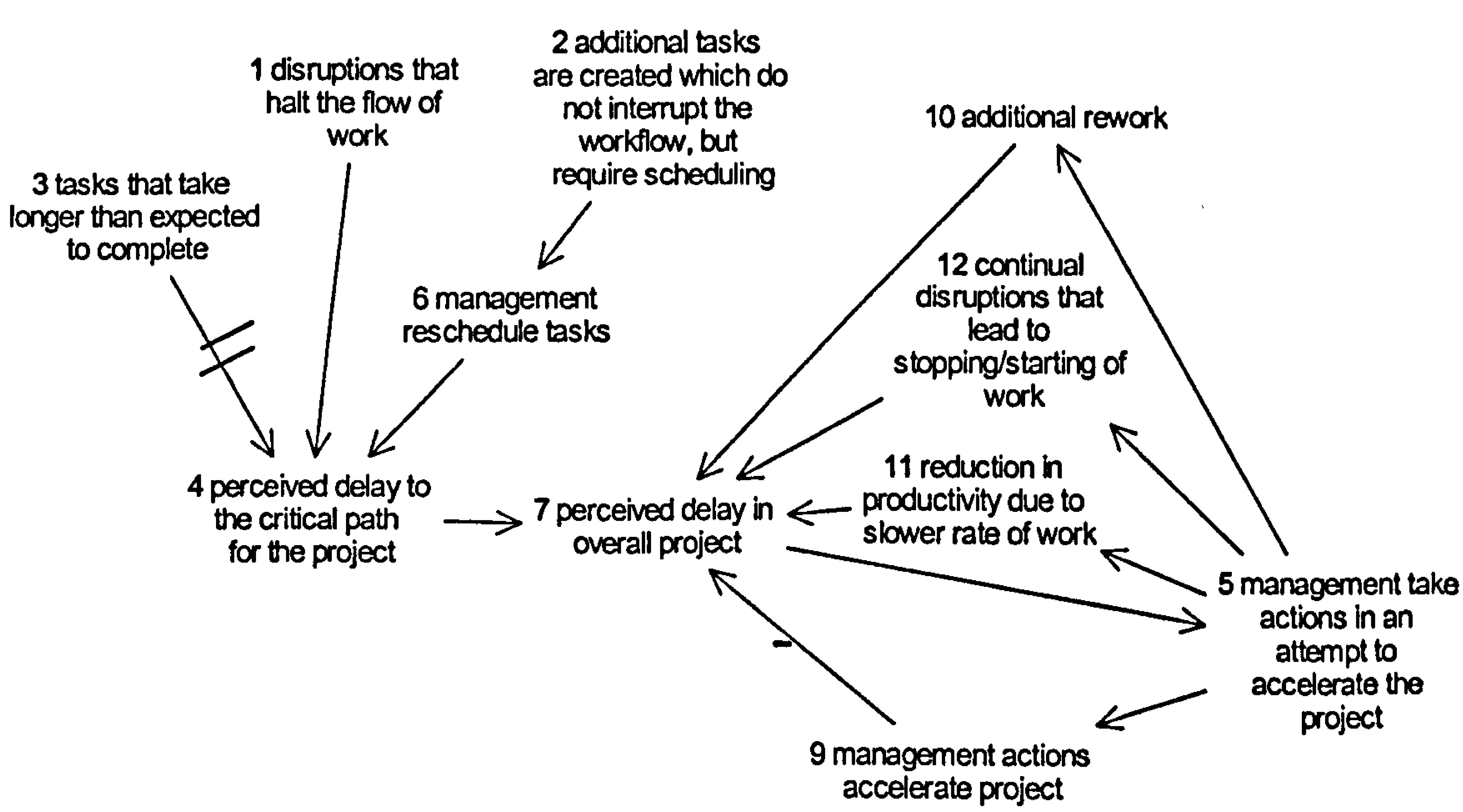


Figure 6.1 – A Summary of the Main Groups of Triggers and their Outcomes in the Castle and Pirate Projects

Based on figure 6.1, the modelling method would need to be able to model the triggers shown on the left-hand side of the diagram as well as the outcomes and the feedback shown on the right-hand side of the diagram. The capability of SD to model this is captured by Coyle (1977, p5) when he states that SD "...is that part of management science which deals with the controllability of managed systems over time, usually in the face of external shocks". This highlights that SD was intended to be used to model managerial actions that occur when exogenous triggers hit a system.

However, to be fully capable of modelling the concepts and the feedback dynamics shown in figure 6.1, additional criteria need to be met. Rather than simply modelling an overview of triggers and their outcomes, the full paths of argument from triggers to outcomes need to be modelled. Also, if SD is to prove to be a successful tool in a litigation environment, then it needs to be able to quantify the outcomes of D&D so that the plaintiff can support a claim for compensation. These issues will be discussed in sections 6.3 and 6.4.

6.3 Model the Paths of Argument from an Action to an Eventual Outcome

When discussing different types of claims in construction projects, King and Brooks (1996, Italics added) state that "In order to prove damages, the claimant must *establish a causal link* between the event that gave rise to the claim and the increase in project costs". When considering the need to model paths of argument in a litigation environment, this means that SD should be able to support the establishment of causal links between an action and an eventual outcome. Before fully addressing this criterion, a clarification will be sought of what is meant by a *causal link* when modelling D&D in a litigation environment.

6.3.1 Causal Links in Litigation

It is first necessary to determine what is required in a litigation environment when discussing the modelling of paths of arguments. There is a need to provide a clear route between an action and the contribution the action has towards an outcome. Documentary evidence is very useful in helping to support the route between action and outcome. However, this type of evidence is normally used to support individual events rather than generic links. For example, evidence may be used to support a specific change to the contract leading to specific rework being carried out. However there may not be any *one piece* of documentary evidence to support changes in the contract *generally* causing rework. If the occurrence of a number of individual events can be supported through documentary evidence, then a proposition can be made that the generic link exists.

In any project there will be thousands of events that occurred which could be traced to discover the eventual outcome on the project. To support a case for causal links between events and outcomes, an efficient recording system is required so that any useful documentary evidence can be easily found and used to provide proof of the occurrence of the events and their outcomes.

During the experience of the Castle project a mass of documents, largely consisting of communication between the defendant and plaintiff, were gathered to support statements made by the plaintiff. Another use of these documents was to support work carried out by external consultants hired by the plaintiff's team. The consultants analysed disruptions that affected the project's critical path. Events that occurred during the project were associated with disruptions to the critical path. For example, an extension on the time taken to carry out a group of activities was explained by various events such as the delay in the plaintiff receiving Customer Furnished Information (CFI) from the client. In arguing such a case, a lawyer would use the documentary evidence to attempt to illustrate that there was a clear, non-contentious route between the delay in the CFI and the outcome on the critical path.

6.3.2 Causal Links when Modelling D&D

As part of modelling D&D, two different forms of causal linking have been used:

- Detailed stories are constructed which trace the outcomes of individual events (Williams et al 2000). Documentary evidence is very useful in constructing these stories so that the route between each cause and effect can be substantiated. These individual cause and effect stories are then used to help support the more generic causal links developed in the generic cause maps.
- Generic cause maps which can then lead to influence diagrams. In the modelling of D&D, cause mapping has been used as a qualitative tool to model the lines of argument from an event occurring in a project to the consequential outcomes of that event (Ackermann et al, 1997). Eden (1994) discusses the process in moving from a cause map to an influence diagram and, in particular, how each informs the other during the whole of the modelling process. In a litigation environment, the cause map and

influence diagram are used to illustrate a plausible story which attempts to explain the outcomes of the project.

6.3.3 Causal Links in SD

The above has considered how cause mapping models the causal links between an event and an outcome. However the question being addressed in this section is whether or not *SD* is capable of modelling the causal links between an event and an outcome. Qualitative cause maps enable paths of arguments to be modelled and observed. However, the quantitative SD model also needs to be able to model these paths of arguments. Not only does the model need to fulfil this requirement, but the model's audience also needs to be able to observe and understand the links in the quantitative model. This leads to the criterion that will be discussed in section 6.6; the transparency of a model to lay people.

The criterion considered in this section leads towards a preference for the use of *structural* rather than *black box* modelling approaches (Mitchell 1993, p120). Mitchell explains that structural models identify real world relationships, whereas black box models do not explicitly take cause and effect into account, but look for patterns in the data instead. The need for a structural model is further emphasised as Mitchell continues to explain that a black box model "... is, on the whole, less likely to be sympathetically received by a layman than is a structural model". This leads to the question of whether or not SD models could be categorised as structural models. The SD process captures the cause and effect relationships within a system, particularly focussing on any feedback loops created by the relationships. To enable a computer simulation to be constructed, quantitative data is required to enable the qualitative relationships to be quantified. This process fits with Mitchell's description of a structural model which "... identifies relationships which the modeller believes holds in the real world and uses the data only to fit these relationships with numerical parameters" (Mitchell 1993, p121). SD models therefore do fulfil the criteria to enable them to be labelled *structural* models. When compared to black box modelling approaches, this provides *some* evidence of SD being capable of modelling paths of arguments and having some level of transparency for lay people.

Some SD models in some SD software packages may have a problem with their level of transparency when attempting to trace the paths of argument underlying the model. In many simulation models the paths of argument have to be traced around the levels and rates. This can hinder the visibility of the causal links in the model. An exception to this is the use of the 'Uses' and 'Causes' Tree tools in Vensim[2]. With these tools causal tracing can be used as "a powerful tool for moving through a model tracing what causes something to change" (Ventana Systems 1998). The potential lack of transparency when tracing paths of argument in SD models was considered when the team at Strathclyde University developed a cyclical process (Eden 1994, Ackermann et al 1997) to model D&D for litigation. This process involves the SD model being continually updated to reflect the cause maps and influence diagrams and vice-versa. The SD model therefore becomes a quantification of the cause map or influence diagram. This means that if this approach is closely followed then any causal links appropriate to D&D in the cause map would be reproduced in the SD model. A validation of the causal links in the cause maps should therefore be equivalent to a validation of the causal links in the SD model. The aim of such an approach is to increase the level of transparency in the SD model.

Although the team at Strathclyde University intended their cyclical process to aid the transparency of the SD model, the qualitative models generally provide a clearer route than the quantitative models to trace the paths of argument around concepts and arrows. On the other hand, it is more difficult to visualise a clear path from event to *final outcome* in the qualitative models. The reason for this is that the outcome may arise from a dynamic interaction of feedback loops that may not be immediately obvious from a qualitative model. However, the simulation model provides a picture of the eventual outcomes of each event through its outputs. This means that the audience is not required to attempt to work out the eventual outcome of events from the dynamic behaviour implied by all the interactions within a cause map or influence diagram. This is important, as it has been noted as a particularly difficult process to undertake (Sterman 1989a, Sterman 1989b, Paich and Sterman 1993, Diehl and Sterman 1995). However, this may mean that the audience can lose faith in the SD model, since tracking the seemingly complex quantification process being undertaken by the simulation would be so time-consuming that it would not be a plausible approach to take. This would hinder achieving a transparent *quantitative* link between the events and outcomes when modelling D&D in a litigation environment.

---

[2] Vensim is a proprietary product of Ventana Systems, Inc. Harvard MA

6.3.4 Modelling Causal Links in the Castle Project

As an example of tracing the paths of argument from an event to an outcome, figure 6.2 considers the paths of arguments arising from the occurrence of excessive configuration change requests (CCRs) and engineering change proposals (ECPs) in the Castle project. Both of these caused changes to the designs throughout the project.

Figure 6.2 illustrates how an argument is built-up that relates a plausible story of the outcomes of the occurrence of an excessive number of CCRs and ECPs. The argument is built up step-by-step moving from one concept to the next, along the arrows. This provides a relatively transparent story for an audience to follow.

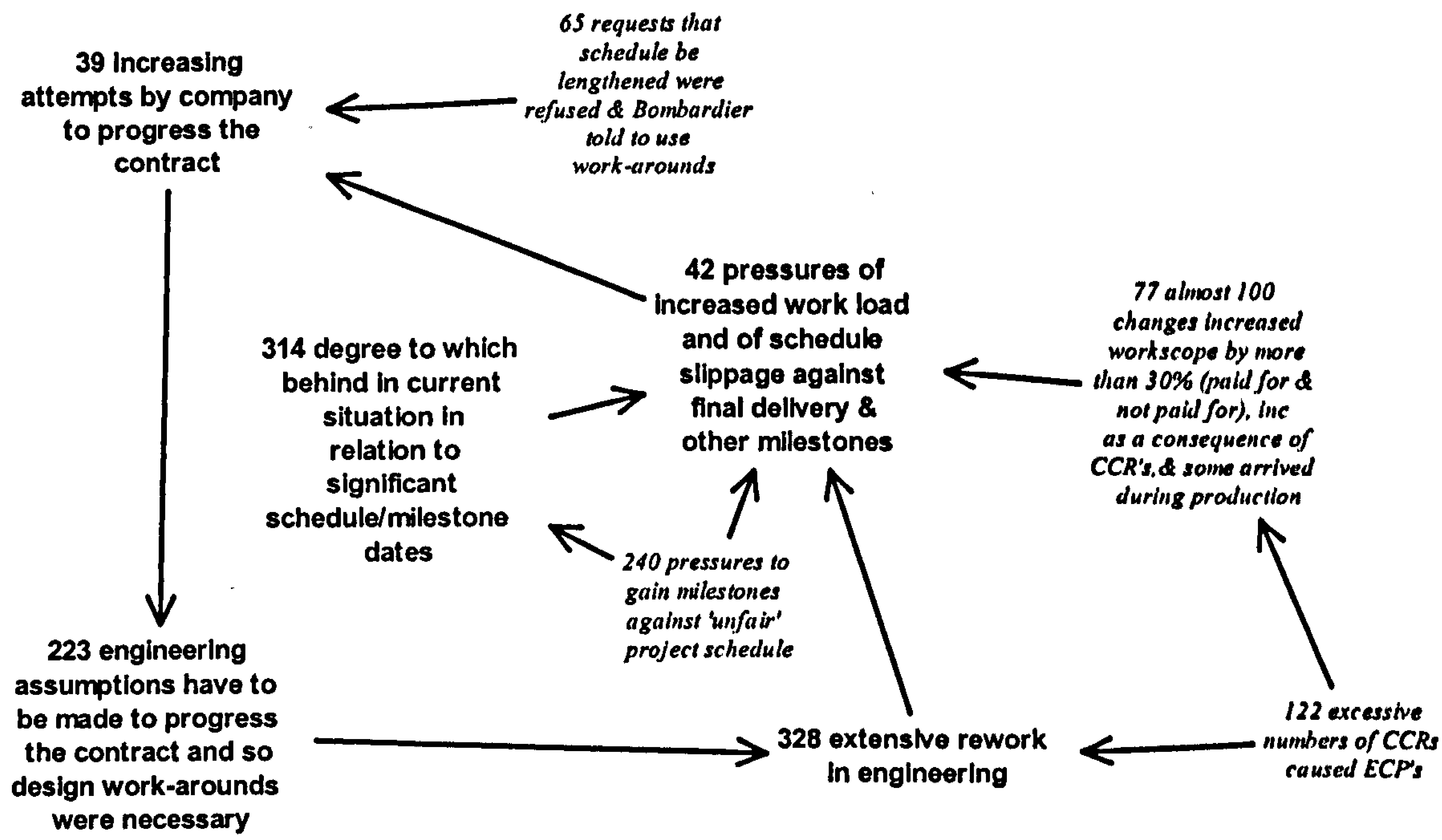


Figure 6.2 – Some of the Outcomes of an Excessive Number of CCRs and ECPs in the Castle Project

When the arguments in figure 6.2 were translated into the SD model, not all elements of the line of argument were modelled. For example, concepts 39 and 240 were not explicitly modelled. Concept 240 was included as a part of concept 42. Also, the 'pressure of increased workload' fed directly into 'engineering assumptions have to be made', omitting concept 39. This concept is implicitly considered when modelling concept 223 and therefore could have been added as background information when documenting the variable. For this example, when comparing the SD model with the cause map model, the reader would not be able to follow the same line of argument in the SD model as clearly as they can in the cause map.

The above gives examples of variables that existed in the cause map but were omitted in the SD model. However, a number of additional variables were also required in the SD model that were not explicitly considered in the cause map. These additional variables also meant that it was more difficult to follow the main line of argument from any part of the model. For example, the section of the SD model that relates to the calculation of the model inputs that were used to represent CCRs and ECPs is shown in figure 6.3.

In this case, the quantification of CCRs and ECPs meant that the 24 variables shown in figure 6.3 were required simply to represent concept 122 and some of concept 328 in figure 6.2 in the SD model. Although it is recognised that this section of the SD model could have used less variables, the number would still be greater than the number of original concepts. Without linking these variables to the one or two concepts they relate to in the cause map, somebody attempting to gain an understanding about this section of the model could spend too much time reviewing each variable and link. However, these variables and links are at a detailed level below that required to understand the main paths of argument relating to the dynamic hypothesis that is being modelled.

The use of the qualitative cause map and, if required, the influence diagram with the quantitative SD model is the preferred approach to be able to model the causal links between an event and outcome. The reason for this is that a SD model, on its own, may not be sufficiently transparent when modelling D&D in a litigation environment.

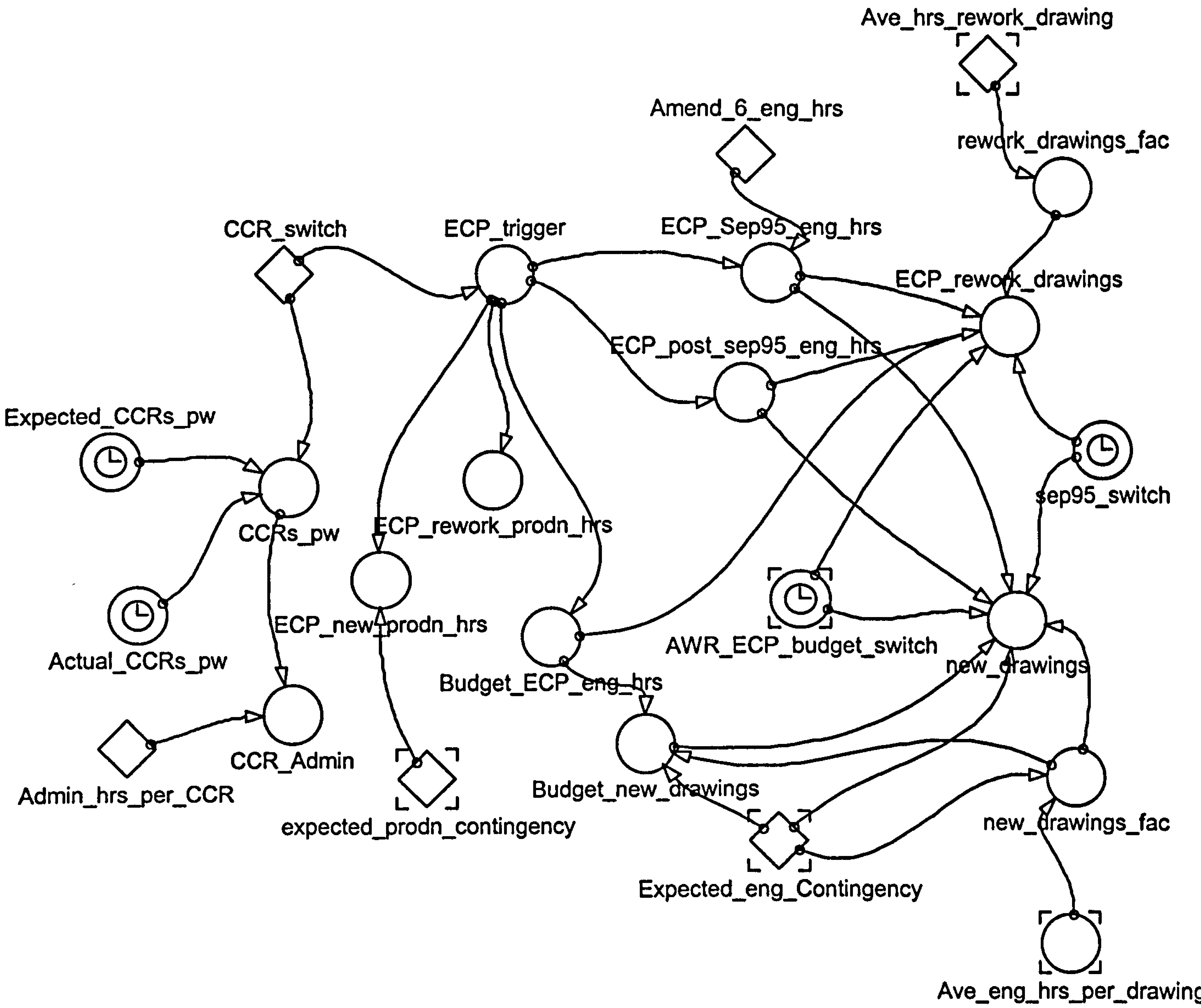


Figure 6.3 – Calculations of the Inputs used to Represent CCRs and ECPs in the SD model for the Castle Project

6.4 Quantify the Outcome of D&D

The first question to ask with respect to this criterion is *what is the outcome of D&D?* As a result of the occurrence of D&D in a complex project, the project may experience delays and cost overruns (Eden et al 2000). The cost overruns may arise from the need for additional labour hours, materials and overheads such as project planning and liasing with the client. There may also be outcomes that go beyond affecting the individual project. For example, other projects the plaintiff is managing may also be disrupted due to labour needing to be transferred from these projects onto the original project. There may also be implications on future projects for the plaintiff. For example, losing out on future work due to the delays experienced on the original project. This criterion is therefore asking whether or not SD can quantify the delays and cost overruns that may occur due to the occurrence of D&D in a project.

6.4.1. Delays

Traditionally, CPA has been used as a tool to de-construct a project into a network of activities before proceeding to evaluate the expected project duration (Lockyer and Gordon 1991). As a project proceeds, this tool has also been used to monitor the project to assess whether or not it is progressing as expected. CPA has also been used during claims to assess the effect of delays and disruptions on the duration of a project (Scott 1993). However, Eden et al (2000) argue that using CPA in such a way does not take full account of the effects of D&D. For example, it cannot take account of 'soft' human elements or the more strategic systemic effects (Morris and Hough 1987, Cooper 1994, Williams 1997).

Delays in the Castle Project

For the Castle project, external consultants used CPA to analyse delays to the project. For each of the major events that occurred during the project, the resulting delay to an activity was assessed and the effect on the critical path analysed.

The SD model constructed for the Castle project also demonstrated a delay to the project. However, this delay was the result of different effects to those assessed by the CPA model. The delays in the SD model resulted from feedback that was created from the effects of managerial actions taken in response to expected schedule slippage. It was therefore not surprising that the overall delays demonstrated by the two models differed. The CPA model experienced the limitations mentioned above. For example, the CPA method could not take account of actions taken to accelerate the project such as the use of parallelism in activities. Therefore, the CPA model would have demonstrated a delay that was greater than the delay actually experienced on the project. Also, the SD model did not allow for the full delay that occurred in practice since it did not capture the detailed operational issues that could be captured by a CPA. Rodrigues (2000) identifies the inadequacies of SD to capture the more detailed, operational issues of project management. In an attempt to combine the advantages of SD and the traditional processes, Rodrigues has defined a framework for integrating the two as an improved project management process. This work was carried out in the context of software development projects. Part of the very detailed framework proposed by Rodrigues formally links SD models with CPA. The links proposed are based on the activities required during the planning and monitoring stages of a project. They are not aimed at a post-mortem analysis of a project. The complexity of the approach means that Rodrigues suggests that if an organisation wishes to implement the framework, it should be done gradually. The detail and complexity of the approach means that it is unlikely to be suitable for organisations to adopt for use in a post-mortem analysis of a project as part of litigation.

6.4.2 Cost Overruns

In the projects that the author has been involved in, the focus of the SD model has not normally been to model *direct* costs. These are additional costs to the project that do not require an analysis of feedback dynamics to explain their occurrence. For example, the cost of additional material due to an increased scope of work, or the cost of additional overheads due to the need for additional management hours due to an extended time frame. The SD models have focussed on the *indirect* costs arising from additional labour hours. These additional labour hours are explained through the feedback modelled in the cause maps.

However, direct labour hours may be included in the SD model. For example, in the Castle project additional direct labour hours accrued in production due to learning being unable to be achieved between the three aircraft. This outcome arose due to unexpected parallel working and higher than expected levels of individuality between the aircraft. The additional direct hours that accrued were required as inputs to the SD model so that the additional indirect labour hours, that were created as a result of the direct hours, could be assessed. The use of direct labour hours in the SD model was therefore only to enable a full assessment of the indirect labour hours that accrued on the project.

Due to the feedback effects in D&D, it is vital that a modelling approach is capable of quantifying feedback. SD is a computer simulation technique that uses mathematical formulae to quantify relationships between variables of a system. This process enables the overall dynamic behaviour of the feedback loops in a system to be quantified. Since SD has been designed with these properties, it *should* be capable of quantifying D&D.

A concern for the modeller, when requesting the SD model to simulate the cost overrun that actually occurred on the project, lies with the level of accuracy of the parameters used as inputs. A potential lack of accuracy has been due to, for example, problems with data gathering or the use of subjective data. This means that a modeller would be reluctant to use the SD model to generate a single cost result. Therefore, it has been normal procedure to provide a range of results as an indication of the range of additional labour hours that could be reasonably supported through the argumentation provided by the models.

As mentioned at the beginning of section 6.4, the D&D that occurs on one project may cause costs to accrue on other projects. These costs are normally omitted from the SD model. The reasons for the omission of these costs are due to the difficulty in modelling them. The costs arise on other projects that are being managed by the plaintiff's organisation. The disruptions in the project may lead to, for example, transferring labour from other projects onto this project. By doing so, the other projects will also be disrupted and hence accrue costs associated with this. Also, future projects may be affected. This can occur due to a delayed project postponing the commencement of work on other projects. Or, alternatively, the outcomes of the disruptions on the project may cause other clients to be put off dealing with the plaintiff as they perceive the cost and time overruns to be caused by incompetent management. Costs associated with issues such as the potential loss of business would be very difficult to quantify. However, they may still be argued to have resulted from the D&D that was experienced on the original project.

The criterion discussed in this section takes the modelling process beyond qualitative modelling and requests the need for a quantitative tool. The previous criteria require that a plausible story be presented and argued. However, this criterion highlights the primary objective of litigation. Even although a plausible story can be argued and agreed, the end result needs to be that, assuming the plaintiff's claim is upheld, a sum of money has to be allocated to the plaintiff as a result of the D&D that the defendant was deemed to be responsible for. This criterion provides aid in arriving at that sum of money.

When testing SD against this criterion it has been found that SD has limited capabilities. For example:

- SD does not incorporate any delays associated with the disruptions at a detailed operational level such as the critical path.
- SD will not incorporate the cost of disruptions to other projects, or the costs associated with more long-term issues, such as the potential loss of future business, arising due to D&D that occurred on the original project.

However, as long as a modeller is aware of the capabilities of the modelling approach, SD can still be used to provide a plausible range of cost overrun on the original project. This range gives an indication of the additional labour hours that could be reasonably supported through the argumentation contained in the models used.

6.5 Replicating Reality Convincingly

The second modelling purpose has been stated as:

- *Replicate over time the additional hours, over-and-above those that were contracted, which were required to carry out the project.*

A modelling approach that will replicate these hours needs to do so in a way that ensures that the audience is convinced that the modelling approach is adequately replicating an historical reality. It is reasonable to assume that there is no need for the model to replicate each hour of the project exactly as it occurred. A model may be defined as a *simplified* representation of the real system and thus will omit elements of the real system. If every detail of the real system were included, the model would be able to replicate the exact behaviour of a system. However, it would then be as large as the real system and the benefits of having a model would then be lost. Therefore, if it is assumed that the model needs to replicate reality *convincingly*, the acceptable level of replication will depend upon what the model's audience believes to be acceptable. It is therefore important to have an understanding of who the model's audience will be.

Depending upon the stage of the litigation process that a project reaches, the model may potentially face the following variety of audiences. The aim of the model for each of the audiences is also given.

- Modellers: The modellers themselves need to be convinced that their model adequately replicates reality before placing it in front of an audience. From a personal point of view, the modellers will be looking to produce a model that reflects 'good', professional modelling.
- Plaintiff's claim team: The plaintiff's claim team is the modellers' client and they would need to be convinced that it should be used as part of their case for compensation. The team will be considering the likelihood that the model will help support a successful claim for compensation from the defendant. Also, they will be considering whether or not the model will help maximise the compensation that they can gain.
- Modelling Experts hired by the plaintiff: The plaintiff may use another modelling expert to audit the model. The experts will attempt to discover possible ways of attacking the model, so that potential attacks on the model can be pre-empted before it is shown to the defendant.
- Defendant and any modelling experts hired by the defendant: Before any litigation case, the plaintiff and defendant may consider a period of negotiation. This process can be used in an attempt to resolve the dispute and avoid a court case. The defendant will judge the model on how well it supports the plaintiff's case and if they believe it could stand up to critique in a court of law. If they are convinced, then they may consider settling during negotiation to save both time and money.
- Judge: A judge can become part of the process during arbitration, a final step before going to litigation, or as part of a court case. At this point in the process, the judge will wish to be convinced that the results of the model are valid if he or she is to take any notice of them when arbitrating between the plaintiff and the defendant.
- Jury: Finally, if litigation cannot be avoided, a jury will need to be convinced of the adequacy of the model. They would need to be convinced to the point that they believe that the model adequately supports the plaintiff's case for compensation from the defendant.

When modelling D&D for any litigation case the modellers will always be the initial audience who need to be convinced that the model adequately represents reality. The majority of the author's experience, as a part of the projects discussed in this thesis, involved convincing this audience. Out of all the audiences, the modeller has the most rigorous process for the model to go through in order to be convinced that it adequately replicates reality. Due to this rigorous process and the importance of this audience (i.e. if this audience is not convinced by the model, no other audience is likely to see the model), this will be dealt with separately in chapter 7.

When considering the 2nd - 6th audiences, the aims of the model for each of the audiences are given above. Convincing the 2nd audience that the model replicates reality sufficiently will be of particular importance to the modeller. If the plaintiff's claim team is not convinced, then since the plaintiff had first-hand knowledge of the project being modelled, it is unlikely that the modeller will be convinced that the model adequately replicates reality.

At the other extreme, if the plaintiff's claim team is convinced that the model replicates reality sufficiently, then the model may play a role in persuading the plaintiff that they have sufficient evidence to support a claim for compensation.

Before the 2nd- 6th audiences each attempt to make a judgement about whether or not they believe that the model replicates reality convincingly, they each need to gain a basic understanding of the mechanics of the model. This means that the next criterion i.e. *Be transparent to lay people* is of vital importance to enable each of the audiences to gain a reasonable understanding of the model.

6.6 Be Transparent to Lay People

As discussed in section 6.3.3, Mitchell (1993) would advise the use of a *structural* rather than a *black box* modelling approach to ensure that the model is more easily accepted by lay people. It was concluded that SD models fulfil the criteria of structural models and therefore should have at least a minimum level of transparency to lay people. However, when Williams et al (2000) consider the use of SD to model D&D in a litigation environment, they argue that the SD model itself can be "black boxish". This is because the

"SD model provides a single figure – the quantum – it does not provide a clear and easily understood representation of what occurred. *Through the necessity of reducing the project's life to its key variables, the model becomes opaque*" (Italics added). However, they continue by stating that the modelling process that they use means that the SD model has been developed through a transparent process. This section will consider whether SD models produced when modelling D&D in a litigation environment *can* be transparent to lay people.

There is evidence to suggest that computer simulation models have been found to be useful in litigation cases. "The advantage of a simulation model is its ability to portray a complicated situation better than any verbal description could hope to do. A model can disclose relationships between various events which might not otherwise be apparent. Also, a model makes it possible to consider all relevant factors simultaneously in the solution of the problem" (Fleming, 1980, p874) "... litigants have successfully used the results of computer simulation as the basis for expert testimony. Though there are few reported decisions dealing with the admissibility of computer-aided simulation models, the courts seem more than willing to accept these models, provided that they are carefully constructed." (Fleming, 1980, p875). These quotes indicate that there is evidence that those present in court cases have understood simulation models and found them useful in portraying situations.

However, difficulties in getting the different audiences to fully understand the SD model have arisen during the projects discussed in this thesis. These difficulties focus on the lack of transparency of the SD model. The causes of these difficulties will now be discussed.

6.6.1 Size of the Model

When modelling D&D in connection with the litigation cases discussed in this thesis, there is a requirement to:

- *Replicate over time the hours due to D&D which were over-and-above those that were contracted, but were required to carry out the project.*

This requirement has meant that the level of detail required from the model has led to relatively large and complex SD models. In particular, additional variables are required beyond the cause map to enable the model to simulate reality as convincingly as possible. For example, in the Castle project, the cross-impact effect experienced in the engineering function was illustrated in the cause map as follows:

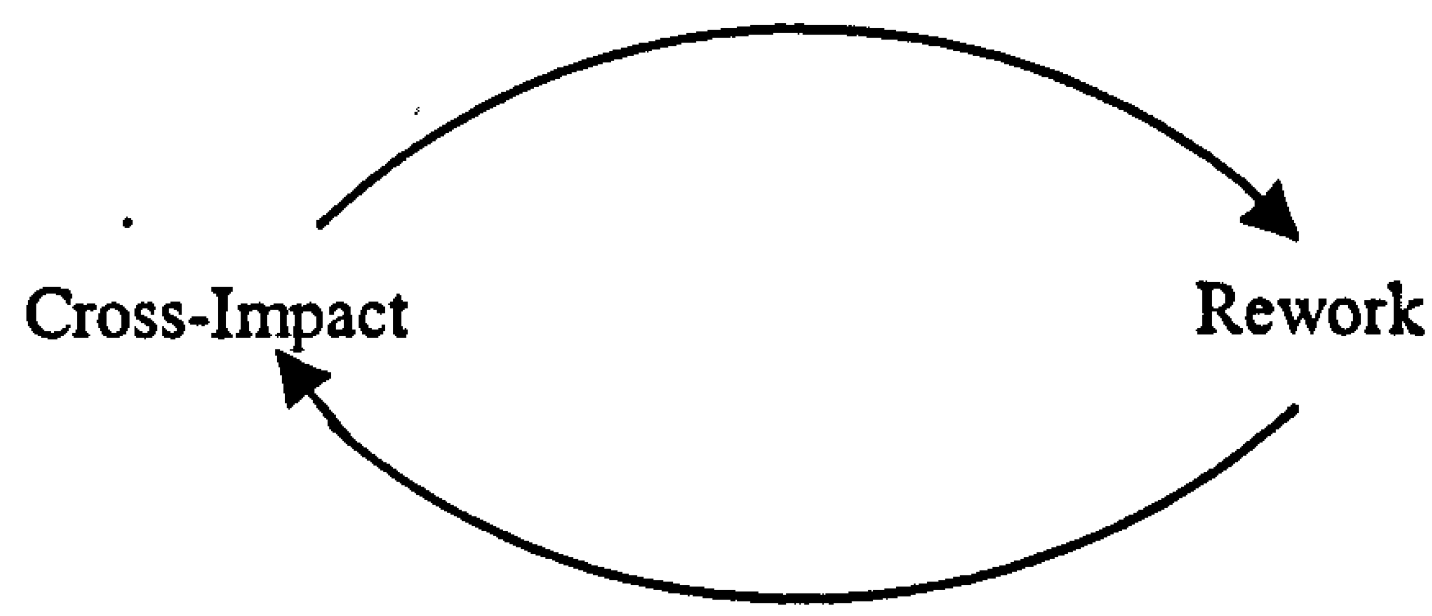


Figure 6.4 – Rework/Cross-Impact Loop

However, the translation of this loop into the SD model required the variables shown in figure 4.4, which is replicated below. It should be noted that many of the variables shown in figure 4.4 were also replicated in the cause map for the Castle project. However, they were modelled to capture other effects in the project rather than the cross-impact /rework loop.

It is not possible to simply model cross-impact and rework in the SD model. Stocks and flows are required to represent the flow of drawings over time, as these are the units that are affected by cross-impact. Also, variables such as *available_productivity* are required to simulate the workforce actually doing work on the drawings. The increase in the number of variables increases the size of the model. This therefore increases the amount of information an audience needs to understand in order to gain an overall understanding of the model.

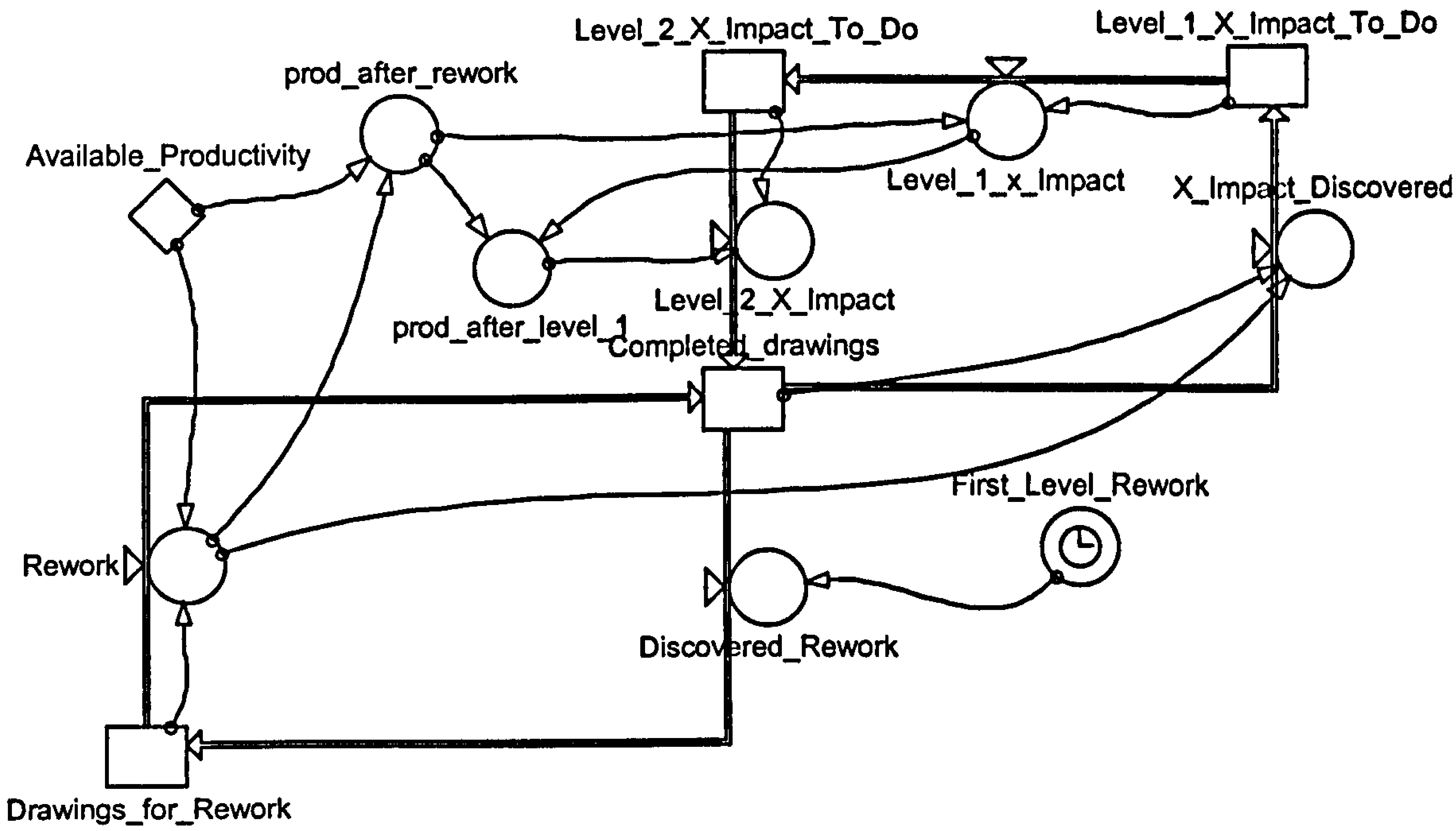


Figure 4.4 – Improved Structure for Modelling Cross-Impact in the SD model for the Castle Project

6.6.2 Quantitative Reasoning Compared to Qualitative Reasoning

To have an understanding of the cause map and/or influence diagram, the audience is required to follow a line of argument. However, to gain a full appreciation of the SD model, the audience would be required to follow quantitative reasoning. This means that in addition to understanding the causal structure of the SD model, the audience needs to be able to gain an appreciation of the manner in which data has been used in the model. An audience may not have the ability to do so, or may find this difficult.

Although SD does have some history of being used successfully in a courtroom (Cooper 1980, Weil and Etherton 1990), it is interesting to note the attitudes of the plaintiffs' lawyers who worked on the projects discussed in this thesis. Their views have been that they could perceive a use for the SD models during a period of negotiation. However, they have stated that they would be unlikely to use the models in a courtroom. Their reason for this decision was due to the models' level of complexity. They would not feel comfortable presenting the mathematical complexity of the models to a judge or jury. They felt that the complexities of the model might alienate a jury since they did not consider the SD model to be sufficiently transparent for courtroom proceedings. Although Williams et al (2000) argue that the transparency is embedded in the process taken to construct the SD model, this has not been a satisfactory argument for the lawyers to be confident that this would work in a courtroom. For example, one of the lawyers in the Castle project was particularly concerned that the lack of understanding of the quantitative relationships in the SD model by lay people could lead to the misconception that the model could lead to "anything we want it to". The modellers may be used to prove that this was not the case or expert modellers may be brought in to verify the model. However a judgement in a court case may be made based upon what the judge or a jury feel comfortable with and fully understand themselves.

However, lawyers have felt that the model may be useful during a period of negotiation. The detail of the model was not seen to be as important in this environment. The aim during this period would be to avoid a long drawn out courtroom process and the model would be presented as expert analysis to support testimonies, argumentation and a mass of documentation. The purpose here would be to attempt to illustrate that the plaintiff had gathered a mass of evidence to support a court case to request compensation for D&D caused by the defendant. Of course, the audiences still need to gain some understanding of the SD model during negotiation. This has been approached in the projects discussed in this thesis through reports and/or presentations by the modellers.

6.6.3 Lack of Time the Audience is Exposed to the SD model

The literature suggests that a lack of transparency in a SD model can be avoided by involving the audience in the construction of the model (Lane 1992). Indeed Sterman (1994, p320) suggests that "In practice, effective learning from models occurs best, and perhaps only, when the decision makers participate actively in the development of the model." Participation in model development would therefore improve the audience's chance of understanding what actually drives the behaviour of the system. However, since D&D models are relatively detailed, a full appreciation of the model would mean that any audience would need to be involved with the model over a long period of time. The modeller is certainly exposed in this way and, depending upon the working relationship and involvement in the model, some members of the plaintiff's team could also have a reasonable amount of exposure to it. However in a litigation situation, transparency of the model must also be required to extend to the defendant's claim team, a judge and a jury. A lengthy involvement with the model would not be practical for any of these audiences since they only have a very limited time available to be exposed to the model.

A lack of available time to understand what occurred in a system can be a reason for using a model. A simplified representation should be easier to understand than the real system. However, the restricted time exposure to the model in this situation, means that there are still difficulties with the audience fully understanding the model. Not only is the time exposure to the model a constraint, the plaintiff would be unlikely to expose the model to the defendant, a judge or a jury until the construction of the model was at an advanced stage. The plaintiff will want to present as robust a case as possible. This means that they would not wish to expose the model to other audiences until they were confident that the model was in a position to fully support the plaintiff's case. This reinforces the issue that many of the model's audiences are unable to be involved in the construction of the model.

6.6.4 Overall Behaviour of a Dynamic System

Although qualitative diagrams can provide significant assistance to enhance people's thinking about a system, attempting to appreciate the overall behaviour of a feedback structure over time from the qualitative diagrams is very difficult. Indeed, behaviour is often counter-intuitive to what an audience expects (Forrester 1972). Quantitative tools are therefore used to ascertain the behaviour of a system over time. However, attempting to get an audience to gain an appreciation of how the results of the quantitative simulation are arrived at is then a difficult process due to constraints such as *misconceptions of the impact of feedback* (Sterman 1989a, Sterman 1989b, Paich and Sterman 1993, Diehl and Sterman 1995). In a litigation environment this constraint is particularly problematic due to the difficulties discussed above i.e. due to the size of the models being considered and the lack of time the audience is exposed to the model. The complex dynamic behaviour of a system therefore contributes towards the potential for the model's lack of transparency.

When considering whether or not the SD models discussed in this thesis can be transparent to lay people, four particular difficulties have been highlighted. These are difficulties that have arisen with the models in the projects discussed in this thesis which it was felt had constrained the transparency of the models to some audiences. This has meant that the lawyers involved in the projects have concluded that they would be unlikely to use the models in a courtroom. The lawyers have been comfortable with the arguments contained in the qualitative models. However, they have had issues with the transparency of the SD models for the reasons discussed above. But, the lack of transparency has not been of such great concern as to prevent their use as part of negotiation periods between a plaintiff and a defendant.

One method of attempting to enable the various audiences to gain additional confidence in a SD model, without having to have a full appreciation of it themselves, is to get other experts in SD modelling to pass approval on the original modeller's work. This means that the model needs to be able to stand up to expert critique. This will be discussed in the next section.

6.7 Stand up to Expert Critique

The *experts* referred to in this section could have a variety of professional backgrounds. If a model is to withstand the rigour of a litigation process, then it has to be able to deal with the potential attack from experts in law or in modelling. If the model under question is based on a SD approach, then the modelling experts may be aficionados in SD. However, they may also be experts in project management, or general modelling. A litigation process is likely to involve experts from the defendant's side attempting to discredit any models presented by the plaintiff. In such a process the experts will be attempting to discover not only if an appropriate methodology is being used, but also if that methodology was carried out in a correct manner. The process of answering questions asked by such experts can be hindered when experts raise as many issues as they possibly can. These questions may also be of a nature that cannot be sensibly dealt with by the plaintiff. For example, they may be requesting hundreds of pages of information, or data of such detail that it would take weeks to gather. This was an approach that one SD modelling expert suggested when he was brought in to audit the Castle model. When asked how he would attack the model if he were to be hired by the defendant, he suggested that if he could not find something that he believed to be fundamentally incorrect with the model, then he would probably raise lots of small issues that may be very difficult to deal with. In this way he would attempt to gradually increase doubt in any audience's mind of the validity of the model.

An example of doubts that could be cast on the model lies with data collection. Gathering data on 'soft' behavioural variables such as *the impact of pressure on productivity* could always be questioned, since the data is likely to be subjective. With variables such as *pressure*, data may be of a judgmental nature. However, as long as the defendant admits the causal relationship exists then the use of this type of data represents a more informed position than having no data. Although the modeller cannot claim that the method used to model pressure is the only plausible way of doing so, if experts want to attempt to discredit this, then the onus is on them to prove it incorrect and to present data which is more convincing than that which has already been gathered.

As previously mentioned, the type of experts that may be used to either audit the model or attempt to discredit it could be general modelling experts, SD modelling experts or experts in project management. The following will consider specific issues relating to how each of these experts may critique the SD model.

6.7.1 General Modelling Expert

At this stage it is difficult to be able to come to any final conclusions on what an expert would conclude regarding the appropriateness of the use of the SD modelling approach since this is the aim of this thesis.

A general modelling expert is likely to want to ensure that a rigorous, logical modelling approach has been adopted. This would involve being convinced that each step in the modelling process has been carried out in a reasonable manner. For example, the collection of data had been carried out in a manner that considers any potential biases that may exist or there was evidence that a thorough validation process had been carried out.

In an attempt to discredit the model the expert may attempt to discover the weaker elements of the model. For example, the model outputs may be particularly sensitive to particular variables and the expert may want to focus on the process of collecting data for these variables with the aim of illustrating that the data is not trustworthy and hence the model results are not trustworthy.

Another area that the expert may pick up on is the objectives that the modellers will have based their work on. Eden (1989, p43) argues that "OR is client orientated not solution orientated". This suggests that the objective of an OR modelling process should be directed specifically at the needs of a client. If the modelling process carried out is *too* client-focused and does not have a sufficient *scientific* focus, it may be perceived that the model represents a particularly biased point of view, as only the plaintiff's views of what occurred on the project has been taken account of. Even if a strong client focus is not intended by the modeller it is likely to occur due to the lack of access to the defendant's views.

6.7.2 SD Modelling Expert

A SD modelling expert is likely to be interested in any of the topics which an expert in general modelling would be interested in. In addition to these, the expert would also focus on the attention that the modellers have placed on following a "traditional" SD modelling approach. At each stage of the modelling process, the expert would wish to ensure that publicised procedures on SD modelling had either been adhered to (for example the basic principles of SD laid down by Forrester in 1961) or that the modeller had good reasons for taking an alternative approach. In particular, the expert would want to see that the model had been thoroughly validated. Therefore, any of the validation tests that will be discussed in chapter 7 - *Replicating Reality Convincingly for the Modeller*, are also appropriate tests for a SD expert to expect the model to have been exposed to.

For the Castle project an expert modeller was brought in by the plaintiff to audit the SD model. The expert's explorations consisted of exposing the model to various SD validation tests that are detailed in the SD literature. Once the expert was familiarised with the model, he was asked to consider how he would mount a hostile attack against it. Areas that the expert highlighted were as follows:

- Testing the sensitivity of key behaviour loops around project pressure. The majority of the feedback loops in the Castle project focussed upon compression. However, actions taken to compress the project arose from the pressure that was felt to attempt to reduce the expected delays on the project. The nature of pressure means that any quantification is subjective and hence could be an area where the defendant will contest the data used. Also, since it is a key driver of the feedback in the model, it is likely that altering the value of pressure will have a large impact on the model output. This highlighted a particular need for sensitivity analysis during the modelling process. This will be discussed in chapter 7.

- In attacking a detailed SD model the expert suggested that he would search through the model in an attempt to find any type of formulation errors. In searching through the model, the expert would also be looking for points of ambiguity. By building up a set of such errors/ambiguities he would hope to cast doubt on the model output and hence the claim for compensation that it supported. This action highlights the need for the modelling process to be as rigorous as possible. For example, the structure of the model needs to be tidy and logical and the model documentation needs to be full and clear. Any ambiguities in the model may cast doubt on the model and hence prove detrimental to the litigation process.

The experience of the use of an SD expert on the Castle project reinforces the importance of following a rigorous modelling approach that includes the use of the validation tests that will be discussed in chapter 7.

6.7.3 Project Management Expert

Examples of the types of tools that have traditionally been used to analyse projects have been, for example, CPA, Gantt charts, the earned value technique and others stated in *The Project Management Body of Knowledge* (PMBOK) (Project Management Institute 1996). SD is also recognised by PMBOK as a project management tool. However, its uses have been given a narrow description as a method to enable the modelling of non-sequential activities. The experience of a project management expert is more likely to have been focused on the more traditional project management techniques.

Criticisms of the SD model by a project management expert may be based upon the operational detail that it does *not* capture. For example, the models constructed for the projects discussed in this thesis have not captured the detailed network of activities that will have existed for the projects. However, it has been widely recognised that the traditional project management techniques do not take account of 'soft' human elements of a project or the more strategic systemic effects (Morris and Hough 1987, Cooper 1994, Williams 1997). These elements are crucial in the analysis of D&D in a large complex project. Rodrigues (2000) gives a full account of the capabilities of the traditional techniques in the analysis of projects that fail and highlights the need for other tools when analysing why a project has overrun in both time and cost.

Due to the limitations of both SD and the traditional project management tools, both of these techniques have been used alongside one another to analyse projects. For example, in the Castle project consultants were hired to use CPA to explain the project overrun. It was hoped that the causes of the project overrun could be highlighted using two different forms of analysis. In this way the weaknesses of one approach would be covered by the other approach. Therefore, if a project management expert were to attack the weaknesses of SD to model D&D, the plaintiff would still have the results of the CPA approach to support their case. Rodrigues (2000) goes further than simply using SD and the traditional project management techniques alongside one another and suggests a way forward for combining the two techniques. However, previous discussions highlighted the unsuitability of this approach as a post-mortem analysis tool for litigation.

A project management expert may also prefer the use of the traditional project management technique when compared to SD due to their simplicity. The potential lack of transparency with SD models has already been discussed in this chapter. The more simplistic approach of the traditional methods is perhaps one reason why there is far more publicised reports of their success in claims when compared to the SD modelling approach (Wickwire and Smith 1974, Scott 1993, King and Brooks 1996, Cushman et al 1996). The expert may be inclined to argue for their use when attempting to meet the purpose:

- *Be Transparent to lay people.*

However, the modeller may argue that they are not capable of meeting any of the other modelling purposes as they do not attempt to model the 'soft' human effects which are an important part of the dynamics of D&D.

This section has explored some of the types of expert critique that a SD model may need to withstand during a litigation process. No expert critique can be fully anticipated, but one major area is the appropriateness of the SD modelling approach to actually model D&D for litigation. However, an opinion on this cannot be fully formed until the end of this thesis.

6.8 Summary of Chapter

"No model can claim absolute objectivity, for every model carries in it the modeler's worldview. Models are not true or false, but lie on a continuum of usefulness" (Barlas and Carpenter 1990, p157). When considering how *useful* a model is, the *purposes* of the model process need to be taken into account.

The purposes of modelling D&D for litigation were explored in this chapter. The appropriateness of SD in being capable of meeting these objectives was considered. The following is a summary of the conclusions drawn from these explorations:

- *Model exogenous triggers and their outcomes as D&D*

  The disruptive triggers that occurred in the Castle and Pirate projects could be grouped into three categories. The outcome of each of these groups of triggers was that managerial actions were taken in response to them. The modelling method used therefore needed to be able to model the triggers and managerial actions as feedback D&D. The SD modelling approach is intended to be used to model the feedback that arises from managerial actions that occur when exogenous triggers hit a system. It was therefore concluded that SD was capable of modelling exogenous triggers and their outcomes as D&D.

- *Model the paths of argument from an action to an eventual outcome*

  This criterion was also phrased as:

  - *Can SD establish a causal link between an action and an eventual outcome?*

  What is meant by a causal link in a litigation environment, the modelling of D&D and SD modelling was discussed. It was noted that not only does the quantitative SD model need to be able to model paths of arguments, but the model's audience also needs to be able to observe and understand the links in the model. This ties in with another of the objectives; the transparency of a model to lay people.

  Although the team at Strathclyde University intended their cyclical process to aid the transparency of the SD model, the qualitative models generally provide a clearer route to trace the paths of argument around concepts and arrows. The use of the qualitative cause map and, if required, an influence diagram with the quantitative SD model is the recommended approach to be able to model the causal links between an event and outcome. However, it should be noted that achieving a relatively transparent *quantified* link between the events and outcomes is very difficult when modelling D&D in a litigation environment.

- *Quantify the outcome of D&D*

  This criterion is asking whether or not SD can quantify the delays and cost overruns that may occur in a project due to the occurrence of D&D. It was noted that there are inadequacies in the SD modelling approach when attempting to capture the more detailed, operational issues of project management. This will cause difficulties when attempting to replicate project delays.

  A concern in attempting to simulate the cost overrun that actually occurred on the project lies with the lack of accuracy of the parameters used as inputs. This means that a modeller should provide a range of results as an indication of the range of additional labour hours incurred rather than a single cost result. Also, the D&D that occurs on one project may cause costs to accrue on other projects. However, these costs are normally omitted from the SD model. The reasons for the omission of these costs are due to the general difficulty in modelling them. Although these costs would be very difficult to quantify, they may still be argued to have resulted from the D&D that was experienced in the original project.

- *Replicating reality convincingly*

  There are a number of audiences that need to be convinced that the model replicates reality. The first audience for the model will be the modeller. The majority of the author's experience as a part of the projects discussed in this thesis involved convincing the modeller. Out of all the audiences, the modeller has the most rigorous process for the model to go through in order to be convinced that the model replicates reality convincingly. Due to this rigorous process and the importance of this audience, the issues associated with this process will be dealt with separately in chapter 7.

  To enable each of the other audiences to make a judgement about whether or not they believe the model to replicate reality convincingly, they each need to firstly gain a basic understanding of the mechanics of the model. This means that the next criterion i.e. *Be transparent to lay people* is vitally important to enable each of the audiences to gain a reasonable understanding of the model.

- *Be transparent to lay people*

  When considering whether or not the SD models discussed in this thesis can be transparent to lay people, the following difficulties were highlighted:

  - Size of the model.
  - Requirement of quantitative reasoning rather than qualitative reasoning.
  - Lack of time the audience is exposed to the SD model.
  - Overall behaviour of a dynamic system.

  Each of these difficulties constrains the level of transparency that particular audiences can reach with the model.

  A lack of transparency in the models for some of the audiences has meant that the lawyers involved in the projects discussed in this thesis have concluded that they would be unlikely to use the models in a courtroom. However, the lack of transparency has not been of such a great concern to prevent their use as part of a negotiation period between the plaintiff and defendant.

- *Stand up to expert critique*

  The *experts* referred to in this section could have a variety of professional backgrounds. Experts in general modelling, SD modelling and project management were discussed in this chapter. Any of these may be hired by the plaintiff to audit the model or hired by the defendant in an attempt to discredit the model. No expert critique can be fully anticipated. However, a modeller should attempt to be aware of how other experts may attempt to discredit the model and take these issues into consideration during the construction of the model.

  A major area under this criterion is the appropriateness of the methodology to actually model D&D for litigation. However, an opinion on this cannot be fully formed until this thesis has been completed.

The above summary highlights Barlas and Carpenter's (1990, p157) statement on models that they "...lie on a continuum of usefulness". SD can fulfil some of the above criteria better than others. For those where difficulties have been highlighted, these difficulties do not prevent SD being used to model D&D for litigation. However, they are weaknesses that the modeller needs to be aware of when considering how *useful* their model is in achieving the purposes of the modelling process.

The focus of this chapter was to explore the appropriateness of SD to aid the exploration of D&D for litigation through consideration of *the purposes of the modelling process*. In doing so, the criterion *Replicate Reality Convincingly* was not fully discussed for the first audience; the modeller. The modeller has the most rigorous process for the model to go through in order to be convinced that the model replicates reality convincingly. Also, the modeller is an important audience since if the model does not convince this audience, no other audience is likely to be exposed to the model. For these reasons, it was decided that the ability of SD to *replicate reality convincingly for the modeller* should be dealt with separately. This will therefore be discussed in the next chapter.

CHAPTER 7: REPLICATING REALITY CONVINCINGLY FOR THE MODELLER

7.1 Introduction

Section 6.5 considered the following purpose of modelling D&D for litigation:

- *A model should replicate reality convincingly.*

As part of this discussion, the different audiences for the model were identified as including the following groups of people:

- Modellers
- Plaintiff's claim team
- Modelling experts hired by the plaintiff
- Defendant and any modelling experts hired by the defendant
- Judge
- Jury

For each of the audiences beyond the modeller, section 6.5 discussed the needs of the audience for them to be convinced that the model adequately replicates reality.

This chapter will consider the first audience for the model; the modeller. The modeller needs to be convinced that the model replicates reality convincingly before any other audiences will be exposed to the model. When considering the replication of reality with respect to a SD model, there are specific confidence building tests that have been used on SD models over the last 40 years and are set down in the literature. These tests will be considered in this chapter.

Before exploring any of the confidence building tests, a modeller will be aware that any model is a simplified representation of reality. For this reason a model will never *fully* replicate reality. Instead, a modeller needs to ask whether or not it a*dequately* replicates reality. Before any judgement can be made on this, the manner in which the model simplifies reality needs to be considered.

7.2 How SD Simplifies Reality

There are three reasons why SD models are simplified representations of real systems:

(i) The aggregation of variables.

(ii) The modeller's choice of system boundary which will represent a finite number of variables.

(iii) The omission of noise.

These simplifications are in line with the purpose of most SD models as they are "basically strategy models... (their) purpose ... is *not to make precise quantitative predictions of the future*, but rather to indicate the trends of the key variables..."(Roberts et al 1983, p535, Italics added). An issue with using SD to model D&D for litigation is that the modelling purpose is different to many other SD models. Here SD is *not* being used as a strategy model, but as a model that can:

- *Replicate over time the hours due to D&D which were over-and-above those that were contracted, but were required to carry out the project.*

To help achieve this objective the quantitative results of key variables from the model *do* matter. This means that the issue of whether or not SD can replicate *sufficient detail* to be able to fulfil this objective needs to be addressed. The three causes of a SD model being a simplified representation of reality will therefore be considered next with respect to modelling D&D for litigation.

7.2.1 Level of Aggregation of Variables

In SD models, variables are normally aggregated as continuous variables. However, when attempting to replicate reality adequately, a modeller has to try to avoid the result that "an aggregated model will produce results which are completely different from a disaggregated equivalent" (Coyle 1977, p360). Therefore, the level of aggregation of the model is important. In particular, a modeller may have to consider whether or not certain discrete events are better modelled as discrete variables, rather than being modelled as aggregated, continuous variables. This would be a consideration when the modeller is concerned that, by aggregating the variables, the model will not adequately replicate reality.

Forrester comments on the inclusion of discrete events in a SD model by stating that "Discreteness of events is entirely compatible with the concept of information-feedback systems, but we must be on guard against unnecessarily cluttering our formulation with the detail of discrete events that only obscure the momentum and continuity exhibited by our industrial systems"(1961, p64). When constructing a SD model Forrester suggests that "As a starting point, the dynamics of the continuous-flow model are usually easier to understand and should be explored *before* complications of discontinuities and noise are included" (1961, p65, Italics added). However, "When a model has progressed to the point where ... there is reason to believe that discreteness has a significant influence on system behavior, discontinuous variables should then be explored to determine their effect on the model" (1961, p66).

In the case of modelling D&D for litigation, questioning the level of aggregation of a variable is important when attempting to meet the modelling purposes. This is due to the concern of being able to include sufficient detail to be able to meet the modelling purposes.

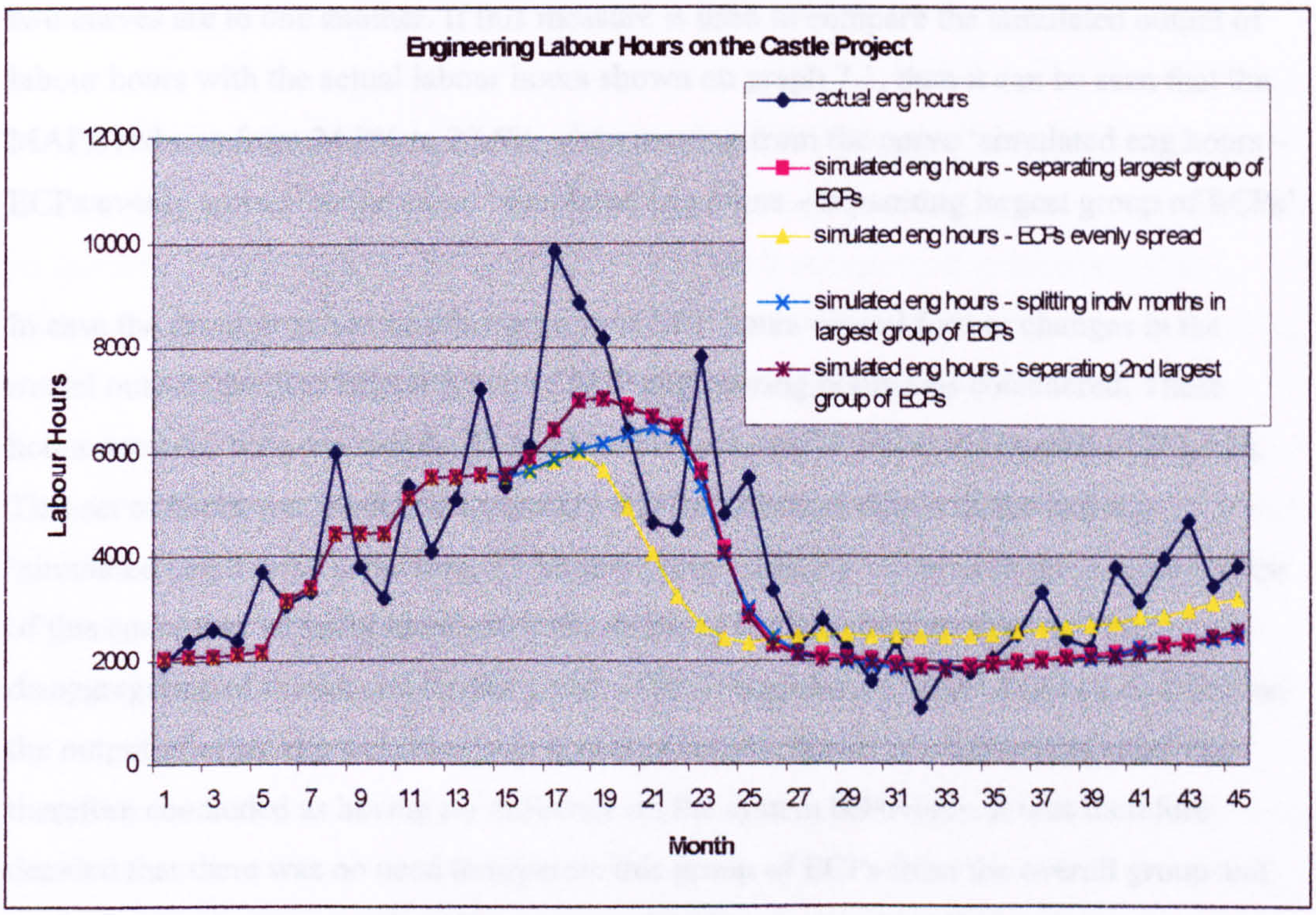
The modelling of exogenous variables in the Castle project is an example where the level of aggregation of the variables needed some consideration.

Exogenous Variables

In the Castle project model, contract changes made by the defendant during the life of the project were known as *amendments*. These were modelled as exogenous variables in the model. When considering the amendments that affected the engineering function, ECPs were one cause of the additional engineering hours that occurred during the life of the project.

ECPs occurred during a large proportion of the life of the project. The first assumption that was made when including ECPs in the model was that they occurred as a continuous stream of changes. The level of changes was then initially assumed to be at an average amount throughout the time period that the ECPs occurred. When such an assumption is used in the Castle simulation model the output, in terms of monthly labour hours, can be seen in graph 7.1 below.

The yellow curve 'simulated eng hours – ECPs evenly spread' represents the simulated engineering hours based on the assumption that the occurrence of ECPs was evenly spread during the period that they occurred. This time series can be compared against the actual engineering hours for the project represented by the dark blue 'actual eng hours' curve.



Graph 7.1 – The Modelling of ECPs in the Castle Project

When considering the data regarding ECPs, one period of the project; months 14 to 20, represented a period of time when a large percentage (85%) of the engineering hours that were budgeted for ECPs occurred. This represents such a large proportion of the hours, that it was believed that allowing for this group of hours separately from the remaining ECP hours may have a *significant effect on the system behaviour*. This group of hours was therefore modelled separately from the remaining hours. When modelling this group of ECP hours separately, the model output alters to the pink 'simulated eng hours – separating largest group of ECPs' curve on graph 7.1. The shape of the curve has changed significantly.

A measure that can be used to assess the overall closeness of the points on two curves is the mean absolute percent error (MAPE). This is a measure that will be used further in chapter 8 when discussing behaviour reproduction. The smaller the MAPE, the closer the points on two curves are to one another. If this measure is used to compare the simulated output of labour hours with the actual labour hours shown on graph 7.1, then it can be seen that the MAPE reduces from 24.8% to 22.5% when moving from the curve 'simulated eng hours – ECPs evenly spread' to the curve 'simulated eng hours – separating largest group of ECPs'.

In case the disaggregation of other groups of ECP hours caused further changes in the model output, the next largest group of ECP engineering hours was considered. These hours occurred between months 42 to 45 and represented 4.7% of the overall ECP hours. This set of hours was modelled separately and the output is shown as the maroon 'simulated eng hours - separating 2nd largest group of ECPs' curve in graph 7.1. The shape of this curve was virtually identical to the shape of the last simulated curve. The disaggregation of the second largest group of ECP engineering hours also had no effect on the output of other key variables (e.g. levels of rework, levels of productivity) and was therefore concluded as having no influence on the system behaviour. It was therefore decided that there was no need to separate this group of ECPs from the overall group and no further separating out of smaller groups of ECPs was considered necessary.

The result of separating groups of ECPs was that only the months 14 to 20 ECP hours would be separated in the model. The next question was whether or not it was worth further disaggregating the engineering hours within this period of time. However, the problem with doing so was that there was no data available on how the actual hours of additional work were introduced into the system during this period of time. The only available data gave an indication of dates for when ECP work had been planned to be carried out. However, in reality, the work could have been distributed in any way during the given dates. Bearing this in mind, the planned dates for the work were used to further split the work between the individual months of 14 to 20. The results from the simulation that reflect this change is shown as the light blue 'simulated eng hours - splitting indiv months in largest group of ECPs' curve on graph 7.1. This time series shows a changed behaviour in the model during the period months 14 to 20. However, the information on the planned dates used to distribute the hours of work during June to December 1995 was not considered to be sufficiently reliable to include it in the model. If it had been more reliable, then pulling out discrete ECP hours during months 14 to 20 would have been worthwhile since the model behaviour was altered by doing this. However, due to the unreliability of the data the previous pink 'simulated eng hours – separating largest group of ECPs' curve was taken to be the best match to the actual data.

The above shows that when considering the modelling of ECPs in the Castle project, the level of disaggregation was considered by disaggregating the available data until there was no further *significant effect on the system behaviour.* It is important to note that what distinguishes a *significant effect on the system behaviour* when modelling D&D in a litigation environment is different to what may be taken to be a *significant effect on the system behaviour* in other types of uses of SD. Many SD models have been used for the purpose of understanding and improving policy making in organisations. In such cases, the difference between the yellow and pink time series in graph 7.1 above i.e. splitting off the largest group of ECPs, or not, may not have been concluded as having a *significant effect on the system behaviour.* The overall shapes of the two curves are similar, with the peak of the pink curve lasting a little longer. However, the purposes of modelling D&D for litigation are different. When considering system behaviour both the shape *and* quantitative results of key variables need to be taken into consideration.

Amendments have been discussed as an example of an area of the Castle project model where the level of aggregation needed some consideration. The production function is another area of the Castle model where the level of aggregation was questioned. However, this is not unique to the Castle model. Similar issues have arisen in other projects that the author has been associated with.

Production

Both of the projects considered in this thesis involve multiple items being manufactured on a production line. In these cases, the production line would be divided into a number of stages and each item would complete one stage before being moved onto the next stage for production work to progress on the item. Therefore, in reality, the production line consists of work being carried out on the different stages, then at some discrete point in time work stops and all the items get moved onto the next stage before production work commences again.

The issues of using SD to model discrete events in a production environment are not new. Wolstenholme and Coyle came across such an issue when modelling mining in coal mines (Wolstenholme and Coyle 1980, Coyle 1985). This work considered the modelling of machinery breakdown. The model was required to reflect discrete changes in states from production to breakdown and back to production. Coyle (1985, p310) noted that "... although the example is set in a coal-mining context, similar phenomena occur in many other production processes".

The production functions in the projects discussed in this thesis could have been modelled in a continuous manner. To do this, the models would not have been concerned with the movement of items between stages, but with the overall work required on an item. The work would therefore be modelled as a continuous flow over time. When developing the models discussed in this thesis, the modellers considered this type of modelling approach. However, some of the reasons for not taking this approach are as follows:

- To be able to model some of the disruptions that affect the production function, the individual stages needed to be replicated. For example in the Pirate project, one of the reasons it was believed that productivity had been lower than expected in the production function was due to workers having to carry out work in a stage of production that was not the stage where the work had been originally planned to be carried out. D&D had caused delays in the project that meant that when an item was planned to be moved from one stage to the next, the work on the first stage may not have been completed. However, in an attempt to try and keep to the schedule of the project, the items were moved onto the next stage before completing all the work on the previous stage. This meant that this work had to be carried out during a stage where it had not been planned to take place. This was when productivity levels in the production function were affected. The ramifications of out-of-stage work could only be simulated if the various stages of the production function were individually modelled. Also, to be able to know when an item should be in one stage or another, the discrete movement of items from one stage to another was required to adequately replicate reality.

- The matching of historical data is a useful exercise to carry out during the development of the SD model to help the modeller discover areas of the model where improvements are required (this is discussed in detail in chapter 8). Therefore, to be able to do this, reliable historical data is required. There are many different variables that can be used for this test. For the production function, relevant historical data may include the dates when individual items commence and complete production. However, if the modeller discovers that the model is not matching this data, to be able to focus on where the improvements are required in the model, the modeller may want to then go on to attempt to match the dates where items move from one stage to another. Therefore, in order to be able to do so, the model has to actually simulate the movement of an item from one stage to the next.

- When simulating what occurred on a project, the intent is not only to replicate the additional labour hours that occurred on the project due to D&D but also to replicate any time delay that occurred on the project. This is particularly important if the defendant is enforcing liquidated damages on the plaintiff due to a delay in the completion of the project. The plaintiff would then want to demonstrate that the delay was due to D&D triggered by the defendant so that they are not held liable for the liquidated damages. Therefore, adequately replicating the dates that items commence and complete production becomes particularly important. To ensure the accuracy of these dates, the modeller may also need to model the delay between individual production stages. This would create a need to simulate the movement of an item from one stage to the next.

A modeller using SD to model D&D for litigation is going to be concerned with the effect of discreteness on both the behaviour of the system and the numerical output of the system. It may be the case that a modeller will use more discrete variables for this modelling compared to other types of modelling using SD. Indeed, when auditing the Castle model, a SD modelling expert commented on his concern about the use of SD in this way. However, the appropriate level of discreteness in a model should always be considered with the following statement in mind: "When a model has progressed to the point where ... there is reason to believe that discreteness has a significant influence on system behavior, discontinuous variables should then be explored to determine their effect on the model"

(Forrester 1961, p66). In the case of modelling D&D for litigation, the modeller will be considering numerical outputs as a part of system behaviour.

In conclusion to this section, the level of aggregation should not be an issue that excludes SD from being able to model D&D for litigation. A modeller should be able to include sufficient detail in the model to replicate reality convincingly. The level of aggregation should be taken to the point where any more detail would not have a significant influence on the system behaviour. However, this may mean that the level of detail is greater than is normally the case in SD models since system behaviour in this case needs to consider both the shape and quantitative results of key variables.

7.2.2 System Boundary

The choice of the system boundary is an element of the model that the modeller has control over. However, the modeller needs to get a balance between including all relevant elements of the real system and the time and effort required to construct the model. The literature advises that any components that are necessary to create the *behaviour* being explored should be included (Forrester 1968a, Forrester and Senge 1980, Richardson and Pugh 1981, Coyle 1977). However, previous discussions have noted that the purposes of modelling D&D for litigation mean that there is also a focus on quantitative results. Therefore, the concepts included in the boundary will need to go beyond those that are required to replicate the overall dynamic behaviour.

The boundary adequacy test will be discussed later in this chapter. This will explore the elements that should be included in a system boundary when modelling D&D for litigation.

Self-Inflicted D&D

An area of the system that is often omitted when modelling D&D for litigation is self-inflicted D&D. This would represent any D&D that was believed to have been triggered by the plaintiff. Since the model is constructed on behalf of the plaintiff as part of a claim against the defendant, identifying and modelling any errors made by the plaintiff is not part of the plaintiff's modelling objectives. In all the cases discussed in this thesis, the simulated labour hours have been less than the actual hours that accrued on the project. The difference between the simulated and actual hours is seen to be the labour hours which were lacking a causal hypothesis to explain their existence. The plaintiff may be willing to accept this difference as a "give-away" i.e. takes responsibility for the cause of these additional hours. This may be seen as a useful approach during any form of negotiation. If the plaintiff is willing to accept some responsibility for the hours overrun on the project, the defendant may be more likely to also accept some responsibility.

Of course, it is often very difficult for the modellers to differentiate what is and what is not self-inflicted. Indeed, it may not be until the lawyers argue the case in front of a judge that it is finally concluded what is and what is not self-inflicted.

By involving many different staff who worked on the project, the interviewing process attempts to elicit all the triggers of D&D. Therefore, the omission of D&D triggers from a model has not generally been an issue. Instead, the contention generally lies with *who* was originally responsible for the D&D trigger. For example, the plaintiff may make a poor decision that causes an overrun in hours in the project. However, they may argue that this poor decision was due to the pressure that the defendant had placed the project managers under in an attempt to speed up the project to meet certain deadlines. In this case, the D&D trigger would be included in the SD model. It is only those triggers for which the plaintiff cannot push the blame back on to the defendant that would not be included.

By excluding any self-inflicted D&D from the model it is likely that the simulated and actual output from key variables will not match in either shape or overall total hours. Chapter 8 discusses the matching of simulated and actual data as part of a confidence building test. The issue of the exclusion of self-inflicted D&D makes this process difficult. Of course, there is no reason why any self-inflicted elements should not be modelled. However, for the projects discussed in this thesis, this has not been an approach that the plaintiff's claim team has wished to follow. In particular, the plaintiff's lawyers have argued that all additional labour hours should be claimed for, with the consideration that some hours may be "given-away" if they believe this to be an appropriate step to take during negotiation.

In conclusion to this section, the modeller has control over the concepts that should be included in the system boundary to enable the modelling purposes to be met. Some time and effort may be required in the consideration of what should and should not be included to enable *sufficient* detail to be included in the model. This section has highlighted the specific case of the exclusion of self-inflicted D&D from the model. This may lead to a difficulty in matching actual data as part of a confidence building test. This issue will be returned to in chapter 8.

7.2.3 Noise

Forrester describes noise as "... that part of the decision flow for which we have no satisfactory causal hypothesis" (1961, p430). He explains the omission of noise from a SD model by noting that "...we cannot assume a *perfect* model in which *every* relationship is known *exactly*. Therefore, we are committed to models in which *every* decision function has, at least in principle, a noise or uncertainty component. By definition, the exact time pattern of this noise is unknown, and we have not discovered its generating causes" (1961, p124).

Since noise represents events in the real system that a modeller is unable to account for in the model, when considering the modelling of D&D for litigation, examples of areas where noise may exist are as follows:

- Productivity levels of individual workers will vary. For example, a worker may perform better or worse on an individual day due to factors that affect them individually.
- Equipment performance may vary, which again will affect productivity levels on a specific day.
- The time it takes to be able to hire a new worker to the project will vary. This will be influenced by various factors. One factor is that workers' previous employment will enforce differing notice periods.

It should be noted that some noise could be included if a modeller believes this to be important. Sterman (2000, p914) refers to this by stating that "As a general modeling strategy you should first understand the dynamics of your model without noise...Once you understand how and why the system responds as it does you can consider how more realistic inputs such as noise affect the dynamics". However, due to the purposes of modelling D&D for litigation, we also need to focus on how the quantitative results are also affected by the introduction of noise.

As noise is, by definition, unaccountable, the only way in which it can be taken account of in a model is by modelling a variable with a random element. The inclusion of noise to continuous variables in the Castle model did not lead to any significant change in behaviour of the key variables in the system. For example:

- *Engineering productivity* – the equation that calculated the productivity level was multiplied by a normal distribution with mean 1, standard deviation 0.05. The simulation was re-run a number of times. The shape produced by the labour hours output was virtually unchanged. The number of engineering labour hours decreased by a maximum of 0.07%.
- *Delay in hiring engineers* - the equation that calculated the productivity level was multiplied by a normal distribution with mean 1 and standard deviation 0.1. The simulation was re-run a number of times. Again, the shape produced by the labour hours output was virtually unchanged. The number of engineering labour hours increased by a maximum of 0.02%

The results are not unsurprising. The inclusion of a random distribution to a continuous variable is unlikely to lead to a significant change in behaviour of key variables in the system due to the small delta time (i.e. the interval of time between calculations) used when simulating the model. However, of more interest, is the inclusion of noise to variables that are more discrete by nature, since this *is* likely to have a significant effect on the system results. An example of a situation where noise was used in the modelling of discrete variables occurred in the Pirate project.

During the Pirate project, a claim was made that due to the lack of available labour, the management of the project were unable to hire labour at any point in time. Also, management attention was diverted to "fire-fighting" on the project to such an extent that the consideration of the need to hire more labour was sometimes delayed. As a result, the decision to hire labour was only made at monthly intervals. To test the impact that this situation had on the project, a SD model was constructed. The hiring decisions in this model occurred at discrete intervals. Initially, the model differed to reality since the information that the hiring decisions were based on in the model was accurate and immediately available. This meant that the model was more accurate in its ability to forecast the number of staff required when compared to the real project. To reflect the fact that management were never completely able to accurately forecast the number of staff required, noise was introduced into the hiring decision. Once this was added, the model was run a number of times.

The results of the model showed a huge range of results in the outcome of the project. The results indicated that even a small error in the forecasting of labour could result in a large overrun in the project. Due to the complex ramifications of D&D during the project, the accurate forecasting of labour was very difficult. Indeed, the conclusion was that management of the project had performed well, since inaccurate labour forecasts could have resulted in far higher overruns on the project.

Based on the author's experience of the Castle and Pirate projects, noise has been seen to have a significant impact on the outcome of a model if it forms part of a discrete variable. The actual effect that noise will have on the output of a model should be considered alongside sensitivity analysis. Uncertainties around the values of some parameters mean that instead of producing one result for the claim process a modeller is more likely to put forward a range of results that would have been produced by a sensitivity analysis. Therefore, when considering the effect that noise has on the system, the focus should be on the effect it has on the range of results produced by the sensitivity analysis. Sensitivity analysis will be discussed later in this chapter.

The aim of this chapter is to investigate what is required to enable a modeller to conclude whether or not a SD model of D&D for litigation is a convincing replication of reality. This section has explored the reasons for the SD model being a simplified representation of a real system. With this understanding, a modeller can now consider possible confidence building tests that will aid the process of deciding whether or not they perceive the model to be an adequate representation of reality. The SD literature discusses confidence building tests that have been used for this purpose. These tests will therefore be discussed in the next section.

7.3 Confidence Building Tests

The validation of a model is an important step in any model building process (for example Ackoff and Sasieni 1968, Rivett 1972, Mitchell 1993, Pidd 1996). This is no exception in the SD modelling process with a number of authors discussing possible sets of validation tests that can be used on SD models (this will be returned to later in this section).

It has been argued that "SD models are ... judged by how much more "insight" they generate on the nature of the system being modelled, instead of how "valid" they are" (Dash 1994, p93). Generating insight into how certain triggers could have caused an overrun in a project is certainly of interest to the following modelling purpose:

- *Demonstrate that a part of the time and cost overruns were caused by D&D through particular triggers of D&D.*

If a SD model can help an audience gain an insight into how the initial D&D triggers could create the overall time and cost overruns observed on a project, then this would be a useful outcome. However, due to the specific modelling purposes when modelling D&D for litigation, the validity of the model is also important. The model will be used as part of a case that aims to convince an audience that the plaintiff should be compensated for a project which has overrun in both time and cost. This means that the validity of the model is going to be under detailed scrutiny in, for example, a court environment. Any doubt that can be placed on the model may call the entire credibility of the model into question. A modeller should therefore be concerned with the validity of the model.

Sterman (2000, p846) highlights that since a model is a simplified representation of reality it can never be *validated*. Therefore building *confidence* in a model is more appropriate. In chapter 6 the nature of *confidence* in the SD model was discussed in terms of the different audiences for the model. This chapter is concerned with *the modeller having sufficient confidence* in the model that he believes it to be an adequate replication of reality to the extent that the modelling purposes can be met.

Common tests detailed in the general SD literature, which are used to build confidence in a SD model, will now be discussed.

7.3.1 Confidence Building Tests in the General SD Literature

The first fully comprehensive list of tests to build confidence in SD models was set down by Forrester and Senge (1980). Other authors have added to these tests. For example, Barlas (1989, 1996) introduced a set of pattern oriented measures appropriate for behaviour pattern evaluation and Sterman (1984, 2000) introduced summary statistics to evaluate the historical fit of SD models. However, both of these works have concentrated on the advancement of statistical tests in building confidence that a SD model adequately reproduces historical data. The other tests detailed in Forrester and Senge (1980), in particular those focussing on building confidence in the structure of the model, have not been replaced. Coyle and Exelby (2000) mention a full set of 75 tests when they discuss the nature of SD model tests in a consultancy firm. However, due to the commercial value and confidentiality of these tests, they remain unpublished.

Forrester and Senge (1980) split their confidence building tests into three sets:

(i) Tests of model structure
(ii) Tests of model behaviour (where the term "behaviour" represents patterns or shapes of graphs over time)
(iii) Tests of policy implications

When considering these tests in relation to the models discussed in this thesis, the third set of tests is not generally considered, since the model is not built to consider policy changes. Therefore, tests to discover how the real system reacts to policy changes are not appropriate.

Section 7.3.2 – 7.3.4 will examine each of the tests set down by Forrester and Senge and consider their appropriateness when attempting to build confidence in a SD model which models D&D for litigation. As noted by Forrester (1961, p115, Italics added) "The validity (or significance) of a model should be judged by its *suitability for a particular purpose.* A model is sound and defendable if it accomplishes what is expected of it ... validity, as an abstract concept divorced from purpose, has no useful meaning". The importance of purpose when judging validity holds for any form of modelling (Rivett 1972, Mitchell 1993, Pidd 1996).

The modelling purposes when modelling D&D for litigation are therefore of great importance when considering which validity tests should be considered appropriate and will be considered throughout the next three sections.

7.3.2 Tests of Model Structure

1. Structure Verification

The structure of the model should adequately represent the structure of the real system. The modelling process used in the projects discussed in this thesis aids the verification of the model structure. Eden (1994) describes this process of moving from cause map to influence diagram to SD model. Once the structure of the system has been initially captured by the cause map, the structure may then be replicated by an influence diagram and then a SD model. The process is cyclical, with each of the models helping with the verification of the other models. Since the process helps with transparency in the model building for both the client and the modellers (Eden 1994), a change in any of the three models will automatically be fed to the other two models. Any potential contradictions that may then arise between models can be immediately resolved. Therefore, the process of verifying the structure of the SD model can be checked through any of these models.

For any SD model that aims to replicate reality, this is an important model test to carry out.

Understanding the details of the engineering processes in the Castle project

The author does not have any formal background in engineering or manufacturing and had only been involved in one engineering project prior to the work carried out for this thesis. Therefore, for the first project the author was involved in for this thesis (the Pirate project) an amount of effort was required to gain an understanding of the systems and processes involved in an engineering and manufacturing project. When the author began working on the second project (the Castle project), experience from the Pirate project assisted in gaining an understanding of the processes involved. However, the Castle project still contained very individual processes that were different to those used on the Pirate project. The individuality of projects has also been highlighted in the author's involvement in the design and construction of a project management simulation. This simulation was created as a senior management training tool to be used with one organisation. However, during the design of the simulation, discussions with various managers in the organisation highlighted the fact that very different terminology, documents and processes are used in different parts of the organisation and therefore on different projects. It is likely that no matter how many projects a modeller works on, each individual project will contain specific processes that will be new to the modeller. Whilst constructing models for the projects discussed in this thesis, experience has shown that it is important to ensure that the modeller has a sound understanding of the various processes involved in the project. A misunderstanding of the project can lead to a misrepresented model structure.

To aid the verification of the structure of the model surrounding design changes in the Castle model, a visual model was constructed to demonstrate how all the documents and processes related to one another. This was constructed by the modeller from an interview with an engineer on the client's claim team. The software used to create this model was Powersim[3]. Although the Powersim software is not intended for this purpose, it was used at the time as a purely visual aid to illustrate the flow of processes. The model can be seen in figure 7.1. This diagram was used to help verify the model structure surrounding, for example, configuration change requests (CCRs) received by the plaintiff which were requests for changes to be made to a number of drawings. The model structure for this element of the Castle model is represented by the cause map shown in figure 7.2.

Although the full detail of figure 7.1 was not required for the SD model, a full appreciation of the various processes was needed to verify that the more aggregated version represented in the SD model did correctly capture the nature of the project. Such level of detail is normally not required for all parts of a project. In this example, the detail was required to enable an appropriate level of understanding of the complex structure of the real system.

[3] Powersim is a proprietary product of Powersim AS, Norway.

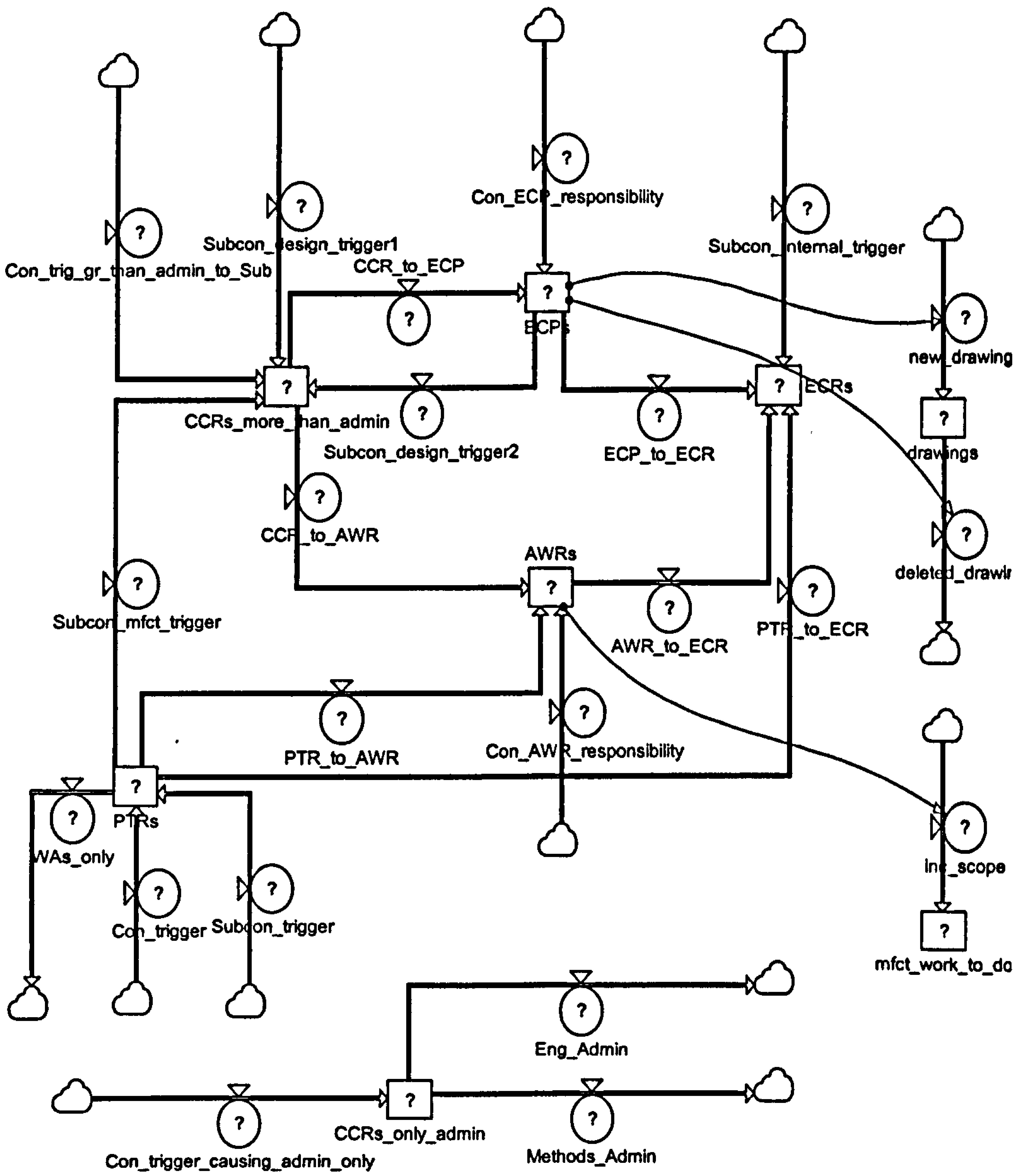


Figure 7.1 – Documents and Processes Involved in Changing a Design in the Castle Project

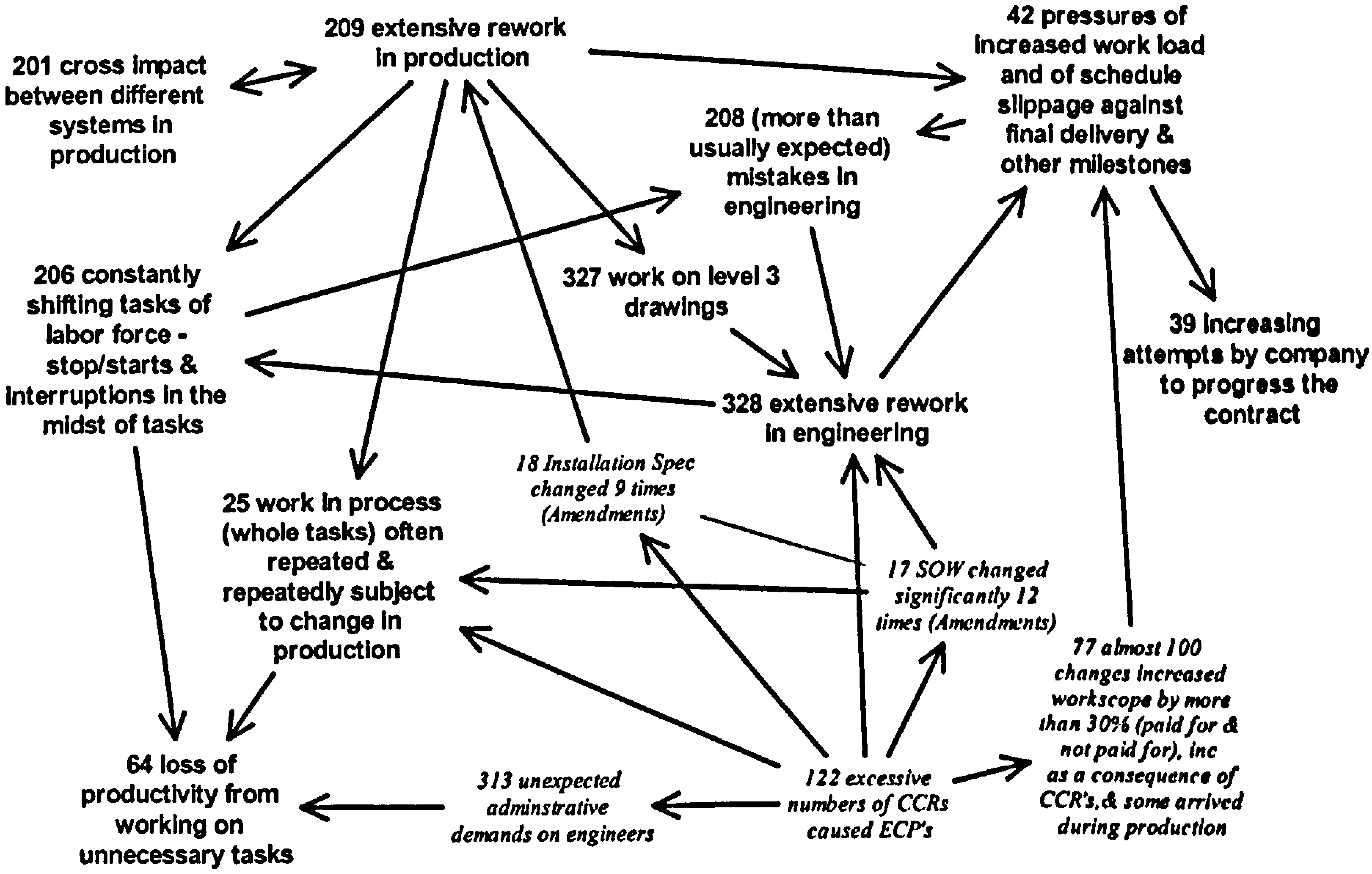


Figure 7.2 - The Effects of an Excessive Number of CCRs in the Castle Project

2. *Parameter Verification*

In the models discussed in this thesis, parameters were based on both numerical data and judgmental estimation. It is important that all parameters have clear, real life meaning so that a modeller is clear about the data or judgements required to quantify the parameters. Also, throughout the litigation process, there is always a possibility that an expert may audit the model. This expert (often hired by the defendant) may be an expert in project management, in SD modelling, or may be an expert in more general modelling fields. In any of these cases, the more obvious the meanings of any of the parameters, the easier the process of validating the parameters by an expert and therefore the less criticism the expert is likely to place on the model.

This is an important model test and in the case of modelling D&D for litigation can aid the model standing up to expert critique.

Classification of Parameters

The process of parameter verification will involve ensuring that any parameters correspond both conceptually and numerically to the real system. However, difficulties can exist in this process. For example, conceptual difficulty may arise when different people involved in the modelling process provide different definitions for the same variable contained in the model. This means that the real life meaning of the variable is not clear. Numerical difficulties can exist when data either does not exist or is difficult to collect.

When carrying out parameter verification, any parameter can be placed on a scale of how easy or difficult it is to verify it both conceptually and numerically. Such a scale is represented in table 7.1. The percentages given in each of the quadrants of this table represent the percentage of variables in the Castle model requiring parameter estimation that were categorised by each of the conceptual/numerical labels. An example variable is also given in each quadrant. In reality, any variable will lie on a *scale* of ease/difficulty. However, for ease of representation, the parameters are split into four quadrants. The four quadrants as follows:

1. *Numerically Easy and Conceptually Easy* – this category represents variables where straightforward information can be requested and easily gained from the client.
2. *Numerically Difficult and Conceptually Easy* – this category represents a situation where, for example, the client does not record the information requested, however it has been relatively straightforward to explain the data required to those who have worked on the project and thereby gain judgmental data.
3. *Conceptually Difficult and Numerically Easy* – no parameters from the Castle project lie in this quadrant.
4. *Conceptually Difficult and Numerically Difficult* – this category can involve a lengthy process to gain useful information. For example, on the Castle project a group workshop was carried out to estimate the relative productivity levels of workers during the project. The session began by spending time defining exactly what was meant by productivity, before continuing onto gaining judgement on numerical estimations.

	Numerically Easy	Numerically Difficult
Conceptually Easy	40% Number of original Drawings	38% Average number of engineering hours per reworked drawing
Conceptually Difficult	0%	22% Production cross-impact factor

Table 7.1 –Classifications of the Parameters in the Castle SD Model

Table 7.1 shows that for 78% of the parameters in the SD model for the Castle project, data gathering did not involve issues of conceptual difficulty. However, although this means that only 22% of the parameters involved difficulties with the conceptual correspondence of the parameter to the real system, this still represented 50 parameters. For each of these parameters, the gathering of data involved lengthy discussions regarding the meaning of the parameter.

From table 7.1, 60% of the parameters from the SD model of the Castle project involved difficulties in gaining numerical data. This was due to the data not being recorded by the organisation during the project. If the data could not be gathered during the litigation process, then the judgement of those on the project was required instead. A number of views were gathered from different people from the project to attempt to gain a wide perspective of views on the parameter. This process means that the verification of parameters with numerical difficulties is not straightforward. Verification would be based upon peoples' perceptions of the real system and their judgement about the particular parameter.

Sixty percent of parameters with numerical difficulties creates the need for a lot of time and effort in attempting to gain data for a SD model. However, the author is aware of SD models for litigation cases where the percentage has been far higher. This percentage is dependent upon how much data the organisation records during a project. If an organisation records the minimum amount they need to carry out the project, then if modelling is required for a litigation process, the percentage of parameters with numerical difficulties may be extremely high. It is only when staff have been involved in a litigation process that they appreciate the extent of data that is needed. Through discussions with managers that have been involved in litigation cases, the author knows of circumstances where managers have changed the data that is recorded during a project due to their experience in previous litigation processes.

Parameter verification is an important test during the construction of SD models. However, conceptual and numerical difficulties mean that the process is not straightforward and may require a large amount of time and effort to carry it out.

3. Extreme Conditions

This model test involves setting input variables or policies to extreme values. Forrester and Senge (1980, p214) state two reasons for carrying out this test; "...for discovering flaws in model structure...for example, extreme conditions aid in identifying nonlinearities and asymptotes which should be incorporated into model structure" and "...to enhance the usefulness of a model for analysing policies that may force a system to operate outside historical regions of behavior".

In both of the projects discussed in this thesis, the modellers were brought in to begin modelling D&D before each of the projects were completed. This meant that at the beginning of the modelling process, historical data only existed for a part of the project that the modellers were replicating. However, as the main modelling effort is normally spread over an extended period of time (for example, the main modelling effort lasted 24 months for the Pirate project and 30 months for the Castle project), the project may be complete by the end of the modelling process. This means that although historical data did not exist for the entire life of the project at the beginning of the modelling process, it would by the end of the process. Therefore the objectives of the modelling process mean that the model will only use historical data.

The only model runs that are anticipated to take the model beyond the actual or budgeted runs are when a limited number of D&D triggers are considered. For example, a model may be used to support the argument for ten different types of D&D triggers. However, to present an idea of the costs associated with each D&D trigger, the model is normally run introducing one D&D trigger at a time. This is also important since during a court case, the judge may only allow the plaintiff to claim for a sub-set of the D&D triggers i.e. those that the judge perceives the defendant to be responsible for. In this case, the plaintiff would wish to attach a cost to the specific sub-set of triggers. Each of these cases will mean that the model includes a set of D&D triggers that did not represent an historic situation. However, it is unlikely that the parameter values for each of the triggers will be taken outside historic ranges. The budgeted run will represent the lowest values for each of the D&D triggers, whereas the actual run will represent the highest values for each of the triggers. It is unlikely that a circumstance would arise where the plaintiff would either wish to run the model with less than the budgeted inputs or that they would be considering being compensated for more or larger triggers than those that actually occurred. Therefore, each of the values of the triggers will be within historic ranges.

However, the modeller needs to check that any functions used to represent variables do not give peculiar output for any input that lies between the lower and upper bounds. Therefore, running the model with different input variables between plausible upper and lower values is important. However, this is not the same as applying an extreme conditions test. The actual use of such a test would mean putting resources into taking the model beyond the purposes of modelling D&D for litigation.

Therefore, the extreme conditions test would not be a test which would aid a model to meet the purposes of modelling D&D for litigation.

4. Boundary Adequacy

Forrester and Senge (1980) define both a structure and behaviour focussed boundary adequacy test. The first aims to ensure that the model includes all the structure relevant to the model's purpose. The second extends this "to include analysis of model behaviour" (p222). It would be natural to consider both of these tests together during the modelling process, therefore the two tests will be considered together in this section.

The cause maps used as part of the modelling process for the projects discussed in this thesis can be used to help carry out this test. When modelling D&D, one of the modelling purposes is as follows:

- *Demonstrate that a part of the time and cost overruns were caused by D&D and identify the triggers of D&D.*

This means that the system boundary would need to include all the triggers of D&D and the ramifications of these triggers on the system. The cause map can be used to gain agreement between the modellers and the plaintiff that they are in agreement that all the triggers of D&D and their ramifications are included in the model. Also, as discussed in section 7.2.2, the modeller needs to ensure that the triggers of D&D are modelled at a level of aggregation that the lawyers are comfortable to use when arguing the plaintiff's case.

This test would be continually referred to throughout the modelling process. As new information is gained, it would need to be incorporated into the cause map and this would automatically be fed down to the SD model, so that if the boundary required alteration, this would be done. When modelling D&D for litigation, ensuring that the system captures the relevant structure to enable the behaviour of the real system to be created is important. However, the system also needs to replicate the numerical output of the real system and this needs to be considered when defining the system boundary.

This model test is important to enable the SD model to adequately replicate reality.

The Importance of Managerial Actions

In section 3.4 it was noted that for both the Castle and Pirate projects the majority of the feedback loops that were identified occurred due to a need to compress the project. This would suggest that when considering the boundary adequacy test when using SD to model D&D, the ramifications of the D&D triggers include all types of managerial actions that were used during the life of the project in response to the need to compress the project. This means that it is important to get clear information on the actions that were used and the ramifications they were believed to have on all parts of the project. Managers therefore have to consider both the positive and negative effects of managerial action (Cooper 1994, Howick and Eden 2001). However, exploring the negative effects of managerial actions can be difficult. Managers may not have fully considered the side-effects that are caused due to the decisions that they take.

For example, figure 7.3 below demonstrates the causes and effects of production workarounds in the Castle project. Getting managers to appreciate the full consequences of such an action can be difficult. Figure 7.3 illustrates how feedback loops are created due to the use of production workarounds. The feedback loops create a dynamic effect that is difficult for a manager (or anyone else) to fully appreciate. Illustrating how such loops are created to managers will increase their understanding of the ramifications of the actions they take and can therefore have an effect on the way in which they manage future projects.

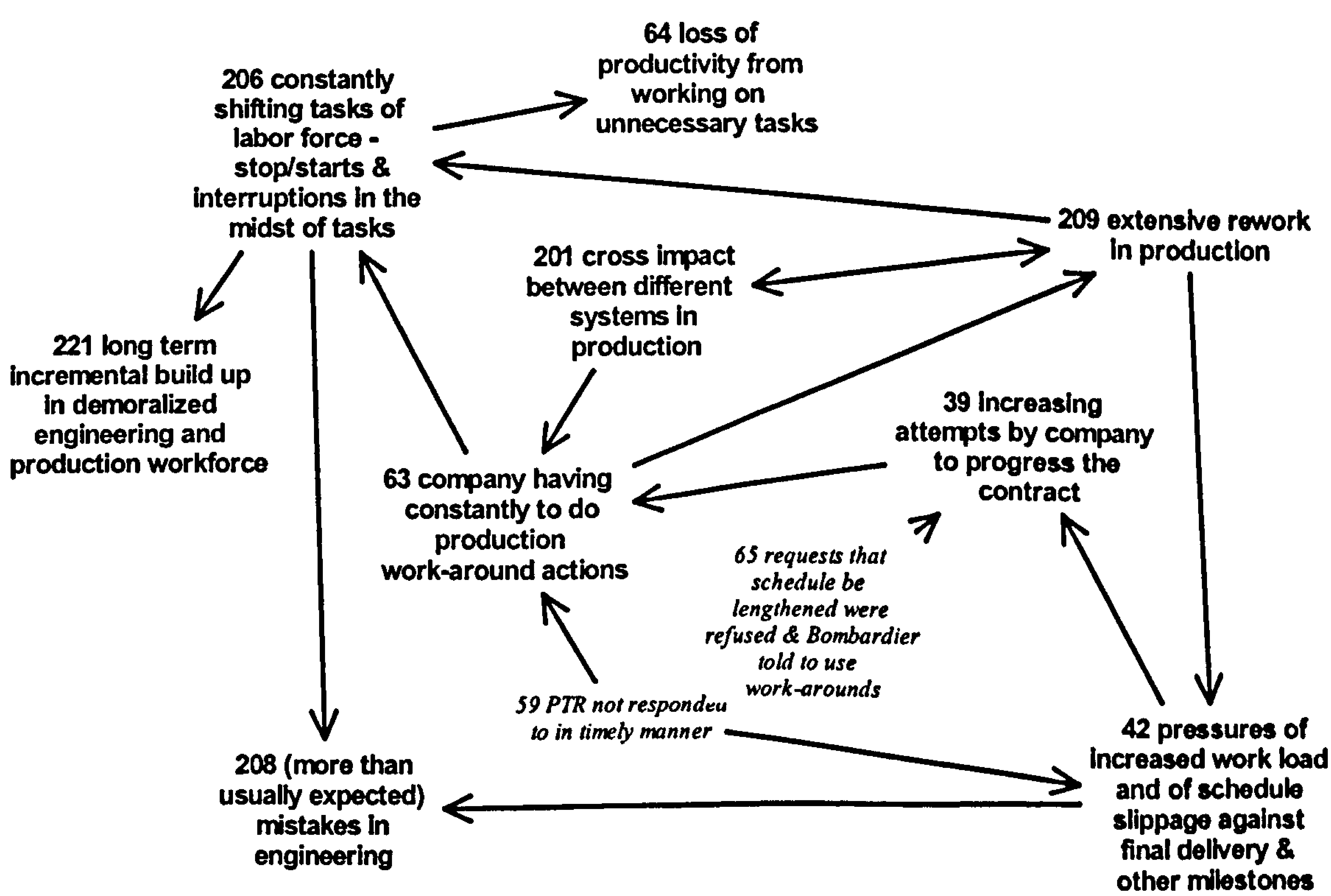


Figure 7.3 – Causes and Effects of Production Workarounds in the Castle Project

5. Dimensional Consistency

This is an important test for *any* SD model and should, in particular, be considered in conjunction with parameter verification whilst constructing the model. Good documentation of each variable, including its dimensions, should always be carried out when constructing a model, especially when the model is as large and complex as those that are discussed in this thesis.

7.3.3 Tests of Model Behaviour

1. Behaviour Prediction

This model test examines whether or not the patterns of behaviour generated by a model *beyond* the period which historical data exists are reasonable. As discussed in the extreme conditions test in section 7.3.2, there are two circumstances where the simulation model may be requested to produce non-historic patterns of behaviour:

- The simulation of a sub-set of the D&D triggers.
- If model construction were completed before the completion of the project. However, in this case, no future disruptive triggers would be included. The model would be simulated to the end of the project based upon the disruptive triggers that had actually occurred. This would be the only circumstance where a modeller would need to consider the *future* behaviour of the model.

Forrester and Senge (1980, p219-220) comment on two types of tests; the event-prediction and pattern-prediction tests. For these tests a modeller would wish to consider whether or not any events and patterns of behaviour created by the SD model, when running it beyond the period which historical data is available, are reasonable. Judgement would then be based on a consideration of the feedback structure of the SD model. Not only could this test be used to consider future behaviour, but also non-historic behaviour created by a different set of D&D triggers. This model test is therefore relevant for either of the situations where the simulation model may be requested to produce non-historic patterns of behaviour.

Forecasting the effects of embedded D&D

The objective of the SD model for the Castle project was to cover the period up until the re-scope occurred on the project. Since the modellers were hired by the plaintiff after the re-scope had occurred, the modellers were only required to construct a model to reproduce *past* behaviour. Therefore the modelling of *future* behaviour was not required during the Castle project.

However, the prediction of *future* behaviour was required during model construction for the Pirate project. For this project a large amount of the modelling effort was carried out before the project had completed. Therefore, during this time, the figures that were used to represent the overrun in labour hours included an element that forecasted the overrun to the end of the project. This forecasted overrun was calculated separately by both the plaintiff and the modellers.

The plaintiff's estimates were based on a linear forecast of the remaining work for the project. For each workstation on each aircraft this involved noting the number of labour hours required to achieve the percentage of the work completed. This number of hours was then divided by the percentage of work completed to indicate the expected number of hours to complete 100% of the work. This is a normal forecasting procedure used by the organisation. However, the modellers' estimates were based upon a different approach. Through a spreadsheet analysis, the modellers had modelled the direct impact that each of the D&D triggers had on the project. An estimate of the forecasted impact of the triggers was also made. Although D&D has not been explicitly modelled, the modellers also included an estimated allowance for the impact of D&D. By adopting this approach, the modellers made an assessment about the future impact of embedded D&D. An allowance was made for the D&D that has been triggered on current production work as well as how the embedded D&D will impact on future production work.

Throughout the modelling process results simulated by the modellers were continually in excess of the plaintiff's normal forecasting procedures. As time progressed and actual labour hour data replaced forecasted labour hours, the actual data was repeatedly found to be closer to the modellers' estimates rather than the plaintiff's estimates. Due to the plaintiff's continual underestimation of the forecasts, they had to continually re-estimate their forecasts and therefore had to repeatedly alter their estimate of the overall time and cost overrun for the project. For example, in only 8 days the forecast of labour hours to completion for one aircraft was increased by 2.4%.

This phenomenon has been observed in other projects with which the author has been associated. The plaintiff does not have any tools to aid the assessment of feedback that had been triggered in the project, but had not yet shown its full effects. Their standard forecasts therefore underestimate the future ramifications of feedback. It is extremely difficult to even identify that D&D has been triggered before even considering the difficulties of understanding the full effects of D&D (Eden et al 2000). Due to the inability to accurately forecast embedded D&D, the forecasted overrun for a project may change dramatically throughout the claim process. This is not helpful to the plaintiff's case if they are regularly changing the amount of their claim for compensation. Obviously this situation is partly due to the occurrence of additional disruptions and delays. However, the inability to accurately forecast the outcome of embedded D&D plays a major role in the inaccurate forecasting. To help this situation, organisations need to improve the forecasting tools they utilise. The tools need to take the dynamic nature of D&D into account, rather than simply taking a linear view of the future.

2. Behaviour Anomaly

This test can be used throughout model development. When constructing the model, the production of graphs demonstrating the behaviour of various variables over time helps the modellers to track whether or not the model output is as expected. For example, does the behaviour contradict any of the information and data collected for the project?

This test can also be used through use of the *loop knockout analysis* (Sterman 2000, p880). This analysis aims to discover the importance of particular relationships. The test involves zeroing out a relationship in the model. Then, if an anomalous behaviour ensues, this would indicate that the relationship is particularly important. In using the *Decision Explorer* software to construct and analyse the cause maps for the projects discussed in this thesis, an analysis that was used is the *potent loop analysis*. This analysis informs the user of the number of feedback loops each concept is contained in. In doing so, the user is able to identify those concepts contained in the most number of feedback loops. Section 3.4 noted that for the projects discussed in this thesis *compression* is an important concept since it is contained in 90% of the feedback loops for the Castle project and 100% of the feedback loops for the Pirate project. This means that the relationships modelled around *compression* are important to the overall system behaviour and the modeller should therefore ensure that the information used to model them are as reliable as possible. The potent loop analysis is very useful to focus the modeller's attention on those relationships that are important to the system behaviour.

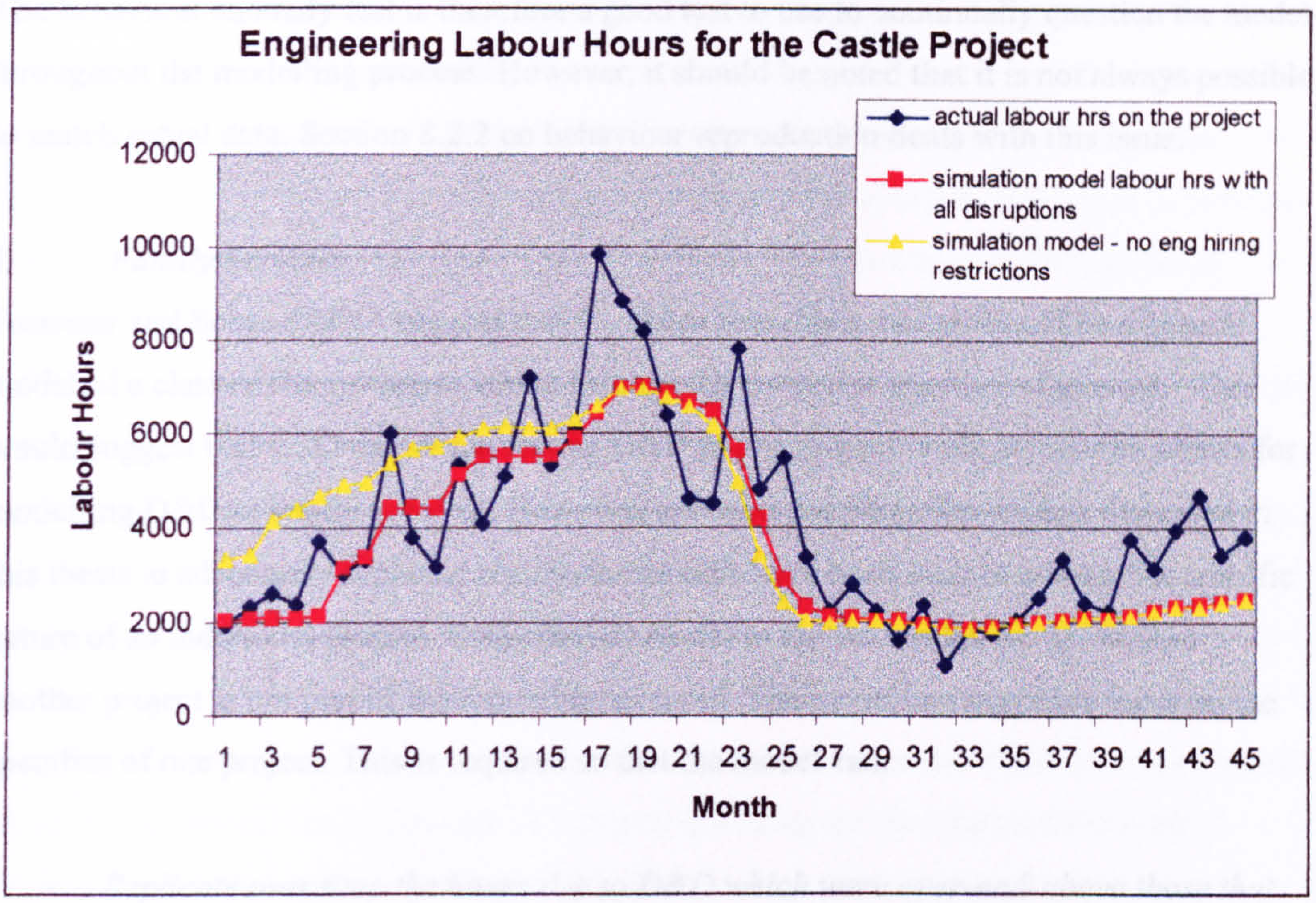
The behaviour anomaly test can aid in setting plausible ranges for parameters and relationships. This would therefore be done in conjunction with the behaviour and numerical sensitivity test.

The impact of one variable on the overall shape of a key variable

In the Castle project, the type of variables that were used as a part of the behaviour anomaly test were historical labour data, productivity levels, overtime figures and levels of rework. The actual behaviour was continually compared to the simulated data of key variables whenever changes were made to the model, to ensure that the model was not creating any unexpected behaviour. The particular variables tracked during the Castle project were dependent upon the data that was available. This data was available either in numerical form or through judgements made by those involved in the project.

An example of the use of the behaviour anomaly test during the modelling of the Castle project occurred when an anomalous behaviour was output by the labour hour variable during the simulation of the project outcome. The simulation of the initial few months produced a very different shape to the actual data. This can be seen in graph 7.2 below.

During the construction of the cause map that captured the events of the Castle project, the existence of caps on the number of engineering staff that were available for work on the project had not been captured. It was only during further discussions with some of the managers on the project regarding the anomaly in the shape between the actual and simulated engineering labour hours, that the existence of caps on the number of engineers that could be used on the project was highlighted. The model was then altered to reflect the new information. The question to be asked following this experience is why did the cause mapping process not capture this information?



Graph 7.2 – Capping the Number of Engineering Staff Available on the Castle Project

The draft claim document obviously did not include this information, otherwise, assuming that the modeller did a thorough job, it would have been included in the cause map. This means that those who drafted the claim document, from which the cause map was produced, did not believe the information to be of importance to the claim. This assumes that staff had recognised its existence when information had been gathered for the claim document. It was only when quantitative modelling was used that the importance of the information was highlighted.

Qualitative modelling may therefore not capture all the events that are important to the modelling process. If anomalies such as the difference in the shape of a variable as in graph 7.2 are highlighted during the quantitative modelling process, then the qualitative modelling may need to be revisited to explore if there are any omissions from the model.

The behaviour anomaly test is therefore a good test to use to continually question the model throughout the modelling process. However, it should be noted that it is not always possible to match actual data. Section 8.2.2 on behaviour reproduction deals with this issue.

3. Family Member

Forrester and Senge (1980) suggest that "... *when possible* a model should be a general model of a class of the system to which belongs the particular member of interest." This would suggest that a SD model replicating D&D in one project could be used as a basis for modelling D&D in another project. However, to enable the litigation models discussed in this thesis to adequately replicate reality, the models have been built to contain the specific nature of an individual project. Using the SD model to aid the modelling of D&D in another project is not part of the modelling purpose. The modelling purposes focus on the specifics of one project. This is required so that the model can:

- *Replicate over time the hours due to D&D which were over-and-above those that were contracted, but were required to carry out (for) the (specific) project.*

When modelling D&D for litigation, the use of this test would take the model beyond the purposes of the modelling process.

Tracking the causal story of a D&D trigger

Detailed causal stories illustrating the ramifications of D&D triggers have been used in the modelling of the projects discussed in this thesis as a part of attempting to fulfil the following modelling purpose:

- *Demonstrate that a part of the time and cost overruns were caused by D&D through particular triggers of D&D.*

Tracking the detailed ramifications of a D&D trigger is a way in which to demonstrate how the individual day-to-day events that have occurred on a project can lead to the more aggregate variables that are captured by the cause map and SD model. The use of such detail has proven to be a powerful tool in demonstrating the full and expansive ramifications of a specific, one-off event. The detail and specific nature of the information included have been found to be powerful in demonstrating the outcomes of defendants' actions. A more generic model would not have the same impact with the defendant. They could argue that it might not be truly representative of the specific project. Experience of working on the projects discussed in this thesis shows that defendants are most responsive to models that represent the specific project and contain the specific events that occurred in that project, so that there is no question over who did what to whom and when. When presenting detailed causal stories to a defendant's claim team, one of the author's colleagues recounts how the expression on one of the team's face visibly changed over the course of a couple of minutes from a confident to an extremely concerned expression as the specifics of the story unfolded. The detail and evidence to support each part of the story left no doubt that the defendant was responsible for a chain of events that had had multiple D&D ramifications throughout the project.

In a litigation environment, the various audiences are not concerned whether or not a model represents a general class of models. Their only concerns are that the model is built to represent the specific project under investigation.

4. Surprise Behaviour

As with the behaviour anomaly test, the modeller needs to continually track the behaviour of various key variables in the model in order to carry out this test. The behaviour of these variables should be compared with the behaviour that can be observed in the real system or a modeller's expectation of what the behaviour should be, based on gathered information. If differences occur between the two, then the reason could be due to flaws in the modelling of relationships and parameters. However, it may be the case that the model is a reasonable replication of reality and that the behaviour observed in the model has simply gone unobserved in reality.

In any SD model, the behaviour of various variables should be tracked during model construction, so that if any unexpected behaviour occurs, the modeller can investigate individual relationships and parameters before concluding whether the behaviour is *a surprise* or *anomalous*.

The number of engineering changes required during the Castle Project

Throughout the modelling of the projects discussed in this thesis, the models did not highlight any major surprises in the behaviour of the systems. However, during the modelling of the Castle project, model output did help to emphasise the number of changes that were faced by engineers. Graph 3.3 in section 3.5 demonstrates the extent of the number of engineering changes that actually occurred on the project compared to the number that had been expected before the project commenced. The claim team knew that the number of changes were large. However, when presented with graph 3.3, they had not fully appreciated the comparison of the actual changes to the number that had been expected. They felt that this graph improved the way in which they could demonstrate one of the main reasons for the project failing. The output contained in the graph had highlighted an area of the system that they had previously not fully appreciated.

5. Extreme Policy

This test "...shows the resilience of a model to major policy changes. The better a model passes a multiplicity of extreme-policy tests, the greater can be confidence over the range of normal policy analysis and design" (Forrester and Senge 1980, p222). This has not been a test that has been used in the construction of the models discussed in this thesis. The reason for this has been that the objective of the modelling process is not focussed on the consideration of different polices in an attempt to change the system.

However, there is a situation where the use of different managerial policies may be useful in the model. If we suppose that the plaintiff is accused of incompetent management, and that other managerial policies were suggested as the actions that the management should have taken, then the model could be used to demonstrate the effect that the alternative policies would have had on the outcome of the project. The plaintiff may carry out this analysis in an attempt to demonstrate that their management did not behave incompetently, rather that the alternative policies actually increased the actual labour hours accrued on the project or, if they did perform better, that the improvement was minimal.

Such a situation has not arisen in either of the projects discussed in this thesis. However, the possibility of it arising should be considered, since the extreme policy test would then become extremely useful in ensuring that the model would behave reasonably to a request of a change in managerial policy.

6. Behaviour Sensitivity

This test considers the change in model behaviour when sensitivity analysis is carried out on the parameters of the model. As well as considering the change in model behaviour, numerical sensitivity is also of concern when modelling D&D for litigation.

In section 4.8, it was suggested that the data required to populate a model may often not exist within the organisation or prove extremely difficult or time consuming to gather. In these cases statistical analysis of sample data or judgmental estimation may be used. The data generated from either of these methods of analysis will contain a range of uncertainty. Any modelling of the way in which decisions are made and the level of aggregation of variables will also contain a level of uncertainty. The various ranges of uncertainty need to be addressed to observe their effect on the behaviour and numerical output of the model.

For the reasons mentioned above, the modellers' report of the output of a SD model of D&D in a litigation environment would not present a single result. Ranges of results are presented which reflect the sensitivity analysis that has been carried out as part of the modelling process. In presenting their findings in this way, the modellers are using the model to demonstrate that it is reasonable to assume that the triggers described in the cause map can cause a cost overrun in the *approximate* range of what was actually experienced. The model is therefore used to support a case that it is reasonable for the plaintiff to request a settlement in the approximate region of the range of costs presented through the sensitivity analysis.

Both behaviour and numerical sensitivity analysis are therefore important when modelling D&D for litigation. Some time will be spent agreeing upon reasonable sensitivity analysis ranges for parameters. When the sensitivity analysis is run, the plausibility of the changes in behaviour of variables are considered. Also, the overall number of labour hours is an important variable to check for numerical sensitivity. This plays a very important role in the modelling process in a litigation process as it contributes directly towards the plausible ranges of money for which the plaintiff will request compensation.

When carrying out behaviour and numerical sensitivity analysis, both univariate as well as multivariate sensitivity analysis should be considered. These tests should be carried out as a part of behaviour and numerical reproduction tests. Behaviour reproduction tests are a part of the tests discussed by Forrester and Senge (1980) and will therefore be considered next.

7.3.4 Behaviour Reproduction Tests

One of the modelling purposes was stated as:

- *Replicate over time the additional hours, over-and-above those that were contracted, which were required to carry out the project.*

This purpose highlights that matching historical data, in this case labour hours as a time series, will certainly play a role in the modelling process when attempting to convince the audience that the model adequately replicates reality.

In fact, experience of the use of this test in a litigation environment has proven to be very important. It is a powerful test in convincing the various audiences that the model adequately replicates reality. Sterman (1984, p52) discusses how the matching of historical data has been viewed by many system dynamicists as a weak test and has thus tended to be ignored. He identifies the importance of matching simulated and actual data with respect to general SD models by stating that "failure to satisfy a client or reviewer that a model's historical fit is satisfactory is often sufficient grounds to dismiss the model and its conclusions". For this reason he believes that "passing the historical behaviour test, while far from sufficient, is a necessary step in the confidence-building process"(p52). In a litigation environment, gaining audience confidence in a model is vital, therefore the reproduction of historic data is important. To gain audience confidence in the model, an adequate replication of what actually occurred in the project is the type of visual evidence that could help an audience with no modelling expertise, such as a judge and a jury, to assess a model. Of course, other tests are required to ensure that the model replicates reality to a satisfactory degree *for the correct reasons*. On the other hand, not being capable of matching historic data could be extremely detrimental for the model.

Behaviour reproduction also needs to be considered alongside numerical reproduction to help support the plaintiff's case for compensation. However, numerical reproduction should be considered alongside numerical sensitivity since providing the audience with a singular numerical output would be misleading. Numerical sensitivity is used to provide a range of results to the audience.

Due to the importance of this test and the fact that the behaviour reproduction test is an area where the most advancement has been made beyond the tests noted by Forrester and Senge (1980), behaviour and numerical reproduction tests will be discussed in full in the next chapter.

7.4 Summary of Chapter

This chapter focussed on the modeller's requirements to be adequately convinced that the model replicates the important aspects of reality. This is an essential step before any of the other audiences are exposed to the model.

It was stated that a model would never *fully* replicate reality, since any model is a simplified representation of reality. Therefore before a modeller can make an assessment of whether or not a model is an *adequate* representation of reality, the reasons for a model being a simplified representation of reality need to be considered.

Three reasons for a SD model being a simplified representation of reality were explored:

(i) The aggregation of variables.
(ii) The modeller's choice of system boundary.
(iii) The omission of noise.

In most SD models replicating reality relates to adequately modelling the behaviour of key variables. However, when modelling D&D for litigation, the modelling purposes also required an adequate representation of the *quantitative* results of key variables. Therefore, the three reasons for a SD model being a simplified representation of reality were discussed in order to consider whether or not SD can replicate ***sufficient detail*** to be able to fulfil the modelling objectives. It can be concluded that:

- When considering the level of aggregation of variables in a model, the following statement made by Forrester (1961, p66) should be borne in mind "When a model has progressed to the point where ... there is reason to believe that discreteness has a significant influence on system behavior, discontinuous variables should then be explored to determine their effect on the model". What is deemed to be a *significant* influence on system behaviour will depend upon the purposes of the modelling process. When using SD to model problems where the modelling purpose is to improve understanding of the system and thereby improve the way in which the system is managed in the future, the replication of the general behaviour of the system is normally sufficient. However, when the modelling purpose involves:

- *Replicate over time the hours due to D&D which were over-and-above those that were contracted, but were required to carry out the project*

the matching of individual points of data becomes more of a concern. This is especially true when the model may be the basis of a claim for a large amount of money in a litigation environment. For example, in a policy model a very slight shift in a time series may produce no difference in the model behaviour, but result in 1,000 additional labour hours. At a typical rate of, say, £30 per hour claimed, this is not an insignificant difference in the outcome for the litigation case.

- The modeller has a reasonable amount of choice about the concepts to include in the system boundary to enable the modelling purposes to be met. Some time and effort may be required in the consideration of what should and should not be included to enable *sufficient* detail to be included in the model. The specific case of the exclusion of self-inflicted D&D from the model has been highlighted. This may lead to a difficulty in matching actual data which will be returned to in chapter 8.

- Noise can be introduced into a model with the intention of observing whether or not the shape and quantitative results of key variables alter due to this. Based on the author's experience of the Castle and Pirate projects, noise has been seen to have a significant impact on the outcome of a model if it forms part of a discrete variable. However, when assessing the actual impact that noise will have on the output of a model, it should be considered alongside sensitivity analysis.

During the modelling process a modeller needs to be convinced that the model under construction is an adequate replication of reality. The SD literature discusses confidence building tests that have been used for this purpose. The relevant tests to consider when modelling D&D for litigation are below. These are listed together with some considerations of their use in this type of environment.

- *Structure Verification*

  A full appreciation of the various engineering processes should be achieved in order to adequately verify that the more aggregated structures represented in the SD model correctly capture the nature of a project. Such an increased level of detail is normally not required for all parts of a project. However, one reason for the detail being required was observed in the Castle project. The structure of one part of the real system was particularly complex and hence required an increased level of detail so that the modeller could gain an appropriate level of understanding of the structure of the system.

- *Parameter Verification*

  Parameter verification is an important test during the construction of SD models. However, conceptual and numerical difficulties mean that the process is not straightforward and may require a large amount of time and effort to carry out.

  Conceptual difficulties arise in litigation cases when lengthy debates occur over the meaning of parameters. Also, numerical difficulties arise when the organisation does not record the data required by the parameters. If an organisation records the minimum data they need to carry out a project, then if modelling is required for a litigation process, the percentage of parameters with numerical difficulties may be extremely high. It is only when staff have been involved in a litigation process that they appreciate the extent of data that is needed and may then go on to change the data that is recorded on projects.

- *Boundary Adequacy*

  The system boundary should include all the ramifications of D&D triggers. This will include all types of managerial actions that were used during the life of the project in response to the need to compress the project and their ramifications.

Identifying all the ramifications of managerial actions may be a process that managers have not previously gone through. Therefore, the process may have a secondary effect of getting managers to question the managerial actions they often take for granted. Carrying out such a process with managers will have an effect on the way in which they manage future projects. Having explored all the ramifications of their actions they should have an improved understanding of the impact of future actions.

- *Dimensional Consistency*

  This is an important test for *any* SD model. Good documentation of each variable, including its dimensions, should always be carried out when constructing a model, especially when the model is as large and complex as those that are discussed in this thesis.

- *Behaviour Prediction*

  The inability to accurately forecast the outcome of a project is a common problem before the completion of the project. This is partly due to the extreme difficulty in identifying any embedded D&D that exists in the project, as well as the future effect that D&D will have on a project. This problem is not helpful in a litigation environment, as it means that the plaintiff may be continually altering the number of labour hours they forecast as the project overrun. To help this situation, organisations need to improve the forecasting tools they utilise. The tools need to take the dynamic nature of D&D into account, rather than simply taking a linear view of the future.

- *Behaviour Anomaly*

  The behaviour anomaly test is a good test to use to continually question the model throughout the modelling process. In particular, it can be used to question whether or not the qualitative model contains all the events that it should. The importance of the inclusions of some information may not be highlighted until the quantitative modelling process is under way. This is one reason for the quantitative modelling process creating a situation where the modeller needs to re-visit the qualitative model to explore whether or not information has been omitted.

The behaviour anomaly test can also be used to help to set plausible ranges for parameters and relationships. This would be done in conjunction with the behaviour and numerical sensitivity test.

- *Surprise Behaviour*

  In any SD model, the behaviour of various variables should be tracked during model construction, so that if any unexpected behaviour occurs, the modeller can investigate individual relationships and parameters before concluding whether the behaviour is *a surprise* or *anomalous.*

- *Sensitivity Analysis – Behaviour and Numerical*
- *Behaviour Reproduction*

Both behaviour and numerical sensitivity analysis were seen to be of particular importance when modelling D&D for litigation. Some time is normally spent agreeing upon reasonable sensitivity analysis ranges for parameters. When the sensitivity analysis is run, the plausibility of the changes in behaviour of key variables should be considered. Also, the overall number of labour hours is an important variable to check for numerical sensitivity. This plays a very important role in the modelling process for litigation as it contributes directly towards the plausible ranges of compensation that the plaintiff will request in a claim. Experience of the use of the behaviour reproduction test in a litigation environment has also shown that it can be very important. It can be a powerful test in convincing the various audiences that the model adequately replicates reality. Also, this test is an area where the most advancement has been made beyond the tests noted by Forrester and Senge (1980). This test will therefore be considered in the next chapter.

CHAPTER 8: BEHAVIOUR AND NUMERICAL REPRODUCTION

8.1 Introduction

The aim of the last chapter was to consider the confidence building process that a modeller needs to go through when considering whether or not a SD model of D&D adequately replicates reality. Chapter 7 discussed several confidence building tests laid down in the SD literature. One of these tests; behaviour and numerical reproduction test was highlighted as playing a particularly important role in the modelling process. Experience on litigation cases has shown that the behaviour reproduction test is a powerful test in convincing the various audiences that the model adequately replicates reality. Due to the importance placed on this test, it was decided that it should be explored in more detail. This test will therefore be explored in this chapter.

8.2 Behaviour Reproduction Tests

One of the modelling purposes has been stated as:

– *Replicate over time the additional hours, over-and-above those that were contracted, which were required to carry out the project.*

This purpose highlights that matching historical data, in this case labour hours as a time series, will certainly play a role in the modelling process when attempting to convince the audience that the model adequately replicates reality.

The SD literature informs us of tests that have been previously used on SD models to assess the adequacy of model-generated behaviour. Senge and Forrester (1980, p218) note that "in the literature on modeling and simulation, there are a wide range of tests involving point-by-point comparisons of model-generated and observed behavior. Despite widespread acceptance, such tests involving point-by point measures of goodness of fit are generally less appropriate for system dynamics than the symptom-generation, frequency-generation, and multiple-mode test…". These tests suggested by Forrester and Senge are tests that can be used to assess the adequacy of *patterns* that are generated by a model rather than individual *points*. This point is supported by Barlas (1989, p60) as he states that "it is well documented that, by their nature, SD models do *not* predict specific *individual* values of output variables... but they predict major *time patterns* of concern".

The problem with many of the pattern prediction tests is the level of subjectivity required. Judgement is needed to decide whether or not a given model-generated pattern can be deemed adequate and no assistance is given in the literature as to what should be deemed adequate. This has been one cause of one criticism that "system dynamics does not employ formal, objective, quantitative model validation procedures, which are supposed to be fundamental to scientific inquiry" (Barlas and Carpenter 1990, p148). However, this criticism has been judged to be a subjective view itself as Barlas and Carpenter (1990) conclude that whether or not an individual will accept that the system dynamics method is truly scientific depends on their philosophy of science.

Forrester (1961, p121) summarises the need for subjectivity in the system dynamics validation process by stating that, "what is close enough depends on the purpose of the model and the nature of the discrepancy". If pattern prediction tests were used on a SD model in a litigation environment, the implication would be that the modeller would have to make a judgement on what he or she believed that the different model's audiences would consider to be adequate.

Barlas (1989) notes the problems with qualitative pattern prediction tests and suggests an alternative set of "appropriate and simple" quantitative tests. However, he later (1996) states that "if the problem involves a transient, highly non-stationary behavior (such as a truncated S-shaped growth, or a single boom-then-bust pattern), then it is impossible to apply any standard statistical measure. The problem is of *no* statistical nature to start with and therefore no general statistical test can be offered in this case. The best approach is to compare graphical/visual measures of the most typical behavior-pattern characteristics... If, on the other hand, the problem involves a long-term steady-state simulation, then, it is possible to apply certain standard statistical measures and tests" (Barlas, 1996, p194).

Forrester and Senge (1980, p209) state that behaviour reproduction is one of the core validation tests for a SD model. Behaviour reproduction tests examine "how well model-generated behavior matches observed behavior of the real system" (Forrester and Senge, 1980, p217). Such tests not only include point-by-point comparisons, "the test usually focuses on the character of the simulated data: does it exhibit the same *modes, phase relationships, relative amplitude,* and *variability* as the real data" (Sterman, 1984, p52, Italics added).

As discussed in section 7.3.4, Sterman (1984, p52) discusses how the matching of historical data has been viewed by many system dynamicists as a weak test and has thus tended to be ignored. However, Sterman identifies the importance of matching simulated and actual data with respect to general SD models by noting the importance the client places on this test. In a litigation environment, gaining audience confidence in a model is vital, therefore the reproduction of historic data is important. Of course, other tests are required to ensure that the model replicates reality to a satisfactory degree *for the correct reasons*. On the other hand, not being capable of matching historic data could be extremely detrimental for the model.

To be able to assess how well a model replicates historic data, Sterman (1984, 2000) describes summary statistics which can be used to evaluate the reason why there is point-by-point differences between actual and simulated data for a SD model. Judgement then has to be used, *based on the purposes of a model,* whether or not the modeller is willing to accept the differences. The summary statistics involve the decomposition of the mean-square-error (MSE) when comparing simulated and actual point-by-point values. The MSE is decomposed into the Theil inequality statistics which consist of three measures. These measure are the *bias* between the simulated and actual series, the *differences in the variances* of the two series and the *component of the error* due to incomplete covariation between the series (measures the degree to which the changes in the simulated series fails to match the changes in the actual series point-by-point). Each of these measures is divided by the overall MSE to derive 'inequality proportions' to reflect the fraction of the MSE due to bias (U^M), unequal variance (U^S) and unequal covariance (U^C). These summary statistics provide a quantitative method to enable the modeller to then proceed to a subjective judgement on how adequately a model replicates reality. For this reason the summary statistics are helpful in the exploration of whether the SD models in this work replicate reality convincingly.

8.2.1 Model Tests in Project Based SD Literature

Many of the problems addressed by SD models have been of a long-term nature where the variables considered may represent infinite time series. The examples used to assess SD model tests have therefore also been of a long-term nature. For instance, Forrester and Senge (1980, p217-218) discuss 10 to 25 year fluctuations when examining the frequency-generation test, 3 to 7 year and 18 year fluctuations when examining the multiple-mode test. However, when considering a model of a single project, due to the finite life of a project, the variables being tracked during the project will produce finite time series (for example, the SD model replicating the Castle project simulated a period of 3.75 years). Due to this difference when using SD to model a single project, it is worth focussing particularly on previous SD models that have replicated single projects in an attempt to understand if there are any specific confidence building tests that have been used in this type of modelling.

The following represents the main body of published work discussing the use of SD to model a single project:

Roberts (1964) develops a model to simulate the undertaking of research and development projects. The confidence building process focusses upon being able to explain the model's output in relation to empirical knowledge of how research and development projects behave.

Richardson and Pugh (1981) utilise some of the model tests discussed in chapter 7 to aid the understanding of the behaviour of a project model. They follow this by discussing validity tests in general and tie these tests back into their work on the project model. Statistical tests are not considered as part of the confidence building process of the model.

Weil and Etherton (1990) used SD to model the production of a number of ocean-going tankers. The SD model was used to highlight the causes for the overrun of the project in both time and cost as part of a litigation process. In this article, the only reference that is made to model tests is to mention that "...results were compared statistically to what actually happened in the course of the project".

When modelling software project dynamics, Abdel-Hamid and Madnick (1991) do not specifically address model validation. However, they do "examine the model's ability to reproduce the dynamic patterns of a completed software project" (p139) by using a case study to test the model. They compare four variables from their model with actual data. A visual inspection of the closeness of the model-generated variables with actual data that they have gathered is carried out and plausible reasons given for the model-generated variables slightly differing to the actual data.

Lin (1993) and Lin et al (1997) use sensitivity analysis and historical project comparison tests when attempting to gain confidence in the development of a planning tool for software-project development processes. The historical project comparisons involve the use of root MSE as a measure of accuracy and the analysis of the variances. Other confidence building tests also mentioned in this work are (i) the removal of one of the simulated decision functions which is then replaced with a real decision-maker to test whether or not the model captures a general decision-making behaviour (ii) a Turing test where managers are asked to identify whether output information is based on real or simulated data.

Ford (1995, p222-224) carries out a visual inspection of the closeness of model-generated variables with actual data for a multiple phase project (including product definition, design, prototype testing and reliability/quality control). Plausible reasons are given for the differences. Ford also calculates R^2 (the coefficient of determination i.e. represents the fraction of the variance in the data that is explained by the model) for each of the variables (the R^2 results range from 76 to 99%).

A more detailed account of model tests is presented in Cooper's (1980) account of the use of SD to model a time and cost overrun in a shipbuilding project as part of a litigation case. The model tests that are reported as being carried out on the model are:

- Matching of historical data. This included both "hard" and "soft" data to ensure that the model was consistent with all available information.
- Statistical testing (although the literature does not report what form of statistical testing is used).
- Sensitivity analysis of parameters.
- Shock tests to assess robustness in responding to radically different circumstances.

- Exploring alternative models that may plausibly represent the system by using different combinations of reasonable equations and parameters.

When comparing these tests to those discussed in chapter 7, they relate to:

- Behaviour and Numerical Reproduction
- Behaviour and Numerical Sensitivity
- Extreme Conditions/Extreme Policies. When discussing the extreme policies test in chapter 7 it was concluded that the plaintiff may find the modelling of alternative policies useful in an attempt to demonstrate that their management did not behave incompetently. Therefore, the extreme policies test would be a useful test to perform. However, the extreme conditions test was concluded as a test that would not aid a model to meet the purposes of modelling D&D for litigation. An explanation for the reason for Cooper's (1980) use of the extreme conditions test is noted by Sterman (2000, p61). The SD model was developed by replicating a generic module to represent different parts of the project. This module was required to represent a diverse array of activities. Therefore, to ensure its robustness, the extreme conditions test was used to ensure that the module behaved appropriately under different combinations of inputs and conditions.
- Structural and Parameter Verification

The focus of the model tests carried out in Cooper's work was to ensure that "...the model replicated quite accurately the vast amount of detailed information on the programs' histories". The behaviour and numerical reproduction tests were therefore important. The other tests were used "In order to assure... that the model performed correctly for the correct reasons..." (Cooper, 1980, p27).

Most of the literature discussed above does not place a great deal of emphasis on model tests. The exception is Cooper (1980). However, his emphasis on model tests is likely to be similar to the needs of the projects discussed in this thesis i.e. the need to *convince an audience* that a model adequately reflects a real system to the extent that they will believe in its conclusions and agree that one party should pay compensation to another party as part of a litigation case. A litigation environment forces a rigorous validation process before an audience accepts a model as an adequate representation of reality.

Abdel-Hamid and Madnick (1991) and Ford (1995) provide a detailed account of the behaviour reproduction tests they have carried out. However, the work that they recount is an informal visual inspection of key variables. A project based SD model that has used *formal* statistics rather than informal visual inspection is discussed by Ford and Sterman (1998). This work details the use of the MSE statistics (introduced in section 8.2) proposed by Sterman (1984) in the validation of a SD model built to represent the development phase of a project.

As Sterman's MSE summary statistics (1984) are the only statistical tests that have been used in the literature to help validate an SD model of a single project, they will be used to discuss behaviour reproduction in the Castle project. However, before moving onto this discussion, it should be noted that the behaviour reproduction tests should be used *throughout model construction* to identify areas where the model requires improvement. They are not tests simply applied at the end of the process to prove that a model is or is not correct. Comparing actual and simulated data throughout model construction can lead to various amendments and improvements to a model. However, a final comparison between actual and simulated data can be used as part of the argument that a model is an adequate replication of reality. A model that can visually reproduce the types of behaviour observed in a real system is visually important when illustrating reasons for having confidence in a model.

8.2.2 Difficulties in Attempting to Reproduce Behaviour in a SD Model for Litigation

Before moving on to discuss the MSE summary statistics it is important to consider the difficulties that exist with behaviour and numerical reproduction in SD models in a litigation environment.

Self-Inflicted D&D

The most important issue to consider is whether or not the simulated time series *should* replicate the actual time series. One of the main reasons that the simulated time series would *not* replicate the actual time series is due to the existence of self-inflicted D&D. In section 7.2.2 self-inflicted D&D was described as any D&D that was believed to have been triggered by the plaintiff.

In taking a thorough approach to the modelling process, the modeller will wish to include all the main triggers of D&D. By involving a cross-section of people who worked in the project, the interview process aims to elicit all the triggers of D&D. However once these triggers have been elicited, it is likely that there will be D&D which could be blamed on actions taken by either the plaintiff or the defendant. This makes it very difficult for the modellers to differentiate what is and is not self-inflicted. Indeed, it may not be until the lawyers argue the case in front of a judge that it is finally concluded what is and is not self-inflicted. However, there may be some events that the plaintiff believes to be clearly self-inflicted or, at least, the blame cannot be assigned to the defendant. For example, due to the plaintiff taking on too many projects, work on the specific project is delayed. If the claim team do not wish to explicitly identify this event, then the modellers may be instructed that these events should not be included in the final SD model. In this case, the output of the key variables of the model will not match the behaviour or numerical values of the variables in reality.

Due to differences in what is included in the model when compared to reality, the modeller needs to consider what actual and simulated time series should be used for comparisons. There are three possible approaches:

1. Include the self-inflicted triggers in the model. The actual and simulated time series will then both include self-inflicted triggers.
2. Compare the actual time series minus an allowance for the self-inflicted triggers against the simulated time series excluding self-inflicted triggers.
3. Compare the actual time series with the simulated times series which excludes self-inflicted triggers. Plausible stories are then required to explain the differences between the actual and simulated time series.

Considering each of these approaches in turn:

1. The problem with this approach is that the plaintiff's claim team may not wish to explicitly model the self-inflicted triggers. They may be willing to take responsibility for these triggers during a period of negotiation, however the lawyers may wish to keep these as *give-aways* and *not show their hand* immediately from the beginning of the litigation process.

2. The problem with this approach is that the modeller would still need to be explicit about what was deemed to be self-inflicted triggers and would have to quantify them in order to reproduce an actual time series that excluded them.
3. This approach has been the most preferred during the projects discussed in this thesis. This approach does not require an explicit modelling of the self-inflicted triggers. Some of the difference between the actual and simulated time series can then be explained due to the model being a simplified representation of reality. It may then be explained that minor elements not included in the model include the self-inflicted triggers of D&D.

Of course, the issue with not including everything in the model is that behaviour and numerical reproduction becomes very difficult.

If there were large differences between the actual and simulated time series, then it is likely that the modellers would want to explain this, even if it means modelling self-inflicted triggers. As previously mentioned, what is and is not self-inflicted is dependant upon how the case is argued. For example one of the triggers of D&D identified in the Pirate project was an underestimation in the original bid for the project. Since the plaintiff constructed the bid, it could be argued that the underestimation was self-inflicted. However, the plaintiff's argument was that the defendant had previous knowledge of the type of work that was involved in the project, whereas the plaintiff did not have this experience. The defendant then proceeded by not informing the plaintiff of their knowledge. The plaintiff's lawyers therefore argued that if the defendant had made their knowledge known to the plaintiff, the plaintiff would have sought their advice and would not have submitted such a naïve estimate.

Modelling Sub-sections of the Project

Another difficulty in behaviour and numerical reproduction occurs when the model does not include all elements of the project. One reason for this may be that the model only includes those parts of the project that are relevant to the *indirect* costs of D&D. The parts of the project omitted from the model may be believed to only contribute towards a direct cost overrun in the project (see section 2.4 for a discussion concerning the difference between direct and indirect costs). Those elements may be omitted since SD models are normally used to aid the understanding of the indirect consequences of D&D. They are often not required to model the direct consequences of D&D since they are more easily thought through.

Another reason for only modelling a sub-section of a project is that it may be believed that the D&D that occurred in other parts of the project was minimal. It may therefore not be worth the time and effort involved in modelling the other elements of the project. The modelling effort may therefore focus on those sections where the greatest impacts from D&D were believed to occur.

Modelling sub-sections of a project is an important issue when considering behaviour and numerical reproduction. For any labour hours being used as part of a behaviour reproduction test, only that part of the hours that represents those elements that have been included in the model should be used in a comparison. However, it can prove difficult to distinguish between the hours that can be attributed to those elements contained within the model and those that have not been included. For example, the plaintiff's record of labour hours may not be split between those elements included in the model and those that are not included. This further complicates the confidence building process as the modeller must first be convinced which parts of the hours should be used for comparisons before then going on to carry out the comparisons. An example of this is the labour hour data used for the Castle project. Only the work carried out by certain types of workers was included in the model. The engineers drawing the designs were included. However, work not directly related to design work was omitted. For example, the production of technical publications or the time associated with management meetings. These were seen as overheads to the project. They were not modelled as part of the main feedback that occurred in the project. Instead, the increase in the hours spent on these overheads was seen to be a consequence of the feedback.

A further complication with this approach arises when gathering subjective data. Gaining subjective judgement regarding, for example, pressure and morale that was experienced on the project is a difficult process itself. However it would then be even more difficult to get those involved in the project to consider the pressure and morale issues for a specific subset of workers. This would create the problem of people attempting to mentally disaggregate the effects of pressure and morale between sub-sections of workers on the project. In the Castle project, the subjective judgements gathered were therefore considered for engineering as a whole. The assumption that the modellers then had to make was that the data for the specific subset of workers considered by the model could be represented by the overall data gathered for engineering. This can be discussed with those providing the expert judgements to consider whether or not there are any reasons for those included in the model to have experienced something different to engineering as a whole.

Bearing the discussions in this section in mind, this chapter will continue by considering behaviour and numerical reproduction in the SD model for the Castle project.

8.2.3 Sterman's Summary Statistics for the Castle Project

Before we can apply Sterman's summary statistic to the Castle project, the variables for which simulated and actual data will be compared have to be chosen. The variables which are most relevant for consideration are those variables which would be of most importance to the model's audience i.e. any variable demanded by the modelling purposes. However, practical issues also need to be taken into consideration i.e. for which variables is there reliable, reasonably accurate actual data available. Therefore, the variables that should be used will differ depending upon the project that is being modelled. This depends upon the amount and type of data that is actually recorded by the project management team.

One of the objectives of the modelling purpose is to be able to:

- *Replicate over time the additional hours, over-and-above those that were contracted, which were required to carry out the project.*

Therefore, labour hours is a variable for which it would be beneficial for the model to replicate. This variable can be directly translated into dollars of compensation in the litigation claim, therefore an audience needs to be convinced that the model is replicating reality convincingly for this variable.

Figures 3.1 and 3.2 in section 3.5 show graphs of actual and simulated data for the labour hours representing the budget and project outcome for the Castle project. Before proceeding to consider the MSE statistics for the time series shown on these graphs, it is worth recalling Barlas' (1989) quantitative tests. These tests represent the other advancement since Forrester and Senge (1980) in behaviour reproduction tests other than

Sterman's summary statistics. However, when inspecting the time series in figure 3.1 and 3.2, it can be concluded that the time series exhibit non-stationary behaviour. The requirement for a stationary time series includes a constant mean and variance over time. However, the time series can be seen to increase, then decrease producing a varying mean over time. This pattern which is a "single boom–then-bust pattern... is of *no* statistical nature to start with and therefore no general statistical test can be offered in this case " (Barlas 1996, p194). Of course time series which exhibit such patterns can undergo statistical transformations to enable further considerations on whether or not general statistical tests could be used. Therefore, unless the data could be successfully transformed into a stationary non-transient time series by some form of statistical transformation, Barlas would conclude that the quantitative tests he suggests should not be applied to these time series.

Returning to the MSE statistics, the summary statistics for the data shown on graphs 3.1 and 3.2 are as follows:

	Budgeted Data (Graph 3.1)	Project Outcome Data (Graph 3.2)
U^M : Bias	1.7%	15.2%
U^S : Unequal Variance	1.7%	5.5%
U^C : Unequal Covariance	96.7%	81.3%
R^2 : Coefficient of determination	87.2%	69.1%
MAPE : Mean Absolute Percent Error	7.5%	22.5%

Table 8.1 - Behaviour Reproduction Summary Statistics for the Castle Project Using Monthly Data

It should be noted that the simulated time series on graphs 3.1 and 3.2 are based on the most likely set of values for the input variables. However, uncertainty exists for many of the values of these variables. Sensitivity analysis was therefore used to consider both numerical and behaviour sensitivity of the simulated time series (see section 7.3.3).

When considering the statistics in table 8.1, the previous discussion on the existence of self-inflicted D&D in a project may lead us to expect a difference between the actual and simulated time series. However, in the Castle project, the plaintiff believed that the majority of D&D could be argued back to events triggered by the defendant. The plaintiff was therefore hoping to claim for all of the triggers that were included in the cause map. It was therefore expected that the model should capture the majority of the behaviour of the actual project. The statistics in table 8.1 may therefore be examined without the concern that one time series includes major triggers that are not included in the other time series.

In table 8.1, MAPE is reasonably small for the budgeted data, but is larger for the project outcome data. This is confirmed by visual inspection of the graphs, where the point-by-point errors on the project outcome graph are relatively large for certain points. There is no set percentage that a MAPE statistic should be below to enable a modeller to decide whether or not simulated data is a good representation of the real data. Each time MAPE is used, the result should be taken into account along with other statistics and evaluated for the specific modelling circumstance. For example, a result of 22.5% may be deemed acceptable if the modeller believes there are good reasons for the differences between the actual and simulated data.

For both the actual and budgeted graphs, the majority of the error is concentrated in unequal covariance. This suggests that point-by-point values of the simulated and actual series do not match even though the model captures the average and dominant trend in the actual data well. It could therefore be concluded that the differences between the actual and simulated series are due to the model not tracing some large random component of noise that is present in the actual system (Sterman 1984).

However, the above conclusion does not appear to be a reasonable conclusion to make for the Castle labour hours data. The reasons for this statement are as follows:

- The use of monthly data. Genuine noise occurs continuously over time. The smaller the time unit used for reporting data, the more noticeable the noise will appear in the data. For example, if labour hours are reported daily, then the data is more likely to demonstrate peaks and troughs due to a random component of noise when compared to monthly labour hours data. Each monthly data point will include noise that occurred over a 1 month period. Therefore, the noise contained in each monthly data point is unlikely to vary substantially.
- The size of the differences between the point-to-point values. It does not seem reasonable to suggest that the differences between the two time series are only due to unsystematic errors. The large differences must surely be at least partly caused by elements that the modeller has omitted from the model. This omission of the causal hypothesis of elements of the real system from the model may be through choice. However, they may also be due to the modellers' lack of knowledge of the existence of the elements in the real system.

The differences between the actual and simulated data are therefore due to both noise *and* the omission of elements of the real system from the system boundary for the model. In this case the MSE statistics could be concluded to be informing the modeller that sufficient elements have been included in the system boundary to enable the underlying behaviour of the variable to be captured. However, additional causal hypothesis would need to be included in the model if the more detailed variable behaviour is to be explained through the modelling process.

8.2.4 Unit of Time between Time Series Data Points

As can be seen by graphs 3.1 and 3.2, the time unit used to display labour hours was months. This unit was chosen since it was the time unit used by the company to record labour data. However, simply because the company records data in such a manner does not mean that this is the most appropriate time unit to choose for the SD model.

When graphing a variable output, the time unit used will have an effect on the resulting pattern of behaviour. A modeller therefore needs to make a decision on the time unit that should be used to analyse the data. This is of particular interest when making point-by-point comparisons such as in the MSE statistics since these points will differ depending upon the unit of time between data points. When MSE statistics have been discussed in the

SD literature, this issue has not been considered. It is therefore worth considering whether or not this would make a difference to the MSE results for the Castle project.

There may be an ideal time unit that the modeller would wish to work with, however practical constraints may hinder the modeller's choice. For example, the manner in which the company record labour data in the Castle project means that weekly labour figures are not available. Monthly data was the smallest time unit of data available to the modellers. Therefore, when considering whether or not different time units would make a difference to the MSE results for the Castle project, the only experiment that can be carried out is to move towards the use of a larger time unit. The MSE statistics have been re-calculated using a two monthly time unit. The results of this experiment can be seen in table 8.2.

	Budgeted Data	Project Outcome Data
U^M	1.1%	18.3%
U^S	0.6%	1.7%
U^C	98.3%	80.0%
R^2	90.7%	75.1%
MAPE	5.1%	19.7%

Table 8.2 - Behaviour Reproduction Summary Statistics for the Castle Project Using Two-Monthly Data

Budgeted Data: Both U^M and U^S have reduced and U^C has increased. This means that a two-monthly comparison places even more emphasis of the differences on a large random element not captured by the model. Since the monthly data explained a little more of the differences (although a minor amount) on an element of systematic errors in the model, perhaps weekly data would highlight this even further.

Project Outcome Data: U^M has increased and U^S has decreased whilst U^C has also slightly decreased. Although the majority of the difference is still explained through unsystematic errors in the model, the secondary emphasis of the differences in the data being due to a constant bias has increased whilst the minor emphasis on the difference in variance has reduced. Does this mean that the reduction on the secondary emphasis of error in the bias would reduce further if weekly data could be explored?

It would have been useful to expand these experiments to weekly data for the Castle project. However, weekly data is not available for any of the projects discussed in this thesis. It is likely that the existence of weekly rather than only monthly data would be useful to the modelling process. This would provide more data points for the modeller to take into consideration during the behaviour reproduction test. The disaggregation of the monthly figures into weekly figures may highlight patterns that are not seen through the monthly data. However, unsystematic errors that may have been hidden by the monthly data may also be highlighted due to the use of a smaller time unit.

Even if the modeller feels that the monthly data is more suitable for illustration purposes, the weekly data should be checked to see if the results it produces in the behaviour reproduction tests enables additional insights to the analysis.

When the appropriate time unit is decided upon and the MSE statistics calculated, judgement has to be used, based on the purposes of a model, whether or not the modeller is willing to accept the differences highlighted by the statistics. The question to then ask is whether or not the model's system boundary *should* or *could* be expanded in order to explain more of the system behaviour through the model. In order to be able to consider this question, plausible reasons are required for the differences between the actual and simulated data. The next section will therefore discuss this in terms of the data being explored for the Castle project.

8.2.5 Visual Inspection of the Actual and Simulated Data in the Castle Project

The results of the MSE summary statistics now lead us towards a more subjective form of analysis. A judgement has to be made whether or not, based upon the statistics, the modeller is convinced that the model does indeed replicate reality sufficiently. The discussion above suggests that the discrepancies between the actual and simulated series are a result of both noise and the omission of elements of the real system from the system boundary for the model. If plausible reasons for any large discrepancies could be formulated, then the modeller can make a decision on whether or not to expand the system boundary to include the additional causal hypothesis. Otherwise, the plausible explanations may at least help alleviate any potential uncertainties felt by the model's potential audience caused by the differences between the time series.

For the Castle project, the differences between the actual and simulated data will now be considered for graphs 3.1 and 3.2.

Budgeted Data – Graph 3.1

Visually, the budgeted simulation model tracks the budgeted figures well:

- Months 1 to 7: The simulated data appears as a smoothed representation of the actual data.
- Months 7 to 10: The simulated data is very close to the actual data.
- Months 11 to 15: Month 11 is where the largest difference occurs between the actual and simulated data. When discussing the difference between the two data points with the plaintiff, no specific events on the project could be highlighted which would have caused the estimators to allow for a 1 month increase in engineering labour hours. Unfortunately, interviewing those people who estimated the original budgeted hours was not possible. However, as the discrepancy between the two data points is large in relation to the discrepancies between the other data points, it is an obvious place in the graph where doubt may be cast that the model is missing, or not correctly replicating, an important event. Assuming that the estimator did not simply make an error when constructing the budget, then one plausible story is based on the milestones that were set for the project. (Note that a *plausible* story is all that can be considered since a lack of documentation of the reasons for the budgeted hours means that only the original estimator knows the actual reason for the increase in labour hours as it was not documented). The plaintiff received money continuously throughout the course of the project, based on the project reaching certain targets. There were many milestones throughout the project. However, when interviewing one of the program managers it was ascertained that only certain milestones were taken to be crucial (based on the amount of financial pay-off or the criticality of the milestone to the project schedule). These were the dates that managers on the project used as short-term project deadlines to work towards and therefore determined where the pressure points of the project were. However, the project manager did not label month 11 as one of the crucial

milestones. When considering this issue, the modeller noticed that the plaintiff not only documented individual financial milestones, but also the annual cash flow expected each financial year to 31 March. Months 1 to 11 represented the first year of cash flow from the project. It is likely that the company would wish to demonstrate the project as being as profitable as possible over the first year of its execution in its end of year accounts. Therefore, if possible, it may have been considered important to put in a little more effort in the final month of the year to enable the cash flow position to look as healthy as possible. Therefore, this may have been incorporated in the planning of the project as a slight increase in labour hours in month 11. Although the plaintiff could not verify this story, they believed it to be a possible explanation for the 1 month increase in engineering labour hours.

If the SD model were to be altered to reflect this argument, then an additional milestone would be placed at the end of month 11. This would increase the engineering work required by this date and therefore reduce the remaining work on the project, thereby altering the simulated data between months 11 to 15 closer to the actual data.

This plausible story is based on wider organisational issues, rather than specific project events and may therefore be a reason for its explanation not being immediately recognised by the claim team. Although the system boundary did not include wider organisational issues, such as pressure to demonstrate the profitability of the project, it may be the case that when considering reasons for discrepancies between the actual and simulated data, these types of issues may be relevant.

Project Outcome Data – Graph 3.2

Visually, and as concluded by the summary statistics, the simulated data for the actual run is capable of following the overall shape of the actual data well, but does not fit individual points particularly well. The simulated data appears as a smoothed version of the actual data. However, from a visual inspection of the graphs, the question that is likely to be asked is why the actual and simulated figures differ *so* much on a point-by-point comparison. The following provides plausible explanations for some of the differences between the time series. The explanations suggest that the differences are due to elements that have been omitted from the model.

- Months 19 to 22: When interviewing a program manager on the Castle project regarding the spikes that existed between months 17 and 22, he recalled that "I made a mistake – I let some key designers go too early". The explanation given was that "Finance was after me to reduce the wages bill". However, very shortly after the decision had been made the manager realised that he needed the staff back and therefore set about re-hiring staff. These actions are one plausible story that may account for the dramatic drop followed by a sudden increase in the actual labour hours over this period of time. The reason why the simulation model does not replicate this effect is that it would not account for the error in the decision-making. Based on certain decision rules, the simulation model anticipates that a certain amount of staff are still required on the project and would therefore not let these staff go as early as the program manager.

The system boundary for the simulation model did not include the pressure that the engineer was under from the finance department during that period of time. The decision on whether or not the model's system boundary would be expanded to allow for this event, would be based upon whether or not there is evidence that the cause of the decision-making could be traced back to a defendant or plaintiff trigger. If it could not be argued to have been caused by the defendant, then it may be considered to have been caused by a self-inflicted trigger. In the case of the program manager's decision-making, if the plaintiff's claim team were happy to support the argument that an overspend on the project (due to other triggers of D&D) forced the Finance dept to request that cuts be made and hence caused pressure on the program manager which forced him (and any other reasonable manager) to cut back on staff, then the system boundary may be expanded to allow for this. However, the program manager himself felt that the decision was a one-off event due to poor managerial judgement. It was therefore not modelled as part of the overall decision-making policies.

- Spikes in Project Outcome data including month 17: Most of the other differences between the actual and simulated data exist where a 1 month peak or trough occurs in the actual data. These peaks and troughs represent overtime or staff being increased or decreased for a 1 month period in response to a specific situation. Month 17 is the exception to this as it is the commencement of a *3* month period of overtime and additional staff which is not adequately matched by the simulated data. Discussions with those involved in the Castle project highlighted that this peak was due to pressure to meet an important milestone, which was set for the end of this 3 month period. This milestone was important both to the progress of the project as well as being a financial milestone for the plaintiff. However, no matter how the pressure to meet the milestone was modelled, the simulation model was unable to replicate the extreme peak that existed for month 17. The failure of the model to adequately replicate this peak indicates that the number of labour hours that the managers believed to be needed during this month were in excess of the model's forecasts. Perhaps more intense pressure to achieve milestones was being placed on the managers in the real system than compared to the model. This may indicate that the model does not include all the pressures that were placed on the managers, such as the pressure being felt from other parts of the organisation.

Attempting to replicate the sharp peaks and troughs shown in graph 3.2 through a simulation model is certainly not straightforward. Providing plausible reasons for the differences between actual and simulated data relies on information that can be gathered from those involved in the project.

As well as considering the behaviour of the time series, the overall labour hours were also considered. In the case of the Castle project, the difference was only a small proportion of the overall overrun in labour hours.

When reporting on the work that they have carried out, it is likely that the modellers cannot provide a defendable quantifiable model for *all* of the labour hours that have occurred on the project. They will therefore present a report based upon the labour hours that *can* be replicated. The report will discuss the replication of a reasonable proportion of the actual hours that occurred on the project and provide a range of labour hours that their modelling process can support. However, in the cases discussed in this thesis, the plaintiff's claim teams have put cases forward that request compensation for a number of hours that is greater than the range indicated by the model. The claim teams acknowledge the importance of negotiation in a litigation process. Therefore, their strategy is normally to request compensation for a larger number of labour hours than they would actually be willing to settle for.

The above discussions have focussed on the behaviour reproduction of the variable *engineering labour hours*. This variable is of particular importance to the type of litigation cases discussed in this thesis due to the need to:

- *Replicate over time the additional hours, over-and-above those that were contracted, which were required to carry out the project.*

However, this variable would not be the only variable included in the behaviour reproduction test. Examples of other variables that were also observed over time in the Castle project were *overtime hours* and *productivity levels*. Overtime hours give another example of a variable where reliable data was recorded over the life of the project. However, productivity levels are based upon subjective data gathered from those involved in the project. Although this data may not be considered to be as accurate as that available for overtime hours, the information gathered about the behaviour of productivity levels over time can still be monitored as part of building confidence in the simulation model.

8.3 Optimisation

In section 7.3.3, sensitivity analysis was highlighted as an important step when using SD to model D&D. The analysis used should consider both numerical and behaviour sensitivity for the models. Also, the process should involve both univariate and multivariate sensitivity analysis.

A technique that could be used as a part of sensitivity analysis is optimisation, but this has not been used in the models covered in this thesis. A summary of the literature regarding the use of optimisation in SD is given by Wolstenholme (1990) and Dangerfield and Roberts (1996).

Optimisation in SD has been used for two purposes; to determine the optimum policy to adopt in a system relative to some objective and to calibrate the model by fitting model variables to past data. When modelling D&D for litigation a consideration of these two purposes produces the following possible uses of optimisation:

Optimum Policy: To find the managerial policies that minimise the number of labour hours and time taken to complete the project. Keloharju and Wolstenholme (1989) demonstrate the use of optimisation on Richardson and Pugh's project model (1981). By optimising on both project time and cost, they were able to improve upon Richardson and Pugh's policy designs.

When discussing the extreme policy test in section 7.3.3, it was noted that a SD model could be used to test alternative managerial policies during litigation. For example, the defendant may claim that the project was managed incompetently. In addition to this the defendant may suggest alternative ways that the project could have been managed. In response to this, the plaintiff may wish to run the alternative policies through the model. This may be in an attempt to demonstrate that their management did not behave incompetently, rather that the alternative policies actually increased the actual labour hours accrued on the project or, if they did perform better, that the improvement was minimal. However, as an alternative to this, the plaintiff may wish to use optimisation to discover the optimum managerial policies for the project. This may be in an attempt to demonstrate that the actual managerial policies used on the project were close to the optimum policies for the model.

However, the modeller should be aware that the discovery of improved policies may not be particularly helpful in a litigation environment. The existence of optimum policies may not mean that they should have actually been used on the project. During the project, managers do not have a simulation model to test policies on or to determine the best policies to use. Instead, they base the policies they should use on experience and judgement. The discovery of the optimal managerial policies, through the use of optimisation, may therefore provide misleading information.

Data Fitting: To fit variables to historical data. When simulating the project, either as budgeted or as its actual outcome, there will be a number of input variables where uncertainty exists in their value. It may be the case that when gathering data, rather than getting a specific value, a range of values is gathered. The modeller may then, for example, choose the average of the range to input into the model, using the range as an input to a series of sensitivity analysis runs. Optimisation may be a method of illustrating that the plausible ranges given for the variables can produce simulated output that fits the behaviour of the actual data.

This process may also help with the plausibility of the ranges of the more subjective variables. For example, the objective when optimising in this way would be to minimise the difference between the actual and simulated data. However, if the smallest difference between the two time series is not satisfactory then the modeller may want to consider the plausibility of the ranges gathered for the subjective data and whether or not more information should be gathered about the variable.

If optimisation is used in this way, then historical data is required to fit the simulated data. However, this may cause a problem with the process. If the modeller believes that self-inflicted triggers existed in the project, but they are not explicitly modelled, then the question to ask is what historical data should be used to fit the data to? Only in cases where the number of labour hours due to self-inflicted triggers is minimal and where the model is believed to include the majority of the triggers that occurred in the project could this form of data fitting be used.

Optimisation may be a useful method to use to illustrate how the model can fit historical data, however it could not be a replacement for sensitivity analysis. If plausible ranges of variable values are obtained from experts, then the full range of these values needs to be tested on the model, not simply the one value that best fits the historical data. The defendants could argue for any of the values contained in the ranges given and hence the plaintiff's need to have knowledge of the consequences of any of these values being input into the model. For example, the defendants are likely to argue for the most conservative values for any of the variables. If they win their argument, then the plaintiff needs to be aware of what the lowest labour hour output from the model would be.

Optimisation has not been used in the projects discussed in this thesis. One reason for this is that it is not seen as a replacement for sensitivity analysis and this process would still need to be undertaken. Also the models used have been relatively large, containing hundreds of variables. To be able to use optimisation to fit data, a powerful tool would be required. Also, as mentioned above, there is the issue of what historical data to use to fit to the simulated time series.

The advantage of using optimisation to data fit when modelling D&D for litigation is to have the capability of illustrating how well the gathered data can produce a model that closely fits the actual data. This would provide help in demonstrating that the triggers caused by the defendant led to the overrun in hours seen on the actual project. However, before this benefit can be gained the modeller needs to be certain that the correct historical data can be used for the data fitting i.e. that the number of labour hours due to self-inflicted triggers is minimal and the model is believed to include the majority of the triggers that occurred in the project.

8.4 Summary of Chapter

When exploring confidence building tests in chapter 7, the behaviour and numerical reproduction test was highlighted as playing a particularly important role in the modelling process. Experience on litigation cases has shown that the behaviour reproduction test is a powerful test in convincing the various audiences that the model adequately replicates reality.

When considering behaviour reproduction in SD models of a single project, the importance of the test was highlighted by Cooper (1980, p27) when discussing a SD model that explored the reasons for cost and time overruns in a litigation environment. Cooper stated that the focus of the model tests was to ensure that "...the model replicated quite accurately the vast amount of detailed information on the programs' histories." The other tests were used "In order to assure... that the model performed correctly for the correct reasons..." (Cooper, 1980, p27). Literature that recounts tests that have been carried out have focussed on a visual inspection of key variables or the use of MSE summary statistics. When using these model tests on SD models of D&D in a litigation environment, the MSE summary statistics of key variables need to be used in conjunction with visual inspection of the time series so that plausible reasons can be produced for the differences between actual and simulated data.

This chapter highlighted that there may be difficulties involved in attempting to reproduce behaviour in a SD model in a litigation environment. An important issue to consider is whether or not the simulated time series *should* replicate the actual time series. Two of the reasons for this occurring were highlighted as:

- The existence of self-inflicted D&D. If this was believed to occur on the project, then the preferred way in which to deal with this was seen to be to compare the actual time series with the simulated times series which excludes self-inflicted triggers. Plausible stories are then required to explain the differences between the actual and simulated time series.
- The model does not include all elements of the project, only those relevant to D&D. It may be believed that some elements of a project only contribute towards a direct cost overrun in the project. This needs consideration to ensure that the correct actual data is being used for comparison with the simulated data.

When carrying out a behaviour reproduction test, a consideration that has not been covered by the literature was the choice of the unit of time between time series data points. For example, simply because a company records data monthly does not mean that this is the most appropriate time unit to choose when comparing actual and simulated time series. When graphing a variable output, the time unit chosen will have an effect on the resulting pattern of behaviour. A modeller therefore needs to make a decision on the time unit that should be used to analyse the data. This is of particular interest when making point-by-point comparisons such as in the MSE statistics since these points will differ depending upon the unit of time between data points. For example, the disaggregation of monthly figures into weekly figures may highlight patterns that are not seen through the monthly data. Even if the modeller feels that the monthly data is more suitable for illustration purposes, it was suggested that the weekly data should be checked to see if the results it produces in the behaviour reproduction tests enables additional insights to the analysis.

Optimisation was discussed as a possible extension to confidence building tests. Some limited uses were seen to exist for the plaintiff to optimise the managerial policies in the project in an attempt to demonstrate that the actual policies taken were close to this and hence management behaved competently during the project. However, a wider use was seen to exist in the use of optimisation to calibrate the model by fitting model variables to past data. This technique was seen to be a way of illustrating that the plausible ranges given for the variables can produce simulated output that fits the behaviour of the actual data and to also help with the plausibility of the ranges of the more subjective variables. However, only in cases where the number of labour hours due to self-inflicted triggers is minimal and where the model is believed to include the majority of the triggers that occurred in the project, could this form of data fitting be used.

This chapter focused on numerical as well as behavioural reproduction. In placing more emphasis on point-to-point matching of data, a question to consider is whether the SD models discussed in this thesis are using the SD tools, but not the SD methodology. For example, Meadow (1980, p55) states that "If the problem addressed by a computer model reflects an open system view, if the model variables are observables, if the validation procedure involves detailed matching with historic data, then I would say the model is in the econometric paradigm, no matter what mathematical technique or computer language is used. If the problem is centred on generic dynamic behavior of a mostly closed system, if the variables include motivations and goals, if the validation includes assessment of the realism of the model structure, then it is a system dynamics model". To place the work discussed in this thesis in terms used by Meadows, the problems are:

- Centred on dynamic behaviour of a mostly closed system.
- The validation procedure involves some detailed matching with historic data.
- Some of the model variables are observables and some include motivations and goals.

Therefore, Meadows would not consider the use of SD in modelling D&D for litigation to be clearly defined as either using or not using SD methodology. The use of SD in this way focuses on modelling the underlying dynamic structure of a project in a more detailed manner than SD has traditionally been used. However, the conclusions in chapter 5 taken from the exploration of the SD literature do not rule out the use of SD to model D&D for litigation. SD may be being used in a different manner from which it has traditionally been used and there may be limitations to its application in this way. However, if the modeller takes account of all the limitations of using SD in this particular environment and it can still be used in a manner which is of benefit to decision-makers, then there is no reason why SD cannot be used to analyse D&D for litigation.

CHAPTER 9: CONCLUSIONS AND FURTHER WORK

9.1 Introduction

The overall aim of this thesis was to explore the role of SD in analysing D&D for litigation.

Litigation is an area where an organisation can gain or lose millions of dollars. For this simple reason, it is important for the organisation to ensure that it has the best support during the process to optimise its chance of gaining compensation. Mathematical modelling can play a role in supporting organisations during litigation. However, if it is to do so, then a thorough understanding of how it can be used is essential.

SD is one form of mathematical modelling that has been used to support organisations as part of litigation. The minimal literature in this area describes its use in modelling large engineering projects (Cooper 1980, Weil and Etherton 1990, Ackermann et al 1997). Its role in these few litigation cases has been to model D&D that occurred during the projects. Reported claim settlements indicate that the use of SD to model D&D has contributed towards overall settlements of up to US$440m on any one claim.

However, there is a lack of literature in the use of SD to model D&D for litigation and this is a concern. This lack of literature is not due to the lack of use of SD to model D&D for litigation. PA Consulting Group have used SD in the modelling of D&D in many litigation cases; 35 in the last 35 years (Graham 2000a). From their experiences in these and other projects, a number of papers have been published (Cooper 1993a, 1993b, 1993c, 1994, 1999; Graham 2000b). However, the focus of these papers has mainly been on project management issues such as the so-called *rework cycle* that can occur in projects. It is perhaps a pity that they have not published more on their experiences of the use of SD in modelling D&D for litigation. However, this is perhaps understandable since, as a consulting firm, they will view their knowledge as commercially valuable.

Due to the lack of literature in the use of SD to model D&D for litigation, this thesis aimed to improve this situation. By gaining a better understanding of the role that SD can play in the analysis of D&D for litigation it is hoped that more informed support can be given to organisations to assist them through litigation processes.

In structuring these explorations, the work was split into two separate sections:

1. Explorations of the SD literature for any criteria that could help to assess whether or not SD should be used in the analysis of D&D for litigation.

2. Explorations of the purposes of modelling D&D for litigation to gain an understanding of the issues involved in using SD to analyse D&D for litigation.

The author has had the experience of being involved as both researcher and consultant in modelling D&D in two large litigation cases, spanning four years of work. This privileged opportunity, which is rarely made available to researchers, provided valuable data for the explorations covered in the two sections.

Through the explorations, lessons have been gained that will contribute to both the knowledge of SD and the analysis of D&D. Therefore, any modeller either using SD or wishing to analyse D&D, will benefit from the conclusions reached in this thesis.

9.2 Conclusions from the Explorations of the SD Literature

9.2.1 Lessons for SD

In searching the SD literature for potential criteria to use to assess whether or not a situation should be modelled using the SD modelling approach, there appeared to be a distinctive lack of material. This is because many authors in the SD literature avoid a detailed discussion on the suitability of SD as a modelling approach. For example, Forrester (1968b) appears to believe that SD can be used to model virtually all organisational situations. This is concluded from the very narrow set of limitations that he defines for SD.

This situation was understood by considering Vennix's (1996, p104) point of view on the difficulty of defining a set of criteria to assess the suitability of SD to model a situation. He states that one of the most difficult questions asked, even to an experienced model-builder, is when to use SD. This implies that he believes there are situations when SD should not be used, but it is difficult, practically, to assess the criteria to be used. Nevertheless, modellers need to be able to make a reasonable judgement on this question. In particular, when modelling D&D for litigation, they need to be able to defend their choice of modelling approach. This could potentially be in a courtroom. Therefore, no matter how difficult it is to assess the suitability of the use of the SD approach, it is a task that has to be undertaken.

When criteria *are* detailed in the literature some of them are introduced with minimal explanation. This was found to be particularly true when exploring Forrester's and Flood and Jackson's criteria. This may be a reflection of the difficulty involved in arriving at such criteria. However, the lack of detail in the criteria left the reader with the impression that the set of situations which SD could be used as a modelling approach is very large. However, criteria introduced by Coyle (1977) are an exception to this. They provide a detailed, practical set of criteria that formed the basis of useful explorations.

However, the difficulty in providing detailed criteria may be partly due to the fact that a modeller may only be familiar with SD, or at least more familiar with this tool than other types of modelling tools, and thus be inclined to use it automatically. However, the modeller then has to be aware of the danger of translating *any* problem into a SD problem.

It should be noted that SD is not the only modelling approach where a lack of detail applies when defining criteria regarding the approach's suitability in modelling a situation. In any form of modelling the choice of one approach over another is not a straightforward process. It is not simply in the field of SD that we find modellers demonstrating a bias towards one modelling tool, it is a general modelling issue within the Operational Research field.

Although there are many issues involved in providing criteria to assess the suitability of SD to model a situation, a summary of the most informative criteria was produced. This should enable any modeller to assess whether or not SD could be used to model D&D for litigation.

Summary of the Most Informative Criteria

In producing this set of criteria, it was noted that it is extremely important to consider the purpose of the modelling process when forming a decision on whether or not SD is an appropriate method to use to model a given situation. This was generally not captured by the criteria that existed in the SD literature but formed an important part of the revised set of criteria.

The intention of the following set of criteria was to provide a modeller with a comprehensive, practical method with which to assess the appropriateness of SD to model a D&D for litigation. The criteria are as follows:

1. *Does the situation contain feedback loops and are they of importance to the purposes of the study?*

2. *Does the situation exhibit changes through time and are they of importance to the purposes of the study?*

3. *Practically, can it be done?*
   In particular:
   (a) *Can the project be visualised at an early stage in terms of stocks and flows?*

   (b) *Is there sufficient, reliable information and data available which will enable a model to be populated to such a level that the plaintiff's case is in a stronger position when compared to the plaintiff not having the model?*

When applying these criteria it should be noted that they are not intended to be used as formal tests. Instead, they should be used as guidelines to assist the discussion between those who are involved in deciding whether or not SD should be used to model D&D for the litigation case. Rather than a straightforward 'yes'/'no' response, the result of applying these criteria is likely to be an indication of the *extent* of suitability of SD to model a situation.

The above criteria were chosen as the most powerful questions to be asked at the beginning of a study. If any of these questions do not receive a very positive response, then it is likely that SD would not be an appropriate modelling approach for the situation. It should be noted that these criteria do not answer the question whether SD would be the *most* appropriate modelling approach.

It should be noted that the above criteria are still very subjective and are dependent upon the worldview of the modeller. An experienced SD modeller, who has a preference to view the world in terms of feedback and SD, will more easily be satisfied that *the criteria hold* compared to modellers with a different worldview. It is unlikely that any set of criteria could be defined which did not include the need for a certain amount of subjectivity and thus separate the decision from the modeller's personal view.

If the criteria above were used as standards defined by SD to assess whether or not a SD approach can be used to model D&D in the projects discussed in this thesis, then they would suggest that SD could be used.

9.2.2 Lessons for Modelling D&D for Litigation

When the various criteria that were explored were applied to the modelling of D&D for litigation, insights were gained regarding the modelling of D&D. These insights provided an improved understanding of D&D for modellers wishing to model it for litigation.

The Type of System being Modelled

To gain an appreciation of the type of system being modelled an exploration of Flood and Jackson's criteria led to a labelling of the participants and the system when modelling D&D for litigation. The result was that the participants were believed to have a unitary relationship. It was noted that this is important for a successful litigation process since cohesiveness between the participants in the plaintiff organisation is desirable to enable their case to have the best possible chance of success. The systems labelling was not so straightforward. The conclusion reached was that there is evidence to suggest that the systems lie closer to a simple categorisation with the potential for a few elements to display strong complex characteristics. It is important to note that this conclusion was based upon the purpose the observer of the system has for considering the system and is therefore a subjective measurement.

When exploring the extent of the unitary relationship between the participants of the plaintiff organisation, the following issues can be highlighted as areas where cohesiveness between the participants of the plaintiff team may not occur:

- The interests of each individual of the plaintiff's claim team are likely to be in line with the overall common interest of gaining compensation from the defendant.

However, individuals may have different opinions about what they would prefer as a detailed outcome of the claim.

- In both the Castle and Pirate projects, difficulties arose with witnesses that were being used from the plaintiff's organisation. The beliefs of some witnesses about the reasons for the failure of the projects were different to other members of the claim team. If this occurs, the plaintiff's claim team needs to consider the causes of these differences. The claim team then needs to consider whether each of the differences need further research or whether they are due to a lack of appreciation of the system by the witnesses.
- Generally, when a decision has to be made on how the plaintiff's claim team should move forward, the member of the team with the most knowledge in the area of discussion is likely to direct the way forward. However, differences of opinions can occur when, for example, decisions overlap with different individuals' knowledge. Of course, the reduced number of members of the plaintiff team does help in reducing the number of differences that occur. However, if we consider the power of each of the individuals in the claim team, in both the Castle and Pirate projects the senior manager from corporate office was the individual who had ultimate responsibility for the final decisions. Therefore the opinions of each of the individuals in the claim team are not equally weighted when making decisions.

When exploring the extent of the complexity of the system, the following issues were highlighted as areas where there is potential for increased levels of complexity:

- The triggers of D&D. Complexity arises in the system since the occurrence of the triggers are either unexpected or the way in which they occur are unexpected. Also, a number of D&D triggers may cause a portfolio effect where the triggers lead to an outcome that is not the effect of any individual trigger, but is caused due to the occurrence of the specific portfolio of triggers.
- Sub-systems of the overall system can cause complexity by pursuing their own goals. For example, workers on a project may leave the project to increase their salary or to seek promotion. Although this will be in line with their own goals and ambitions, it may prove detrimental to the goals of the project. This occurs since the project is losing the learning that the individual has gained about the project.

- Behavioural influences will exist in any large project. Decision-makers are individuals with their own experiences, beliefs and goals. Each of these factors will affect the decisions they make. For any individual project, a consideration has to be made as to *the extent* that the decisions are affected by these factors.

Modelling Approach

The explorations in chapter 3 concluded that a systems analysis approach is required if all of the effects of D&D are to be taken into account when modelling D&D. However, in using a systems analysis approach there were some concerns over how easily litigation audiences are able to understand the approaches and therefore be convinced by them. This is an area where additional work is required.

When using a systems analysis approach, the feedback structure of a system is an important area of consideration. For the projects explored in this thesis, the majority of feedback was seen to arise from managerial actions in response to a delay in the project. The delays in projects were seen to originate from the triggers of D&D. The triggers were categorised based upon the type of disruptions they caused.

Disruption	**Likelihood of immediate managerial action being taken in response to disruption**
Halts the flow of work	MOST LIKELY
Additional tasks	
Tasks take longer than expected	LEAST LIKELY

Table 9.1 - Categorisation of the Managerial Actions Taken in Response to the Disruptions from D&D Triggers – Extract from Tables 3.1 and 3.2

The categories shown in table 9.1 can be used as a useful assessment of the level of feedback there is likely to be in a system. The more disruptions that immediately cause a halt in the flow of work, the more immediate managerial responses there will be and hence the more disruptions from managerial actions. It was therefore concluded that it is more appropriate to model the feedback structure in a project if managerial actions were likely to be immediately taken in response to disruptions in order to compress the project. This process therefore provides a means of assessing the importance that feedback plays in the outcome of a project.

Qualitative and Quantitative Modelling of D&D

The explorations highlighted an issue with the movement between qualitative and quantitative modelling when considering D&D. A common concept in engineering projects is *cross-impact*. When considering how this has been modelled in previous quantitative models it was concluded that it has been incorrectly represented. The problem occurs since the *rework/cross-impact* loop in the qualitative model represents an infinite knock-on effect between the two concepts. However, in reality this is not the case. Therefore, rather than translating this loop as an infinite knock-on effect in the quantitative model, the modeller needs to consider the average number of times the loop sustains itself and attempt to gather data and populate the model to reflect this. This illustrated a limitation in the qualitative modelling where a feedback loop that only sustains itself for a finite number of times around the loop cannot be clearly represented.

9.3 Conclusions from the Explorations of the Purposes of Modelling D&D for Litigation

The purposes of modelling D&D for litigation were defined as follows:

1. Demonstrate that a part of the time and cost overruns were caused by D&D through particular triggers of D&D.
2. Replicate over time the hours due to D&D which were over-and-above those that were contracted, but were required to carry out the project.
3. Demonstrate the extent to which the client management of the project was reasonable and the extent that overruns could not have been reasonably avoided.

4. All the above have to be demonstrated in a way which will be convincing to the several stakeholders in a litigation audience.

In exploring these purposes, lessons were arrived at regarding the modelling approach. Also, the models' audiences were seen as a vital part in the modelling process and therefore conclusions were also drawn about their participation in the process.

9.3.1 Qualitative and Quantitative Modelling

The use of the qualitative cause map and, if required, an influence diagram with the quantitative SD model was concluded as the recommended approach to be able to model the causal links between an event and outcome. However, it was noted that achieving a relatively transparent quantified link between the events and outcomes is very difficult when modelling D&D for litigation.

When considering the quantification of the outcomes of D&D, some limitations of the SD modelling approach were identified as follows:

- SD can prove inadequate when attempting to capture the more detailed, operational issues of project management. This will cause difficulties when attempting to replicate project delays.
- The lack of accuracy of some of the parameters used as inputs means that a modeller should provide a range of results as an indication of the range of additional labour hours incurred on a project rather than only presenting a single cost result.
- The D&D that occurs on one project that cause costs to accrue on other projects are normally omitted from the SD model. The reasons for the omission of these costs are due to the general difficulty in modelling them.

9.3.2 The Audiences for the Models

There are a number of audiences that need be convinced that the model replicates reality. These are identified as the modeller, the plaintiff's claim team, the defendant's claim team, expert modellers, a judge and potentially a jury. Each of these audiences have different aims whilst playing a part of the litigation process.

To enable each of the audiences to make a judgement about whether or not they believe the model to replicate reality convincingly, they each need to gain a basic understanding of the mechanics of the model. This means that the model should be *transparent to lay people* to enable each of the audiences to gain a reasonable understanding of it.

When considering how transparent SD models are to lay people, the following difficulties were highlighted with SD models for litigation:

- The models are usually relatively large.
- There is a requirement for lay people to use quantitative reasoning rather than only qualitative reasoning.
- There is a lack of time for the audience to be exposed to the SD model.
- Individuals have a difficulty in gaining an appreciation of the overall behaviour of a dynamic system.

Each of these difficulties constrains the level of transparency that particular audiences can reach with the model.

A lack of transparency in the models for some of the audiences has meant that the lawyers involved in the projects discussed in this thesis have concluded that they would be unlikely to use the models in a courtroom. This is an area where further research could be focussed to attempt to increase the transparency of the models, so that they could become more useful during the litigation process.

As well as requiring transparent models, the first audience for the model i.e. the modeller was highlighted as having a particularly rigorous process for the model to go through in order to be convinced that the model replicates reality convincingly. Due to this rigorous process, the issues associated with this process were dealt with in detail.

9.3.3 Replicating Reality Convincingly for the Modeller

Confidence building tests can be used to assess whether or not the modeller believes that the model adequately replicates reality. Relevant tests for modelling D&D for litigation were discussed in chapter 7. In particular, both *behaviour* and *numerical* sensitivity analysis were seen to be important when modelling D&D for litigation.

During numerical sensitivity, the overall number of labour hours is an important variable to check. This plays a very important role in the modelling process in a litigation environment as it contributes directly towards the plausible ranges of compensation that the plaintiff will request in the claim.

Experience of the use of the behaviour reproduction test in litigation environments has shown that it can also be very important. It has been seen to be a powerful test in convincing the various audiences that the model adequately replicates reality.

The role of numerical reproduction places particular demands on the SD approach. When using SD where the modelling purpose is to improve the understanding of the system and thereby improve the way in which the system is managed in the future, the replication of the general behaviour of the system is normally sufficient. However, when attempting to replicate the hours on a project due to D&D, the matching of individual points of data becomes more of a concern. This is especially true when the model may be the basis of a claim for a large amount of money in a litigation environment. Therefore, the modelling process may become more detailed than in other uses of the SD approach. For example, variables may need to be disaggregated to a reasonably detailed level.

Although behaviour and numerical reproduction are important tests, there may be difficulties involved in carrying them out in a SD model for litigation. An important issue to consider is whether or not the simulated time series *should* replicate the actual time series. For example, the existence of self-inflicted D&D in a project may mean that the model will not include all elements of the real system. In this case, the preferred way in which to deal with this was seen to be to compare the actual time series with the simulated times series which excludes self-inflicted triggers. Plausible stories are then required to explain the differences between the actual and simulated time series.

Another example of a situation where the simulated time series should not replicate the actual time series is where the model may not include all elements of the project, only those relevant to D&D. It may be believed that some elements of a project only contribute towards a direct cost overrun in the project. This needs consideration to ensure that the correct actual data is being used for comparison with the simulated data.

When carrying out a behaviour reproduction test, a consideration that has not been covered by the literature was the choice of the unit of time between time series data points. When graphing a variable output, the time unit chosen will have an effect on the resulting pattern of behaviour. A modeller therefore needs to make a decision on the time unit that should be used to analyse the data. This is of particular interest when making point-by-point comparisons since these points will differ depending upon the unit of time between data points. For example, the disaggregation of monthly figures into weekly figures may highlight patterns that are not seen through the monthly data. Even if the modeller feels that the monthly data is more suitable for illustration purposes, it is suggested that the weekly data should be checked to see if the results it produces in the behaviour reproduction tests enables additional insights to the analysis.

Optimisation was discussed as a possible extension to confidence building tests. Some limited uses were seen to exist for the plaintiff to optimise the managerial policies in the project in an attempt to demonstrate that the actual policies taken were close to this and hence management behaved competently during the project. However, a wider use was seen to exist in the use of optimisation to calibrate the model by fitting model variables to past data. This technique was seen to be a way of illustrating that the plausible ranges given for the variables can produce simulated output that fits the behaviour of the actual data and to also help with the plausibility of the ranges of the more subjective variables. However, only in cases where the number of labour hours due to self-inflicted triggers is minimal and where the model is believed to include the majority of the triggers that occurred in the project, could this form of data fitting be used.

9.4 A Final Conclusion

By using SD to model D&D for litigation, SD is being used in a different environment to which it has traditionally been used, for example for policy analysis. This raises issues and limitations for SD being used to model D&D for litigation. However, if the modeller takes account of all the limitations of using SD in this particular environment and SD can still be used in a manner which is of benefit to decision-makers, then there is no reason why SD cannot be used to model D&D for litigation. As long as the plaintiff's case is in a stronger position with the SD model than compared to the plaintiff not having the model, then this is a good enough reason for SD to be used to model D&D.

9.5 Limitations of the Research

The conclusions to any research can only be regarded as being useful when the limitations of the research are understood.

The main limitation to this research is that the conclusions that have been reached are based upon the explorations of only two projects. These were both large engineering projects that were concerned with the modification of aircraft. Therefore, there may be a limited area of use for the conclusions. It was noted in chapter 1 that gaining access to projects as part of a litigation process is extremely difficult. Therefore, getting access to other projects is likely to prove very difficult. However, if access to other projects is possible, then further research could be used to validate the findings in this work. If access could be gained to engineering projects that are concerned with other types of products, then conclusions beyond the narrow view of aircraft modification could also be gained. However, it should be noted that the author has been associated with six further litigation cases (although not integrally involved in any of these) that involved the production of different types of engineering products. This association has not led to the knowledge of any information that would contradict any of the conclusions reached in this thesis.

The limited number of projects explored in this research leads to an inability to generalise the various conclusions. For example, during the research only a limited number of different audiences were considered. Therefore, there is a possibility that the conclusions may not hold for other individuals who form part of another model's audience. Also, the research has highlighted some problems that can occur when modelling D&D in a litigation environment. However, due to the limited exposure to projects, this list has certainly not been exhausted.

Both the projects discussed throughout this research were carried out for the same organisation, although for different divisions of that organisation. The findings may therefore be biased towards the way in which this particular organisation deals with litigation cases. For example, they may record particular data, the people involved in the claim team may not be similar to other organisations, their attitude towards litigation may be different to other organisations etc. However, in defence of this point, the other projects that the author has been associated with were carried out for different organisations. Knowledge of these projects has not led to any contradictions to the conclusions reached in this thesis. Nevertheless, an opportunity to be more integrally involved in other projects for *other organisations* would provide the author with a chance to give more rigour to the findings of this research.

The subjective view of the modeller has been mentioned on a number of occasions during this thesis and cannot be ignored when considering the limitations of the research. For both of the projects discussed in this thesis, the modellers used to analyse D&D were the same. This means that, for example, when interviewing witnesses during the creation of a cause map, or when translating a draft claim document into a cause map, or when translating the qualitative diagrams into a SD model, the same modellers and hence the same personal views existed. Since individuals have differing personal views, intervention from a different modeller may have created different models. These may have led to different conclusions being reached. This is a limitation with many pieces of research and other than gathering data from many different projects with different modellers, would be impossible to avoid.

The nature of the research meant that it was not possible to follow a particularly defined and structured approach to the work. The empirical data came from the researcher's participation in and observation of two litigation cases. Although chosen as the most appropriate method to use in this research, it is not without its' limitations. The observations by the researcher are biased by the researcher's worldview. One individual may perceive a situation very differently to another individual. Detailed note taking and triangulation of information from different sources can help to reduce this subjectivity, but it is unlikely that this can ever be eliminated. Hence, this should be borne in mind when using any of the conclusions from the research.

The set of criteria produced to assess whether or not SD is suitable to model a project with D&D for litigation is a further area where limitations exist. Any conclusions gained from the use of the set of criteria could be very subjective and dependent upon the worldview of the modeller. An experienced SD modeller, who has a preference to view the world in terms of feedback and SD, will more easily be satisfied that the criteria hold compared to modellers with a different worldview. However, it is unlikely that any set of criteria could be defined which did not include the need for a certain amount of subjectivity and thus separate the decision from the modeller's personal view. A further limitation of the criteria is that, although giving us a sense of whether or not SD could be used to model D&D for litigation, they do not inform us whether SD is the *most* appropriate modelling approach.

Although this section has highlighted the limitations of the conclusions reached in this thesis, the research is robust within these limitations and hence the usefulness of the conclusions should not be forgotten. There is a limited amount of literature regarding the role of SD in analysing D&D for litigation and any contribution towards increasing this should not be ignored. Even although the research in this thesis may have only moved the knowledge in this area forward to a limited degree, the important point is that it *has* moved the knowledge forward. It is hoped that future work in this area can further increase the knowledge about the role of SD in analysing D&D for litigation.

9.6 Further Work

The work covered in this thesis represents only the tip of the iceberg in potential research that could be carried out when considering the role of SD in the analysis of D&D for litigation. There are many different avenues the author would therefore like to follow on completion of this thesis.

One avenue the author would like to pursue is to continue gathering data in an attempt to reinforce or further the conclusions arrived at in this work. As mentioned in the limitations of the research, further litigation cases could be used to validate the findings in this thesis. If possible, involvement in additional projects for other organisations would also be beneficial. However, the difficulties involved in gaining access to other projects have been noted.

The aim of the revised set of criteria from the first half of the explorations was to provide modellers with a practical set of criteria to enable them to make an assessment of whether or not the D&D for a particular project can be modelled using SD for litigation purposes. However, a possible extension of this is to create a set of practical criteria that any modellers can use to assess the suitability of SD for *any* situation.

Some specific areas have been highlighted throughout this thesis as areas where attention could be placed in order to attempt to improve the modelling process undertaken for decision-makers. For example, it was concluded that:

- SD can prove inadequate when attempting to capture the more detailed, operational issues of project management. This will cause difficulties when attempting to replicate project delays. Therefore, attention could be focussed towards attempting to improve this limitation of SD.
- The lack of accuracy of some of the parameters used as inputs was highlighted as an issue during the modelling process. Perhaps attention could be placed on methods to improve the accuracy of these parameters.
- The D&D that occurs on one project that cause costs to accrue on other projects are normally omitted from the SD model. The reasons for the omission of these costs are due to the general difficulty in modelling them. Again, attention could focus on improving this situation by researching methods of including the assessment of the costs that accrue to other projects.

The models' audiences played an important role in the explorations in this thesis. However, more could be done to appreciate their individual aims. Modelling for litigation presents a unique opportunity to consider the potential for a model to be able to convince many different audiences with differing objectives. For this reason, a fuller appreciation of the different audiences would aid an improved understanding of each of their needs. This improved understanding could help modellers build models that are better at meeting the purpose of replicating reality convincingly for all the audiences.

When considering the models' audiences, concerns were raised in chapter 6 regarding the difficulties involved in getting litigation audiences to understand SD models used for litigation and therefore be convinced by them. The author would therefore like to research further into this area in an attempt to find ways of improving this situation. In doing so, this may also help to convince the lawyers that SD can be used beyond negotiation, since the models will be more transparent to the audiences.

An area that may prove useful to SD modellers, since it is not touched upon in the SD literature, is the choice of the unit of time between time series data points when carrying out a behaviour reproduction test. It was noted that when graphing a variable output, the time unit chosen would have an effect on the resulting pattern of behaviour. This is of particular interest when making point-by-point comparisons since these points will differ depending upon the unit of time between data points. Even if the modeller feels that monthly data is more suitable for illustration purposes, it has been suggested that weekly data should be checked to see if the results it produces in the behaviour reproduction tests enables additional insights to the analysis. Unfortunately, weekly data was not available on either of the projects that the author has been involved in. However, this is an interesting area to pursue, since the SD literature does not cover it. The author would therefore like to attempt to gain relevant data to test the ideas surfaced further.

When attempting to reproduce behaviour, optimisation was discussed as a possible extension to confidence building tests. This could be used to calibrate the model by fitting model variables to past data. A use for this technique was seen to be to illustrate that the plausible ranges given for the variables can produce simulated output that fits the behaviour of the actual data and to help with the plausibility of the ranges of the more subjective variables. This is a technique that has not been fully discussed in the context of modelling D&D for litigation. The author would therefore like to do further work to identify the possibilities of the use of this technique in this environment, to discover what it can bring to the modelling process.

The author finds the modelling of D&D for litigation a fascinating area of research and hopes to continue to work in this area as both a researcher and a consultant. The general lack of literature regarding the modelling of D&D for litigation is something that the author would like to help to change. This section has shown that there are many different areas of research that could help to achieve this. In following this research it is hoped that an improved understanding of the models that are capable of supporting decision-making in a litigation process is gained. This will no doubt keep the author busy for many years beyond the completion of this thesis.

GLOSSARY OF PROJECT SPECIFIC TERMS

CCR - Configuration Change Request – A request by the customer or contractor to change a design. This is normally raised due to a problem in the current design.

CFE – Customer Furnished Equipment - Equipment required by the contractor from the customer in order that production work can be progressed

CFI – Customer Furnished Information - Information required by the contractor from the customer in order that design work can be progressed.

CO – Change Order -This occurs when the customer requests a change in the work that is being undertaken by a contractor.

Corporate Learning - One element of the learning expected to be gained between products in production. This represents the learning expected to be retained by the organisation rather than by the individual workers.

ECP – Engineering Change Proposal - This is raised when either the customer or contractor requests a change in a design. This document may be raised due to the occurrence of a CCR.

NR - Non-Routine - This represents additional work that is required on a product in order for it to meet the specifications laid down in the project contract.

O&A - Over and Above - Another term used to represent a Non-Routine.

PTR – Problem Trouble Report - This is a report that is raised in production when additional information is required from the client in order that the production work can proceed.

REFERENCES

Abdel-Hamid TK and Madnick SE (1991). *Software Project Dynamics: An Integrated Approach.* Prentice-Hall Inc: Englewood Cliffs, NJ.

Ackermann F, Eden C and Williams TM (1997). Modeling for Litigation: Mixing Qualitative and Quantitative Approaches, *Interfaces* 27(2): 48-65.

Ackoff RL and Sasieni MW (1968). *Fundamentals of Operations Research.* Wiley: New York.

Ansoff HI and Slevin DP (1968). An Appreciation of Industrial Dynamics. *Management Science* 14(7): 383-397.

Argote L and Epple D (1990). Learning-Curves in Manufacturing. *Science* 247: 920-924.

Barlas Y (1989). Multiple tests for validation of system dynamics type of simulation models. *European Journal of Operational Research* 42(1): 59-87.

Barlas Y and Carpenter S (1990). Philosophical roots of model validation: Two paradigms. *System Dynamics Review* 6(2): 148-166.

Barlas Y (1996). Formal Aspects of model validity and validation in system dynamics. *System Dynamics Review* 12(3): 183-210.

Barlas Y (1998). *President's Address to the System Dynamics Society.* 1998 System Dynamics Conference, Quebec City, Canada. http://www.albany.edu/cpr/sds/preaddress98.htm

Beer S (1979). *The Heart of Enterprise.* Wiley: Chichester.

Bennett P, Ackermann F, Eden C and Williams T (1997). Analysing litigation and negotiation: a case study of combined methodology. In: Minger J and Gill A (eds). *Multi-Methodology: The Theory and Practice of Combining Management Science Methodologies.* Wiley: Chichester.

Burrell G and Morgan G (1979). *Sociological Paradigms and Organisational Analysis: Elements of the Sociology of Corporate Life.* Gower: Aldershot.

Checkland P (1981). *Systems Thinking, Systems Practice.* Wiley: Chichester.

Cooper KG (1980). Naval Ship Production: A Claim Settled and a Framework Built. *Interfaces* 10(6): 20-36.

Cooper KG (1993a). The Rework Cycle: How it really works… and reworks... *PMNetwork* February: 25-28.

Cooper KG (1993b). The Rework Cycle: Why Projects are Mismanaged. *PMNetwork.* February: 5-7.

Cooper KG (1993c). The Rework Cycle: Benchmarks for the Project Manager. *PMNetwork* March: 17-21.

Cooper KG (1994). The $2,000 Hour: How Managers Influence Project Performance Through the Rework Cycle. *Project Management Journal* 25(1): 11-24.

Cooper KG (1999). The Power of the People. *PM Network* July: 43-47.

Coyle RG (1973). On the scope and purpose of Industrial Dynamics. *International Journal of Systems Science* 4(3): 397-406.

Coyle RG (1977). *Management System Dynamics*. Wiley: Chichester.

Coyle RG (1985). Representing Discrete Events in System Dynamics Models: A Theoretical Application to Modelling Coal Production. *Journal of the Operational Research Society* 36(4): 307-318.

Coyle RG (1986). Comment on 'System dynamics and operational research: An appraisal'. *European Journal of Operational Research* 23: 403-406.

Coyle RG (1996). *System Dynamics Modelling: A Practical Approach*. Chapman & Hall: London.

Coyle RG (2000). Qualitative and quantitative modelling in system dynamics: some research questions. *System Dynamics Review* 16(3): 225-244.

Coyle G and Exelby D (2000). The validation of commercial system dynamics models. *System Dynamics Review* 16(1): 27-41.

Cushman RF, Jacobsen CM and Trimble PJ (1996). *Proving and Pricing Construction Claims*: Second Edition. The Construction Law Library. Aspen Law & Business. Frederick, MD.

Dangerfield B and Roberts C (1996). An Overview of Strategy and Tactics in System Dynamics Optimization. *Journal of the Operational Research Society* 47(3): 405-423.

Dash DP (1994). System Dynamics: Changing Perspectives. *Systems Practice* 7(1): 87-98.

Diehl E and Sterman JD (1995). Effects of Feedback Complexity on Dynamic Decision Making. *Organizational Behaviour and Human Decision Processes* 62(2): 198-215.

Eden C, Jones S and Sims D (1979). *Thinking in Organisations*. MacMillan: New York.

Eden C, Jones S and Sims D (1983). *Messing about in Problems*. Pergamon: Oxford.

Eden C (1989). Operational Research as Negotiation. In: Jackson MC, Keys P and Cropper SA (eds). *Operational Research and the Social Sciences*. Proceedings of IFORS Specialized Conference on Operational Research and the Social Sciences. Plenum Press: New York.

Eden C, Ackermann F and Cropper S (1992). The Analysis of Cause Maps. *Journal of*

Management Studies 29(3): 309-324.

Eden C (1994). Cognitive mapping and problem structuring for system dynamics model building. *System Dynamics Review* 10(2-3): 257-276.

Eden C and Huxham C (1996). Action Research for the Study of Organizations. In Clegg S, Hardy C and Nord W (eds). *Handbook of Organization Studies*. Sage: Beverley Hills.

Eden C and Ackermann F (1998). Analysing and comparing idiographic causal maps. In: Eden C and Spender JC (eds). *Managerial and Organizational Cognition: Theory, Methods and Research*. Sage: London.

Eden C, Williams T, Ackermann F and Howick S (2000). On the nature of disruption and delay (D&D) in major projects. *Journal of the Operational Research Society* 51(3): 291-300.

Espejo R (1987). From Machines to People and Organisations: A Cybernetic Insight on Management. In: Jackson MC and Keys P (eds). *New Directions in Management Science*. Gower: Aldershot.

Fleming JW (1980). A guide to the use of computers to estimate damages in complex litigation. *Computer/Law Journal* 2: 863-875.

Fleming QW and Koppelman JM (2000). *Earned Value Project Management*: 2nd Edition. Project Management Institute, Inc.: Pennsylvania.

Flood RL and Jackson MC (1991). *Creative Problem Solving: A Total Systems Intervention*. Wiley: Chichester.

Flood RL and Carson ER (1992). Dealing with Complexity: *An Introduction to the Theory and Application of Systems Science*. Plenum Press: New York.

Ford DN (1995). *The Dynamics of Project Management: An Investigation of the Impacts of Project Process and Coordination on Performance*. PhD Thesis: Massachusetts Institute of Technology.

Ford DN and Sterman JD (1998). Dynamic modeling of product development processes. *System Dynamics Review* 14(1): 31-68.

Forrester JW (1958). Industrial Dynamics: A Major Breakthrough for Decision Makers. *Harvard Business Review* 36(4): 37-66.

Forrester JW (1961). *Industrial Dynamics*. Productivity Press: Portland, Oregon.

Forrester JW (1968a). *Principles of Systems*. Productivity Press: Portland, Oregon.

Forrester JW (1968b). Industrial Dynamics - A Response to Ansoff and Slevin. *Management Science* 14(9): 601-618.

Forrester JW (1968c). Industrial Dynamics – After the First Decade. *Management Science* 14(7): 398-415.

Forrester JW (1972). Understanding the counterintuitive behaviour of social systems. In: Beishon J and Peters G (eds). *Systems Behaviour.* The Open University: London.

Forrester JW (1975). *Collected Papers of Jay W. Forrester.* Productivity Press: Portland, Oregon.

Forrester JW and Senge PM (1980). Tests for building confidence in system dynamics models. In: Legasto AA Jr., Forrester JW and Lyneis JM (eds). *TIMS Studies in the Management Sciences* 14.

Goodman M, Kreutzer WB, Kreutzer DP and Sterman JD (1993). Electrifying Learning: Computerizing the Beer Game. In: Zepeda E and Machuca JAD (eds). *The Role of Strategic Modeling in International Competitiveness.* The System Dynamics Society.

Graham AK (2000a). Presentation at the Consultants' Roundtable: Using A Simple Framework To Share The Benefits of Different System Dynamics Modeling Approaches. *2000 International System Dynamics Conference.* System Dynamics Society: Bergen, Norway.

Graham AK (2000b). Beyond PM 101: Lessons for Managing Large Development Programs. *Project Management Journal* 31(4): 7-18.

Howick S and Eden C (2001). The impact of disruption and delay when compressing large projects: going for incentives? *Journal of the Operational Research Society* 52(1): 26-34.

Jackson MC and Keys P (1984). Towards a System of Systems Methodologies. *Journal of the Operational Research Society* 35(6): 473 – 486.

Jackson MC (1990). Beyond a System of Systems Methodologies. *Journal of the Operational Research Society* 41(8): 657-668.

Jarmain WE (1963). *Problems in Industrial Dynamics.* The M.I.T. Press : Massachusetts.

Keloharju R and Wolstenholme EF (1989). A Case Study in System Dynamics Optimization. *Journal of the Operational Research Society* 40(3): 221-230.

Keys P (1988). A methodology for methodology choice. *Systems Research* 5(1): 65-76.

Keys P (1990). System Dynamics as a Systems-Based Problem-Solving Methodology. *Systems Practice* 3(5): 479-493.

Kharbanda OP and Pinto JK (1996). *What Made Gertie Gallop?* Van Nostrand Reinhold: New York.

King RA and Brooks PL (1996). Types of Claims. In: Cushman RF, Jacobsen CM and Trimble PJ (eds). *Proving and Pricing Construction Claims*: Second Edition. Wiley: New York.

Lane D (1992). Modelling as Learning. *European Journal of Operational Research* 59(1): 64-84.

Lane D (1994). With a little help from our friends: how system dynamics and soft OR can learn from each other. *System Dynamics Review* 10(2-3): 101-134.

Lane D (1999). Social Theory and System Dynamics Practice. *European Journal of Operational Research* 113(3): 501-527.

Legasto AA Jr. and Maciarello J (1980). System Dynamics: A Critical Review. In: Legasto AA Jr., Forrester JW and Lyneis JM (eds). *TIMS Studies in the Management Sciences* 14.

Lin CY (1993). Walking on Battlefields: Tools for Strategic Software Management. *American Programmer* 6(5): 34-40.

Lin CY, Abdel-Hamid T and Sherif JS (1997). Software-Engineering Process Simulation Model (SEPS). *Journal of Systems Software* 38: 263-277.

Lock D (1994). The Nature and Purpose of Project Management. In: Dinsmore PC (ed). *Handbook of Project Management:* Second Edition. Gower Press: Aldershot.

Lockyer K and Gordon J (1991). *Critical Path Analysis and Other Project Network Techniques*: Fifth Edition. Pitman Publishing: London.

Maslow A (1966). *The Psychology of Science: A Reconnaissance.* Henry Regnery: Chicago.

Meadows DH (1980). The Unavoidable A Priori. In: Randers J (ed). *Elements of the System Dynamics Method.* Productivity Press: Cambridge, Massachusetts.

Mitchell G (1993). *The Practice of Operational Research.* Wiley: Chichester.

Morecroft J D W (1982). A Critical Review of Diagramming Tools for Conceptualizing Feedback System Models. *Dynamica* 8(1): 20-29.

Morgan G (1997). *Images of Organizations*. Sage: California.

Morris PW and Hough GH (1987). *The Anatomy of Major Projects*. Wiley: Chichester.

Nahmias S (1980). The Use of Management Science to Support a Multimillion Dollar Precedant-Setting Government Contract Litigation. *Interfaces* 10(1): 1-11.

Paich M and Sterman J (1993). Boom, bust, and failures to learn in experimental markets. *Management Science* 39(12): 1439-1458.

Pidd M (1996). *Tools for Thinking: Modelling in Management Science.* Wiley: Chichester.

Pritchard RD (1990). *Measuring and Improving Organizational Productivity: A Practical Guide*. Praeger: New York.

Project Management Institute (1996). *A Guide to the Project Management Body of Knowledge*. Project Management Institute: North Carolina.

Raby M (2000). Project management via earned value. *Work Study* 49(1): 6-9.

Randers J (ed.) (1980) *Elements of the System Dynamics Method*. Productivity Press: Cambridge.

Richardson G and Pugh A III (1981). *Introduction to System Dynamics Modeling*. Productivity Press: Portland, Oregon.

Richardson G (1991). *Feedback Thought in Social Science and Systems Theory*. Pegasus Communications, Inc.: Waltham, Massachusetts.

Richardson GP (1996). Problems for the future of system dynamics. *System Dynamics Review* 12(2): 141-157.

Richardson GP (1999). Reflections for the future of system dynamics. *Journal of the Operational Research Society* 50(4): 440-449.

Rivett (1972). *Principles of Model Building*. Wiley: London.

Roberts EB (1964). *The Dynamics of Research and Development*. Harper & Row: New York.

Roberts EB (ed.) (1978). *Managerial Applications of System Dynamics*. Productivity Press: Cambridge, Massachusetts.

Roberts N, Andersen D, Deal R, Garet M and Shaffer W (1983). *Introduction to Computer Simulation: A System Dynamics Modeling Approach*. Productivity Press: Portland, Oregon.

Rodrigues AG and Bowers J (1996). System dynamics in project management: a comparative analysis with the traditional methods. *System Dynamics Review* 12(2): 121-139.

Rodrigues AG (2000). *The Application of System Dynamics to Project Management: An Integrated Methodology (SYDPIM)*. PhD Thesis: University of Strathclyde, UK.

Schoderbek CG, Schoderbek PP and Kefalas AG (1980). *Management Systems: Conceptual Considerations*. Business Publications: Dallas, Texas.

Scott S (1993). Dealing with delay claims: a survey. *International Journal of Project Management* 11(3): 143-153.

Senge P M (1990). *The Fifth Discipline*. Doubleday: New York.

Sterman JD (1984). Appropriate Summary Statistics for Evaluating the Historical Fit of System Dynamics Models. *Dynamica* 10(2): 51-66.

Sterman JD (1989a). Modeling Managerial Behavior: Misperceptions of Feedback in a Dynamic Decision Making Experiment. *Management Science* 35(3): 321-339.

Sterman JD (1989b). Misperceptions of Feedback in Dynamic Decision Making. *Organizational Behaviour and Human Decision Process* 43(3): 301-335.

Sterman JD (1994). Learning in and about complex systems. *System Dynamics Review* 10(2-3): 291-330.

Sterman JD (2000). *Business Dynamics: Systems Thinking and Modeling for a Complex World.* Irwin/McGraw-Hill: Chicago.

Turner JR (1993). *The Handbook of Project-Based Management.* McGraw-Hill: London.

Veit KP (1978). System Dynamics in Corporate Long-Range Strategic Planning. In: Roberts EB (ed). *Managerial Applications of System Dynamics.* Productivity Press: Cambridge, Massachusetts.

Vemuri V (1978). *Modeling of Complex Systems.* Academic Press: New York.

Vennix J (1996). *Group Model Building: Facilitating Team Learning Using System Dynamics.* Wiley: Chichester.

Ventana Systems (1998). *Vensim 3.0 Manual.* Ventana Systems, Inc.: Harvard, Massachusetts.

Warren K and Langley P (1999). The effective communication of system dynamics to improve insight and learning in management education. *Journal of the Operational Research Society* 50(4): 396-404.

Webster FM (1993). What Project Management Is All About. In: Dinsmore PC (ed). *Handbook of Project Management:* Second Edition. Gower Press: Aldershot.

Weil HB and Etherton RL (1990). System Dynamics in Dispute Resolution. In: Andersen D, Richardson G and Sterman J (eds). *Proceedings of the 1990 International System Dynamics Conference.* System Dynamics Society: Chestnut Hill, Massachusetts.

Wickwire JM and Smith RF (1974). The use of critical path method techniques in contract claims. *Public Contract Law Journal* 7(1): 1-45.

Williams T, Eden C, Ackermann F and Tait A (1995). The effects of design changes and delays on project costs. *Journal of the Operational Research Society* 46(7): 809-818.

Williams T (1997). The Need for New Paradigms for Complex Projects. In : Williams T (ed). *Managing and Modelling Complex Projects.* Kluwere Academic Publishers: Dordretch.

Williams T, Ackermann F and Eden C (2000). Structuring a Delay and Disruption Claim. Working Paper 2000/1: *Theory, Method and Practice Series.* Department of Management Science, University of Strathclyde, UK.

Wolstenholme EF and Coyle RG (1980). Modelling Discrete Events in System Dynamics Models: A Case Study. *Dynamica* 6(1): 21-27.

Wolstenholme EF (1982). System Dynamics in Perspective. *Journal of the Operational Research Society* 33(6): 547-556.

Wolstenholme EF and Coyle RG (1983). The Development of System Dynamics as a Methodology for System Description and Qualitative Analysis. *Journal of the Operational Research Society* 34(7): 569-581.

Wolstenholme EF (1990). *System Enquiry: A System Dynamics Approach.* Wiley: Chichester.

Wolstenholme EF (1999). Qualitative *vs* quantitative modelling: the evolving balance. *Journal of the Operational Research Society* 50(4): 422-428.

Yelle LE (1979). The Learning Curve: Historical Review and Comprehensive Survey. *Decision Sciences* 10: 302-328.