THE UNIVERSITY OF WARWICK

A Thesis Submitted for the Degree of PhD at the University of Warwick

Development and use of simulation models in Operational Research: a comparison of discrete-event simulation and system dynamics

by

Antuela Anthi Tako

A thesis submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy

University of Warwick, Warwick Business School

November 2008

Table of contents

Table of contents

Figures

Tables

Acknowledgements

I am indebted to many people for the help, advice and support provided during the process of completing this PhD thesis. I would like to dedicate this work to my mother, who has always given me the inspiration and the craving for learning. She has been and will always be my motivator in pursuit of knowledge.

Firstly, I would like to thank my supervisor, Professor Stewart Robinson for the continuous encouragement, support and guidance. I am extremely grateful for his direction and timely advice that has enlightened me throughout the process of compiling this work. I also appreciate the comments and the help provided by Professor Ruth Davies, my second supervisor. Thanks must also go to both my examiners, John Morecroft and Frances O'Brein, for having patiently read the thesis and for giving me constructive feedback.

I am extremely grateful to various sources that provided me with financial support and made possible the completion of the thesis. I would like to thank the Warwick Business School and the University of Warwick, for funding me during the course of the doctoral studies. I am also grateful for the financial support provided by Christina Lannara, who has shown a keen interest in my progress and has generously helped me in difficult times.

Thanks must go to the participants of the two studies involved in this work. Even though their names are kept anonymous, I feel indebted to the expert simulation modellers who took the time to participate in the model building study. I now enjoy a good rapport with most of the participants, usually meeting with them in various conferences. I am also grateful to Dr. Susan Howick for her help in contacting potential participants. I would also like to thank the 2 groups of Executive MBA students for participating and completing the questionnaire survey involved in the model use study. Thanks go to the administrative staff at the Executive MBA programme at Warwick Business School, Karen Finlay and Irene Ioannou for their help and support with the preparations of materials for the model use sessions.

I would also like to thank Lanner Group for their help with using Witness.

I am also grateful for the various sources of funding which have enabled me to attend conferences which allowed me to gain a valuable insight into OR. These include: Operational Research and Management Sciences Group at Warwick, Warwick Business School, the UK OR Society, the International System Dynamics Society and the UK chapter of the System Dynamics Society.

I am indebted to many of my colleagues and friends, Natalia Yannopoulou, Martin Liu, Ernie Lee, Suchi Patel, Katy Hoad, Eudokia Koutsoumpa, Michael Nottage, Chris Watts and the staff members of the Operational Research and Management Science group, who have helped to make the time more enjoyable. Thanks go to the secretaries of the Operational Research and Management Sciences group, Sue Shaw and Racheal Monnington for the animated chats and their support.

Finally, special thanks go to my parents and my sisters for their support and encouragement throughout the PhD and in particular to my partner, Fareed for his help in proofreading the chapters of the thesis and for his continuing care and love.

Declaration of authorship

I, Antuela Anthi Tako, declare that this thesis entitled:

"**Development and use of simulation models in Operational Research: a comparison of discrete-event simulation and system dynamics"** and the work presented are my own. I confirm that:

- This work was done wholly while in candidature for this research degree;
- This thesis contains no material which has been accepted for the award of any other degree or diploma in any university;
- I have acknowledged all main sources of help;
- The work described in this thesis has served the material for several published papers and conference presentations which are listed below.

Journal paper:

Tako A.A. & Robinson S. (2009) "Comparing discrete-event simulation and system dynamics: Users' perceptions" **Journal of the Operational Research Society**, 60, (296-312), doi:10.1057/palgrave.jors.2602566.

Refereed conference proceedings:

"Model building in System Dynamics and Discrete-event Simulation: a quantitative comparison", In the proceedings of the **26th International Conference of the System Dynamics Society**, 20-24 July 2008, Athens, Greece. (The paper won the **Dana Meadows Award** for the best student paper submitted at the conference)

"Comparison of Model Building in Discrete-Event Simulation and System Dynamics". In the proceedings of the **2008 OR Society Simulation Workshop (SW08)**, 1-2 April 2008, pp. 209-218, Worcestershire, UK.

"Towards an empirical comparison of DES and SD in the supply chain context". In the proceedings of the **2006 OR Society Simulation Workshop (SW06)**, 28-29 March 2006, pp. 149-156, Leamington Spa, UK.

Conferences and presentations:

Presentation: "Comparing the Use of Discrete-Event Simulation and System Dynamics Models", **PhD Colloquium in the Joint meeting of UK Chapter of the System Dynamics Society & OR Society SD+ Study Group**, 7-8 February 2008, London South Bank University, London.

Workshop presentation: "Comparing SD & DES using prison population models", in the **Joint meeting of UK Chapter of the System Dynamics Society & OR Society SD+ Study Group**, 7-8 February 2008, London South Bank University, London.

Poster presentation: "Comparing the Use of Discrete-Event Simulation and System Dynamics Models", **Winter Simulation Conference 2007**, Washington D.C., 8-10 December 2007.

Presentation: "Discrete-Event Simulation & System Dynamics Models: Comparison from Users Point of View", **OR49** conference, OR Society, Edinburgh, 4-6 September 2007.

Presentation "Comparison of Discrete-Event Simulation and System Dynamics" at **Young OR15 conference** (YOR15), OR Society, Bath, 28-30 March 2007.

Signed:

Date:

Abstract

The thesis presents a comparison study of the two most established simulation approaches in Operational Research, Discrete-Event Simulation (DES) and System Dynamics (SD). The aim of the research implemented is to provide an empirical view of the differences and similarities between DES and SD, in terms of model building and model use. More specifically, the main objectives of this work are:

1. To determine how different the modelling process followed by DES and SD modellers is.

2. To establish the differences and similarities in the modelling approach taken by DES and SD modellers in each stage of simulation modelling.

3. To assess how different DES and SD models of an equivalent problem are from the users' point of view.

In line with the 3 research objectives, two separate studies are implemented: a model building study based on the first and second research objectives and a model use study, dealing with the third research objective. In the former study, Verbal Protocol Analysis is used, where expert DES and SD modellers are asked to 'think aloud' while developing simulation models. In the model use study a questionnaire survey with managers (executive MBA students) is implemented, where participants are requested to provide opinions about two equivalent DES and SD models.

The model building study suggests that DES and SD modelling are different regarding the model building process and the stages followed. Considering the approach taken to modelling, some similarities are found in DES and SD modellers' approach to problem structuring, data inputs, validation & verification. Meanwhile, the modellers' approach to conceptual modelling, model coding, data inputs and model results is considered different. The model use study does not identify many significant differences in the users' opinions regarding the specific DES and SD models used, implying that from the user's point of view the type of simulation approach used makes little difference if any.

The work described in this thesis is the first of its kind. It provides an understanding of the DES and SD simulation approaches in terms of the differences and similarities involved. The key contribution of this study is that it provides empirical evidence on the differences and similarities between DES and SD from the model building and model use point of view. Albeit the study does not provide a comprehensive comparison of the two simulation approaches, the findings of the study, provide new insights about the comparison of the two simulation approaches and contribute to the limited existing comparison literature.

Chapter 1: Introduction

1.1 Introduction

This chapter introduces the topic of the thesis, setting the background of the work undertaken in the chapters to follow. For the readers who are not knowledgeable about simulation modelling, the discrete-event simulation and the system dynamics approach are initially described. A brief description of the thesis follows, which provides the reasoning behind and the initial thoughts that stimulated this study. This chapter ends with an overview of the chapters of the thesis.

1.2 What is simulation?

Talking about simulation as a tool in management science, Pidd (2003) explains that computer simulation is the use of a model to understand and experiment with a system. He also adds that a simulation model mimics the changes that occur through time in the real system. A more complete definition of simulation modelling is given by Robinson (2004), according to which simulation is the "*experimentation with a simplified imitation (on a computer) of a [...] system as it progresses through time, for the purpose of better understanding and/or improving that system*". A system consists of a collection of parts that interact with each other to achieve a specific objective. The notion of simulation modelling in discrete-event simulation (DES) and system dynamics (SD) is fairly represented by the aforementioned definitions. A similar definition of SD modelling is given by Wolstenholme (1990), who defines SD as the study of the behaviour of physical and social systems over time through the creation of diagrammatic representations and computer simulations with the objective to understand system behaviour and to facilitate the design of the improved system. Therefore, when mentioning simulation throughout this thesis, both DES and SD are implied.

Simulation models have been classified as static or dynamic, deterministic or stochastic, and discrete or continuous (Banks, et al., 2001). A static simulation model represents a system at a particular point in time, usually called Monte-Carlo simulation. Dynamic simulation models represent systems that change as time advances. A simulation model of an inventory system for a one-week time period is an example of a dynamic simulation. In addition, simulation models are classified as deterministic, when randomness is not included and they, therefore, have a known set of inputs and a unique set of outputs. Such is the example of an SD model, which at most uses aggregate values (averages) as input variables and therefore, the outputs are consequently deterministic. A stochastic simulation model has one or more random variables as inputs, which lead to random outputs. A simple example of a stochastic model is that of a queuing system, where customers arrive following a random distribution of arrivals and service times vary according to a statistical distribution and as a result waiting times vary.

The next distinction of simulation models, that of discrete and continuous systems is based on how time is handled in the system (Banks et al., 2001; Law, 2007). In discrete systems, state variables change in distinct time-steps (Δt) (i.e. the number of customers in a bank), while in a continuous system state variables change continuously (i.e. the amount of water flowing through a pipe). However, only few systems are completely discrete or completely continuous. For most systems one type of change predominates and it is therefore, possible to classify systems as either discrete or continuous based on the type of change that predominates (Banks et al., 2001; Pidd, 2004; Law, 2007). Hence, the decision as to the most suitable simulation approach to model a specific system varies.

Discrete-Event Simulation (DES) and System Dynamics (SD) are two popular simulation approaches used in Operational Research (Pidd, 2004). They are commonly used in business settings to support management learning and decision-making (Robinson, 2004). With these in mind, the two simulation approaches studied in this thesis, DES and SD are introduced in the next two sections.

1.3 Discrete-event simulation (DES)

A brief introduction to DES modelling is provided, starting from its definition, the key concepts involved and the evolution of the field.

1.3.1 What is DES?

DES involves modelling of systems which consist of discrete entities going through specific states, which change at discrete points in time (Pidd, 2003). These points in time are the ones at which state changes (or events) occur. For example, if customers arrive in a bank system at time 0, 3, 5, 8 and 11 then the model will progress by the corresponding time steps (Δt), 3, 2, 3 and 3.

The variable time is tracked by the *simulation clock*, which gives the current simulation time. The simulation clock advances the variable time based on the time advance approach used. Two different time advance approaches are used, the next-event or the fixed increment time advance approach. In most DES models the next-event time advance approach is used (Law, 2007) and therefore, explained here. At the beginning of the simulation, the simulation clock is initialised at time zero, it then advances to the time when the most imminent future event is planned to occur. In the bank example (Figure 1-1), the first event is the arrival of the first customer at time 0 (who will go straight to the till to be served because no one is waiting in the queue). If service time at the till is more than 3 minutes, then the most imminent event becomes the arrival of the second customer, who will arrive at time unit 3. Therefore, the simulation clock will jump from time 0 to time 3, which is the time when the next event is scheduled to happen. The occurrence of this event is updated in the knowledge of the simulation model and the clock is then advanced to the new most imminent event (if we consider that the till service duration is 4 minutes, then at time 4, which is after a $\Delta t = 1$ time unit, the next event occurs, that of the customer leaving the till) and so on. This illustrates how the simulation clock advances at unequal time slots, whenever an event occurs. Delay times in the form of periods of inactivity are also involved, which are skipped as the simulation clock jumps from one event time to another. An example of delay times is that of customers waiting in the queue to be served.

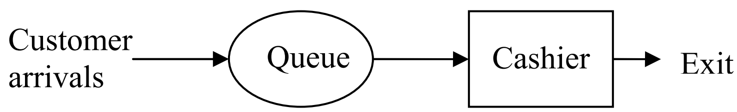


Figure 1-1: Process flow diagram of a simple bank system, where customers arrive, wait in the queue, are served at the till and then leave the system (adapted from Robinson, 2004).

The explanations provided in this section consist only of the basic principles involved in the logic of a DES model. For a more in-depth description of the internal simulation logic the reader is referred to simulation textbooks (Banks et al., 2001; Pidd, 2004; Robinson, 2004; Law, 2007).

1.3.2 Main DES concepts

Entities are objects or components whose behaviour is tracked through the model as the simulation proceeds. Some examples of entities are: customers, patients and products. Entities go through a number of states, which represent their progression in the system. For example, in a bank customers go through a series of states: enter the system, wait to be served, get served and leave the system (Figure 1-1). An entity remains in a state for a period of time and whenever its state changes, an *event* occurs. Events can be endogenous, i.e. to describe events that occur within the system (i.e. completion of a service) or exogenous to describe events in the environment that affect the system (i.e. arrival of customers). Entities are given *attributes,* which can be thought of as properties of each specific entity. Attributes are attached to the entities and these control various aspects such as: routing of objects, priority given to a customer waiting to be served, etc.

Activities are another important element in DES models. An activity takes a specific amount of time, which involves the time needed for an entity to change from one state to another. Typical activities include service times, inter-arrival times, or any other processing times defined in the simulation. In the bank example, an activity occurs throughout the time that the customer is in the ‘being served by the cashier’ state. An activity begins with an event and ends with an event. In the bank example, an activity occurs when a customer reaches the cashier (event 1) until they leave the system (event 2).

In a DES model, before entering an activity, entities wait in *queues (or buffers)*, until the activity centre becomes available. In the bank example (Figure 1-1) the customer waits in the queue for the cashier to finish dealing with the previous customer and to become available. The length of time that an entity spends in a queue is not generally specified. It depends on the end time of the immediately preceding activity, the number of entities in the queue and the start time of the following activity. However, conditional queues (or delays[1]) can also be specified, in which the progress of the entity through the process pauses until some simulated time interval has elapsed (Pidd, 2004).

A simulation control program maintains an event calendar which holds information regarding the state of all the entities in the system. This control program has a list of future events, ordered by time of occurrence. Every time an event occurs the simulation control programme is updated with the current state of entities and also the time left for the entities to remain in a specific state.

From the point of view of DES modelling, most systems contain one or more sources of randomness. Therefore, a central aspect of DES simulation is the inclusion of randomness or variability, from which DES gains its stochastic nature. Randomness is achieved with the use of random variables as inputs and sampling using random numbers. Random numbers consist of a sequence of numbers, integer or real, which appear in a random order. They are generated by a specific method[2] and then used to determine the value of simulation variables from the probability function of empirical or statistical distributions. Two types of distributions are usually used discrete (binomial, Poisson, Bernoulli, etc.) or continuous (i.e. exponential, gamma, uniform, etc.) distributions. For example, in a DES model of a bank typical random variables would be: customer arrivals and service times,

[1] Queues are equivalent to delays in the system.
[2] Most simulation languages have a method that generates random numbers. Various techniques for generating random numbers are described in a number of discrete-event simulation textbooks (Banks, et. al, 2001; Law, 2007).

represented by exponential and gamma distribution respectively. In stochastic models, outputs are presented as statistical estimates of the true system outputs, i.e. average number of customers waiting or average of customer waiting time, etc.

1.3.3 Evolution and applications of DES

The beginnings and evolution of the field of DES is related to the availability and developments in computing (Pidd, 2003; Pidd, 2004; Robinson, 2005). In the late 1950s and in the 1960s the first simulation models were developed in the form of computer code (Robinson, 2005). Later on, with the introduction of programming languages and more powerful computers, specialist simulation software were developed in the 1960s i.e. GPSS, SIMSCRIPT and SIMULA. At the same time, advances in simulation methodology were made with Tocher's (1963) introduction to the three-phase approach in DES, which is still used by various DES packages.

In the 1970s simulation software continued to improve and new languages were introduced. The beginning of visual interactive simulation (VIS) was pioneered with the work of Hurrion (1981; 1991). The use of this technique allows the simulation model to be visually displayed on the screen, enabling the user to watch the model progress through time. The introduction of VIS signalled the time when decision makers and users were becoming more involved with simulation models (Hurrion and Secker, 1978). It started first by giving decision makers the chance to interact with the model and later on getting involved in the model building process. In this day and age, VIS is a common feature of most DES software. However, this was a breakthrough step at that time. This advancement was of course enabled with the rapid progress of computing and the spread of microcomputers. The first commercial VIS software, SEE-WHY was developed in 1979, followed by other packages such as WITNESS, GENETIK, ProModel, etc.

In the 1990s further development of VIS continued, which consisted mainly of improvements in the software interface. Besides expensive simulation packages (WITNESS, Arena, Taylor II, AweSim, etc.), low cost packages were developed, Simul8, Extend, etc., which made simulation modelling more accessible to the wider user base. Some further developments in DES are: simulation optimisation, virtual reality, software integration and parallel distributed simulation (Robinson, 2005). Simulation optimisation became a significant feature in simulation, where the most commonly used optimisation approaches are: metamodelling, neural networks and metaheuristics. A number of simulation packages now incorporate some form of optimisation facility.

Distributed simulation is another important development that has emerged over the last 15 years, linked particularly to the idea of running simulation models on the Wide World Web (WWW). More specifically distributed simulation enables the split of a large model across many computers or the link of separate models on different computers, which can be run concurrently (Cassel and Pidd, 2001). With the use of distributed simulation, useful advantages can be reached, i.e. distributed users can interact with a gaming simulation model, models and software can be shared among distributed users, running model replications or scenarios across a number of distributed computers, etc. (Robinson, 2005).

DES has been traditionally used in the manufacturing sector, however, from the 1990s onwards it has been increasingly used in the service sector (Robinson, 2005). Some key DES applications include airports, call centres, fast food restaurants, banks, health care and business processes. Even though the developments of computing have to a great extent affected developments in DES modelling, Robinson raises awareness about the problems it can incur, including the development of large and complex problems, with a knock on effect on the quality of the models. A lack of a wider methodology of DES modelling has also been pointed out (Robinson, 2005).

1.4 System dynamics (SD)

In the following paragraphs, a brief overview of the field of SD is provided, starting with the key concepts in SD modelling, followed by the applications and the evolution of the field.

1.4.1 What is SD?

SD is a simulation approach with roots from engineering, cybernetics and organisational theory (Meadows, 1980), developed by J. Wright Forrester at MIT. It studies social or managed systems, with the main aim to understand system behaviour and consequently to experiment with policies which will improve their behaviour. As its name implies, system dynamics studies dynamic behaviour of complex systems over time. The main interest of system dynamicists is the general dynamic tendency of the system as a whole, whether it is stable or unstable, oscillating, growing or declining (Meadows, 1980). Hence, SD is considered to be taking a holistic, systems' approach (Wolstenholme, 1990).

A central concept in the SD paradigm is that of feedback, a closed cause and effect chain, where information about the result of an action is fed back to generate further action. This is clearly illustrated in Coyle (1996) with the information/ action/ consequences loop (Figure 1-2). According to this loop, the state of the system i.e. the size of the workforce comes from the choices made by its managers, such as recruiting more people or making them redundant. As a next step the 'state' is acknowledged by the controllers in the system. This knowledge (learning), based on the discrepancy between the actual perceived level of the state and the desired workforce level, leads to a choice of new actions to decrease or increase the workforce to the desired level.

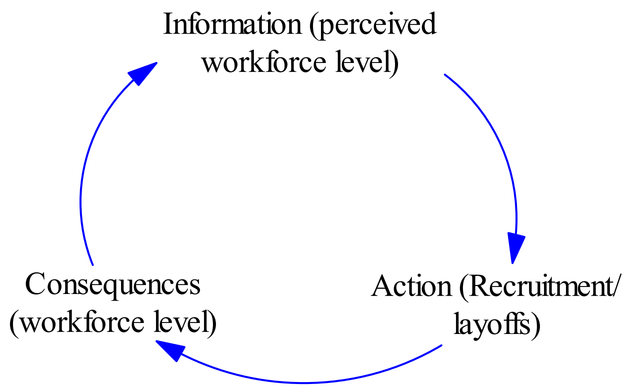


Figure 1-2: Information/action/consequences loop, based on (Coyle, 1996).

Two types of feedback loops are used in SD, positive or reinforcing loops and negative or balancing loops (Wolstenholme, 1990; Coyle, 1996; Morecroft, 2007). Positive loops tend to amplify effects in the system and result in exponential growth. For example, the state of a system, i.e. existing population, grows continually as birth rates increase, which in turn increase as the population increases. Negative feedback loops tend to counteract, seeking to achieve a desired state or equilibrium of the system. This is driven by actions (policies) taken in an attempt to eliminate the difference between the desired and the actual state of the system. In the case of the introduction of a new product in the market, the initial price is high, which incurs high profits, increasing the attractiveness of the market and therefore, attracting new competitors. With the increase of the number of competitors in the market, the supply increases and consequently the product price falls, causing a reduction in company profits, reducing attractiveness of the market and hence, competitors leave the market. Thus the number of competitors reduces, up to a point that the price reaches an equilibrium state.

Nonlinear relationships are an important feature in SD, which can affect the strength of feedback loops depending on the state of the system (Meadows, 1980; Sterman, 2000). In more complex models, the existence of several nonlinear relationships can result in a complex dynamic behaviour, which is more difficult to conceptualise and humans lack the cognitive ability to deduce the resulting dynamic behaviour (Sterman, 2000). Therefore, computer simulation is considered necessary in order to deduce the system behaviour (Lane, 2000).

1.4.2 SD simulation concepts

Defining the structure of the system is crucial in SD because explanations about the system behaviour are sought into the internal system structure rather than in external factors and random events. In the SD world, systems are made up of flows (rates) and stocks (levels). *Stocks* are defined as an accumulation of resources through time in the form of material (people, cash, orders, etc) or information (perceptions, knowledge, motivation, etc.) resources. *Rates* or *flows* consist of the decision, action or change that affects the material or information resources flowing between levels. They directly control the increase or decrease of resource levels. The difference between material and information flows is related to the application of conservation laws. In the case of material flows, resources cannot be lost or gained within the model, whereas information flows are not conserved, for example resources such as motivation are not limited in quantity.

In SD modelling two stages can be distinguished: qualitative and quantitative modelling (Wolstenholme, 1990; Coyle, 1996). Qualitative modelling involves building diagrams or conceptual models as a means of transmitting mental models, without necessarily using a computer. A number of tools are available such as: model boundary diagrams, sub-system diagrams, causal loop diagrams and stock and flow maps (Sterman, 2000). The latter 2 are the diagramming tools most often used in SD modelling.

Causal loop diagrams represent the feedback structure of the system using causal relationships between variables, starting from the cause and ending in the effect variable. Based on the type of effect, a positive or negative relationship between two variables can be determined. Stock and flow diagrams are a continuation of causal loop diagrams, where the physical processes occurring in the system are specified. Stock and flow diagrams represent the accumulation of resources into stocks, controlled by the flow rates (in-flows and out-flows) in the system, which may depend on information from auxiliary variables and other levels.

Quantitative modelling involves the use of a computer, where the relationships suggested in the conceptual modelling stage are mathematically quantified. It uses differential equations, which in mathematical terms represent the integration of rates over a period of time. SD is categorised as a continuous approach because the changes in the accumulated level are represented as continuous flows of material. The accumulated level served by a rate variable is equal to the area under the graph of the rate plotted over that period of time (Coyle, 1996). The progress of time in SD is based on advancement of the clock in equal small time-steps, called Δt, known as the 'simulation' interval. When building SD equation models, three points of time are of interest, the current point in time, the past point in time and the future point in time, all separated by the simulation interval (Δt).

1.4.3 Evolution and applications of system dynamics

Like DES, the field of SD came into being and evolved with the advent of computers. The beginnings of the field are marked with work done by Forrester in the 50s, who applied systems control theory from engineering in the study of management and business problems. Forrester's first published work *Industrial Dynamics* addressed industrial issues, such as corporate planning and policy design, including models of inventory-distribution systems, involving inventory controls, the delays in the system, as well as production capacity and utilisation of workforce. His most famous contribution was the analysis of demand amplification, later on called the Forrester effect. According to this phenomenon, a small change in demand at the customer level can create fluctuations in the stock levels upstream in the supply chain (Forrester, 1958). The greater the number of echelons involved, the more uncertainty is transmitted over the next stage due to delay, noise and bias

(Hieber and Hartel, 2003). Forrester applied the SD approach in the social aspect and the problems involved in his subsequent studies of *Urban Dynamics* and *World Dynamics*. The main aspects involved are: population density, availability of jobs, migration and availability of housing.

The late 70s see a reduction in the number of SD applications, covering a wider range of academic disciplines, looking mainly at socio-economic problems (Wolstenholme, 1990). At the same time, there has been a growing trend in moving away from quantified simulation models towards the qualitative (diagramming) approach. In the light of this, SD has been widely used to study ill-defined problems, forming epistemological and ontological assumptions that support the view that SD share common characteristics with soft systems methodology (Wolstenholme, 1990; Lane and Oliva, 1998). A significant advancement in the 80s was the emergence of user-friendly simulation software with advanced graphical user interfaces (STELLA, Powersim and Vensim) and the use of interactive simulation games (Forrester, 2007). From then onwards, applications of SD modelling have significantly expanded.

SD has been applied to a wide range of problems. Work involves applications in economic behaviour, politics, psychology, defence and criminal justice, energy and environmental problems, supply chain management, biological and healthcare modelling, project management, educational problems, staff recruitment, and also manufacturing.

1.5 What the thesis is about?

Practitioners and academics often get into discussions about which simulation approach is most suitable for modelling which problem. The choice between the two simulation approaches in practice is an on-going problem, however, in order to reach informed answers, a common understanding of the two simulation schools is needed. The current work contributes towards this aspect, looking into the differences and similarities between DES and SD. Hence, the main research question that drives the work undertaken in the thesis is: “What are the differences and similarities between the DES and SD modelling approaches and the use of DES and SD models?”

In this thesis it is accepted that there are fundamental differences between DES and SD, which are derived from the technical concepts behind each simulation approach. It can be acknowledged that the two simulation techniques take different viewpoints to modelling and problem-solving. Whilst one might suppose that this makes them natural antagonists it can be argued that they complement each other (Morecroft and Robinson, 2005). Due to the limited literature and the lack of objectivity in the comparison criteria expressed in the literature so far, the scope of the thesis is to provide an empirical comparison of the two simulation approaches. More specifically, the current work looks at the differences and similarities in DES and SD simulation modelling as observed during the model building process as well as during the use of equivalent simulation models. The three main objectives of the thesis are:

1. To determine how different the modelling process followed by DES and SD modellers is.

2. To establish the differences and similarities in the modelling approach taken by DES and SD modellers in each stage of simulation modelling.

3. To assess how different DES and SD models of an equivalent problem are from the users' point of view.

The work carried out in the thesis is considered as a contribution to knowledge for a number of reasons. Literature on the comparison of the two simulation approaches Discrete-Event Simulation (DES) and System Dynamics (SD) is limited. Very few papers, mostly conference papers, have been published on this topic and some limited references are made in simulation textbooks. Additionally, a common characteristic of most studies is that statements made are mostly opinion-based, derived from the authors' personal opinions and field of expertise. As a result, comparisons tend to be biased towards the DES or SD approach (Brailsford and Hilton, 2001; Morecroft and Robinson, 2005). Therefore, undertaking an empirical study is innovative to the comparison literature and provides a valuable evidence base for making comparisons.

Traditionally, since the beginnings of each field, there has been very little dialogue between the two modelling communities (Sweetser, 1999; Lane, 2000; Borshchev and Filippov, 2004; Morecroft and Robinson, 2005). Most recently some efforts are being made to create a bridge between the two simulation approaches. This is evidenced by the fact that a number of simulation experts from one area are attempting to enter into the other world. Furthermore the number of research projects[3] currently under way reveals the increased interest shown in the comparison of the two simulation approaches. This thesis also contributes towards this aspect, by representing the two worldviews and looking into how they can benefit from each other. Furthermore, DES and SD models of the same problem are developed.

1.6 Thesis outline

The rest of this thesis is organised in the following chapters.

[3] Herbert Daly, 'Investigating Multi Method Modelling by Critical Comparison', Brunel University; Jennifer Morgan, 'Linking Discrete Event Simulation and System Dynamics in Healthcare', University of Strathclyde (EPSRC case award);
EPSRC grant: 'Multi-level simulation models (combining system dynamics and discrete-event simulation) for integrated strategic/operational modelling of health systems', University of Southampton (EP/C531930/1);
Christopher C. Owen, 'Use of simulation to improve the performance of supply chains.' Aston University.
Mazlina Abdul Majid, 'Human Behaviour Modelling for Discrete Event and Agent Based Simulation: A Case Study', University of Nottingham;

Chapter 2 introduces the existing literature on the comparison of the two simulation paradigms, discrete-event simulation (DES) and system dynamics (SD). The chapter discusses the key aspects considered in the comparison literature from the model building, modelling philosophy and model use point of view. A list of the key statements found in the literature is provided and conclusions are drawn with the view to specifying the objectives of the research undertaken.

Chapter 3 sets the scene for the empirical work undertaken in the thesis. The research question is discussed followed by the 3 key research objectives. The research hypotheses, based on the existing comparison literature are stated for the two different aspects of the comparison work undertaken, model building and model use. Furthermore, the research methodology is considered. The methods chosen for the model building and model use study, verbal protocol analysis (VPA) and survey questionnaire respectively, are explained.

In *chapter 4*, a case study on the UK prison population is presented. This case study serves as the research stimulus for both parts of the work undertaken and the reasons for and the suitability of using it are discussed.

Chapter 5 describes the design of the simulation model building study. The steps undertaken are explained. A description of the experimental modelling sessions is provided, including the selection of participating modellers and the data analysis process. An account of the pilot study implemented and the learning achieved is also provided.

The results of the quantitative analysis of the 10 verbal protocols derived from the model building study are presented in *chapter 6*. The aim of this chapter is to explore the differences and similarities in the DES and SD modelling process regarding the distribution of modellers' attention to modelling topics and the sequence of attention among topics.

The results of the qualitative analysis of the 10 verbal protocols are provided in *chapter 7*. The analysis is based on the testable research hypotheses stated in chapter 3. Thus DES and SD modellers' thoughts relevant to the topic of each hypothesis are first presented followed by a comparison of the views expressed by the two groups of modellers. The findings from the quantitative and qualitative analysis (chapters 6 and 7 respectively), are further discussed in *chapter 8*.

Chapter 9 describes the study of model use based on the users' perceptions of two equivalent DES and SD models. More specifically this chapter includes a description of two equivalent simulation models created (one in DES and one in SD), the participant sample, the questionnaire survey as well as the pilot study undertaken. The results of the questionnaire survey are provided in *Chapter 10*. The results are analysed using non-parametric statistical tests and the findings discussed.

Chapter 11 concludes the thesis. It provides a summary of the findings and the extent the research objectives were achieved. Then the contribution of the research to the comparison literature is discussed, as well as its limitations and the potential for further research.

Chapter 2: A comparison of DES and SD in the literature

2.1 Introduction

In this chapter work on the comparison of two simulation approaches DES and SD is reviewed. The aim of this chapter is to provide an overview of the main issues discussed in the literature. The statements found in the comparison literature are considered with a view to specifying the aspects of DES and SD modelling to be investigated in the thesis.

The chapter starts with some general comments on the existing comparison literature, followed with a brief consideration of the technical differences between the two approaches. Then a list of the topics considered in the comparison literature is provided based on the aspects involved in the OR modelling process. Work on the choice of modelling approaches is also presented. The chapter ends with an overall summary of the topics considered in the comparison literature.

2.2 Overview of the comparison literature

While it has been reported that traditionally there has been little dialogue between the two modelling communities (Sweetser, 1999; Lane, 2000; Borshchev and Filippov, 2004; Morecroft and Robinson, 2005), the need for mutual communication has been pointed out (Lane, 2000). In recent years, there has been some change with more academics and practitioners showing an interest in future collaboration between the two fields (Morecroft and Robinson, 2005). Therefore, the need for a comparative study is becoming more necessary.

Literature on the comparison of simulation approaches is scarce, consisting mainly of conference papers, which are not always publicly accessible. Work on the comparison of the two simulation techniques consists mostly of generally accepted statements (Brailsford and Hilton, 2001; Morecroft and Robinson, 2005). It mainly consists of authors' personal opinions, which tend to be biased towards either the DES or SD approach, coming from their own area of expertise (Brailsford and Hilton, 2001). Furthermore, limited empirical work has been done to provide evidence certifying the existing statements in the literature. The only empirical study carried out to date is that of Morecroft and Robinson (2005). The authors built a step-by-step simulation model of a fishery, using SD (Morecroft) and DES (Robinson) modelling, comparing model representation and interpretation. However, one could claim the existence of bias, as the two modellers were aware of each other's views while creating their respective models.

While some argue that DES and SD are quite separate simulation approaches (Brailsford and Hilton, 2001), others see them as complementary to one another (Morecroft and Robinson, 2005). In addition, some authors have created hybrid models with discrete and continuous elements (Lee, et al., 2002a; Lee, et al., 2002b;

Rabelo, et al., 2005; Helal, et al., 2007). However, the technical differences between DES and SD modelling need to be acknowledged. These differences result from the underlying principles of each simulation approach and are briefly considered in section 2.3.

Literature on the comparison of DES and SD is distinguished in three main types of studies:

- Comparison of DES and SD models (Mak, 1993; Taylor and Lane, 1998; Sweetser, 1999; Lane, 2000; Brailsford and Hilton, 2001; Morecroft and Robinson, 2005).

- Combined use of DES and SD (Usano, et al., 1996; Petropoulakis and Giacomini, 1998; Music and Matko, 1999; Martin and Raffo, 2000; Donzelli and Iazeolla, 2001; Lee et al., 2002a; Lee et al., 2002b; Lakey, 2003; Stchedroff and Cheng, 2003; Venkateswaran and Son, 2004; Greasley, 2005; Rabelo et al., 2005; Helal et al., 2007). The main areas of applications are: software development, manufacturing and production and supply chain.

- Representation of discrete aspects in SD models (Coyle, 1985; Curram, et al., 2000).

Work on the combined use of DES and SD is increasingly growing and emphasis is given on the complementary use of the two simulation approaches, where each one can be used to represent different aspects of the problem studied.

2.3 Technical differences between DES and SD

While an introduction to the DES and SD approach has already been provided in chapter 1, in this section the key technical differences between the two approaches are pointed out (Table 2-1).

DES views systems as a network of queues and activities, where state changes occur at discrete points of time. Whereas in SD, models are viewed as a system of stocks and flows where continuous state changes occur over time. In DES entities are individually represented and can be tracked through the system. Specific attributes are assigned to each entity, determining what happens to them throughout the simulation. On the other hand, in SD entities are represented as a continuous quantity, and are ‘indistinguishable’. Specific entities cannot be followed throughout the system.

In DES state changes occur at discrete points of time, the simulation clock jumps in unequal time intervals from the occurrence of one event to the next (next-event technique) (Pidd, 2004). In SD software changes occur continuously at equal small segments of time (Δt). The value of these time steps is chosen as a compromise between the accuracy of the differential equations of the system versus the cost of computer time (Coyle, 1985). Furthermore, Coyle points out that time, as a variable is differently handled in the two modelling approaches. In DES time is usually incorporated as an implicit variable, which is usually of no direct interest to the modeller, while in SD modelling time is an explicit variable of direct interest (ibid). However, it can be argued that this depends on the specific problem situation modelled.

The underlying mathematics used in the equivalent simulation software differs. In DES, the underlying mathematics is part of the software language, of which the user has little understanding. SD models are based on differential equations approximated in discrete time slices, which are entered by the modeller as part of the model building process. DES models are stochastic in nature with randomness incorporated through the use of statistical distributions. SD models are generally deterministic and variables usually represent average values.

Despite the above differences, it is claimed that the objective of models in both simulation approaches is to understand the way systems behave over time and to compare their performance under different conditions (Sweetser, 1999).

Table 2-1: Fundamental differences between DES and SD.

Fundamental differences	**DES**	**SD**
System representation	Queues & activities	Stocks & flows
State changes	Discrete	Continuous
Representation of entities	Distinct	A continuous mass
Variability	Stochastic	Deterministic

2.4 A taxonomy of topics considered in the comparison literature

After having discussed the fundamental differences between DES and SD, this section provides a review of existing comparison work, based on the specific issues raised in the literature. This section aims to represent the opinions expressed in the literature about DES and SD. These opinions are built around the model building process and the use of the respective simulation approaches in practice. The review of the existing comparison work is structured based on the aspects involved in the OR modelling process. Before exploring the specific topics found in the comparison literature regarding how the two simulation approaches compare, the stages of the OR model building process are identified as well as some general aspects regarding the simulation modelling process (sub-section 2.4.1). Then the comparison literature by modelling stage is discussed in the relevant sections (2.5 - 2.11), as well as the choice of simulation approach (section 2.12).

2.4.1 The model building process in DES and SD

The stages followed in generic OR modelling consist of the following (Hillier and Lieberman, 1990; Oral and Kettani, 1993; Willemain, 1995):

- Problem definition
- Conceptual modelling
- Model coding

- Model validity
- Model results and experimentation
- Implementation and learning

In the respective DES and SD textbooks, teaching the art of simulation modelling, the steps suggested are equivalent to the stages followed during OR modelling, similar to the list provided above. This implies that both DES and SD modelling follow similar modelling stages.

Regarding the model building process followed, it is mentioned that in DES modelling emphasis is given in the development of the model on the computer (model coding). Baines et al. (1998) completed an experimental study of various modelling techniques, among others DES and SD, and their ability to evaluate manufacturing strategies. The authors commented on the time taken in building the DES model. The time taken in model building was considerably longer compared to SD and other modelling techniques. Furthermore, Artamonov (2002) developed two equivalent DES and SD models of the beer distribution game model and commented on the difficulty involved in coding the model on the computer. He found the development of the model on the computer more difficult in the case of the DES approach, whereas, the development of the SD model was less troublesome. One possible explanation given by Baines et al (1998) is the fact that DES encourages the construction of a more lifelike representation of the real system compared to the other techniques, which consequently results in a more detailed and complex model.

On the other hand, in SD modelling emphasis is given to understanding the system structure and the dynamic tendencies involved. Consequently, Meadows (1980) highlights that system dynamicists spend the most amount of modelling time specifying the model structure. The specification of the model structure consists of the representation of the causal relationships that generate the dynamic behaviour of the system. This is equivalent to the development of the conceptual model.

Another important feature of DES and SD modelling is the iterative nature of the modelling process. In DES and SD textbooks it is highlighted that simulation modelling involves a number of repetitions and iterations (Randers, 1980; Sterman, 2000; Pidd, 2004; Robinson, 2004). The sequence between modelling stages does not follow a linear progression from problem definition to conceptual modelling, model coding, etc. Regardless of the modeller’s experience, a number of repetitions occur from the creation of the first model, until a better understanding of the real life system is achieved. So long as the number of iterations remains reasonable, these are in fact quite desirable (Randers, 1980).

Next, follows a detailed review of the comparison literature for each stage in the simulation modelling process.

2.5 Problem definition

DES and SD are two simulation approaches used to model social or managed systems with the view to understanding the system behaviour. As Sweetser (1999) mentions, both simulation approaches can be used to understand the way systems behave over time to compare their performance under different conditions. Despite the overall common objective, SD is inherently involved in studying the effect of policies on system behaviour. SD is viewed as a dynamic feedback system and studies the interaction of control policies or exogenous events and the model's feedback structure in producing dynamic behaviour (Mak, 1993). It can also be used as a goal seeking tool in making decisions on a particular variable in the model in order to achieve a desired goal. In DES modelling the system under study is assessed with the view to improving system capacity, resource utilisation or queuing time in the system. A 'what if' philosophy is used to answer questions like: "would additional resources in the system reduce the queue size?" (Mak, 1993)

Another facet tackled in the literature is the nature of problems modelled by each simulation technique, 'strategic' vs. 'tactical/operational'. It is believed that SD focuses mainly on strategic policy analysis, while DES is generally used to study problems at an operational or tactical level (Taylor and Lane, 1998; Sweetser, 1999; Lane, 2000). Based on the differences of discrete and continuous systems, Richardson (1991) maintains that the choice of one or the other approach depends on the conceptual difference from which one views the problem. The SD approach is considered appropriate when taking a 'distant' perspective (meaning strategic) where events and decisions are seen in the form of patterns of behaviour and system structures (Richardson, 1991). Rabelo et al. (2005) point out some of the factors that make SD suitable for high level strategic modelling, including a number of unsupported claims, which are generally accepted in the comparison literature. These consist of the following:

- Takes a holistic approach of systems, integrating many subsystems
- Focuses on policies and system structure
- Use of feedback loops to represent the effects of policy decisions
- Represents a dynamic view of the cause and effect relationships among the system elements
- SD has minimal data requirements to build a model.

Several authors suggest that DES is not suitable for strategic modelling, as it does not normally represent models at aggregate levels (Baines and Harrison, 1999; Lee et al., 2002a; Oyarbide, et al., 2003). To cater for this disadvantage, a number of studies (Lee et al., 2002a; Rabelo et al., 2005; Helal et al., 2007) have suggested the use of hybrid simulation approaches combining DES and SD. Rabelo et al. (2005) in their study of an integrated manufacturing enterprise system, used DES to model local production decisions for selected parts of the enterprise, while the SD model captured the long term effects of these decisions on the entire enterprise and the interactions between decisions made at different levels of management. In another study, Lee et al. (2002a) recommended the use of analytical models for modelling at operational levels, DES for modelling at tactical level, while for modelling at strategic levels recommended the use of hybrid discrete and continuous simulation models. The same authors created a discrete event and a hybrid discrete-continuous simulation model of a supply chain and concluded that the discrete event model overestimated the outputs of inventory levels. The authors recommended the use of hybrid simulation models to model supply chain simulation models, which were shown to be neither completely discrete nor continuous systems.

On the other hand, various authors have expressed the view that, even though it has not yet been adequately exploited, SD can be successfully used in modelling operational systems. For example, Han et al. (2005) represented an operational SD model of an earth-moving system in a construction management study and compared it with an equivalent (already existing) DES model. Their study suggests that an SD-based operational model can address the operational aspects of the model as accurately and reliably as a DES-based model. The advantages of using SD at an operational level are discussed. These include modelling of feedback effects, managerial actions and soft variables. Furthermore, while the use of SD has been rarely considered in manufacturing systems modelling, Oyarbide et al. (2003) comment on the potential of using SD modelling in this context. Taking into consideration the inherent characteristics of the two modelling techniques, the authors suggest that SD would be a better choice in the intermediate stages of decision making when less detailed models or results are required. Some of the advantages of SD modelling with respect to the requirements of decision making at intermediate stages of evaluation are: the simplicity of the data required, ease of building a simulation model and reduced execution time. Obviously, these are statements which represent authors' opinions and have not been empirically verified for their accuracy.

In a study of a manufacturing plant, Greasley (2005) reports on the successful use of the DES approach to investigate the operational aspects of a production-planning facility. The outcome of the DES study was the recommendation of new production sequencing activities. In addition, as a result of the study, it emerged that the disruptions in production planning in the manufacturing plant needed to be further considered. In this case, the SD approach was preferred in order to model the softer aspects related to the problem of disruptions. Greasley considered the SD approach useful in modelling the organisational context of the problem and so extended the already created DES model, using SD.

It is clear that the opinions expressed by the authors referred to in this section tend to suggest that SD modelling is more appropriate in modelling at a strategic level, while DES at a tactical/operational level, however, no empirical evidence has been found to verify these opinions.

2.6 Conceptual modelling

After articulating the problem, the next step in a simulation study is to define the conceptual model, derived from the modellers' mental model of the system, which is then transferred into the simulation software. In SD this stage is also referred to as formulation of the dynamic hypothesis (Sterman, 2000). During this stage, based on the modelling objectives, the boundaries of the system are set by including, inputs, outputs, contents, assumptions and simplifications of the model. The aspects pertaining to conceptual modelling and which come up in the comparison of DES and SD are: diagramming methods, system representation, representation of people and feedback effects. These are discussed in the following paragraphs.

2.6.1 Diagramming

Some diagramming methods used to define SD conceptual models are: model boundary charts, subsystem diagrams, causal loop diagrams, stock and flow diagrams and policy structure diagrams (Sterman, 2000). However, causal loop diagrams or influence diagrams are most often used in practice. Conceptual diagrams are used in order to understand the feedback structure of the system. These diagrams are used to understand the broad system structure and are therefore, kept intentionally simple (Pidd, 2003). These are also called qualitative models (Wolstenholme, 1990), which can at times be adequate to understand the problem situation and thus a further computer model might not be necessary (Brailsford and Hilton, 2001).

With respect to DES modelling, it is suggested that there are no set diagramming methods for representing models (Morecroft and Robinson, 2005). They vary from activity cycle diagrams, process mapping/process flow diagrams, logic flow diagrams, to Petri nets, unified modelling language (UML), digraphs, object models, event graphs, etc. (Robinson, 2004; Onggo, 2007; Onggo, et al., 2008). The most frequently used diagramming methods are process flow diagrams and activity cycle diagrams[4]. Creating a conceptual model is considered beneficial in DES modelling in order to keep the model focused on project objectives and to ensure that the model achieves its requirements (Robinson, 2004).

Mak (1993), in her doctoral thesis, investigated the conversion of DES activity cycle diagrams into SD stock and flow diagrams. She developed a set of conversion guidelines, which were incorporated in the prototype automated conversion software she developed. DES process flow diagrams, which could be considered as more close to stock and flow diagrams were not included in the study. Mak pointed out that SD modelling structures are more flexible than DES Activity Cycle Diagram modelling. In a SD casual loop diagram, one can add as many auxiliary variables and as many information links as necessary in order to represent a situation. While in a DES activity cycle diagram only alternating activities and queues are allowed. However, at a later point, Mak (1993) comments on the flexibility of DES modelling, which allows the modeller to manipulate the events

[4] The activity cycle diagram describes the logic of the simulation model and shows the life cycle of the entities in the system, passing through activities and queues.

and hence its flexibility in representing different components and activities in the model.

2.6.2 Feedback effects

In SD, models are viewed as closed systems, where the outputs have effect on the input and are represented by "a series of stocks and flows" (Brailsford and Hilton, 2001). The system's behaviour is determined by the internal structure of the system, the causal relationships of endogenous variables incorporated into feedback loops (Sweetser, 1999; Morecroft and Robinson, 2005). In SD modelling, the focus is on the feedback processes affecting the changes to the outputs of interest (Taylor and Lane, 1998). Therefore, feedback is an important part in SD modelling. On the other hand, in DES, systems are viewed as "networks of queues and activities" (Brailsford and Hilton, 2001). It is generally claimed that DES follows an open loop structure and feedback is not modelled (Coyle, 1985). It has been argued, however, that feedback is involved in DES models (Sweetser, 1999; Lane, 2000; Morecroft and Robinson, 2005). Robinson (2004, pp.7) explains how feedback effects are present in a DES model, taking a simple example of a Kanban system, where a machine feeds a buffer. The rate at which the machine works affects the number of parts in the buffer, which in turn affects the speed at which the machine works. However, all these effects are hidden behind the computations of the simulation software and are not specifically considered by the modeller or the user. Hence, even though feedback may exist in a DES model, it is not made explicit to the users and modellers are less interested in the events that cause the changes (Sweetser, 1999; Lane, 2000; Morecroft and Robinson, 2005).

2.6.3 System representation

With respect to the representation of systems, it is generally accepted that DES takes an analytic view, whereas SD takes a holistic view of a system's performance. It is believed that SD tends to represent abstract and general systems, while DES can evaluate a variety of issues at a low level of detail (Baines et al., 1998) and so models tend to have a narrower focus (Sweetser, 1999). These beliefs can be explained by the respective philosophy taken during modelling by each simulation approach.

More specifically, SD takes a systems' thinking perspective. The system is seen as a collection of parts and their underlying interrelationships (Bellinger, 2004). SD focuses on the emerging system behaviour over time, by exploring the dynamic implications of the underlying structure of the system. Furthermore, due to the use of systems' thinking in conceptualising the model and the ability to include the interrelationships between various factors, SD is considered more suitable for the representation of systems with a wider focus. Abstract models can be developed with the use of approximations (one such example is the use of average values) and subsequently accurate models are not required (Baines et al., 1998). This approach typifies holistic thinking.

In DES however, a reductionist approach is taken (Han et al., 2005), where system understanding is achieved in terms of its components. DES models break a system down into its constituent parts (Lane, 2000). DES is more oriented to representing distinct objects/people, scheduled activities, queues and decision rules (Brailsford and Hilton, 2001) in the system and the associated interdependencies. Efforts in conceptual modelling do not focus on identifying the interrelationships between the parts of the system. It is hence suggested that DES is more suitable in representing detailed and well-defined processes (Baines et al., 1998). This approach typifies an analytic way of thinking.

2.6.4 People

In SD models, the entities are ‘indistinguishable’ (Borshchev and Filippov, 2004) and the aggregate behaviour of the system population is examined. The SD approach is particularly preferred in the case of models with a very large population (Brailsford, et al., 2004). On the contrary, in DES the entities are distinctly represented and their behaviour in the system is individually modelled. The characteristics attributed to the entities determine their progress in the system from the time they enter until they exit. The history of each entity in the model can be observed and state changes recorded. As a result, Pidd (2004) suggests that DES models are more appropriate when the tracking of individual entities is important.

2.7 Model coding

Model coding involves the conversion of the conceptual model into a computer model. In DES and SD modelling, model coding involves creating the model using the relevant computer software. In this section, statements made regarding aspects of DES and SD model coding on the computer, are discussed. These are: model complexity, modelling structures and other model elements.

2.7.1 Model complexity

Looking at both simulation approaches in terms of model complexity, it is maintained that DES is more concerned with detailed complexity, while SD with dynamic complexity (Taylor and Lane, 1998; Lane, 2000). This is due to their inherent features, where DES can model great complexity and detail, representing specific individuals and the subsequent interactions, while SD represents the aggregate picture of the system.

In DES, complexity is the result of multiple random processes and the endogenous structure of the system (Lane, 2000; Morecroft and Robinson, 2005). DES models represent systems of small operational tasks or individual items, which comprise distinct entities with multiple attributes, individually defined. Complexity results from the interconnections and effects between variables. In SD, a model's behaviour is determined by the feedback structure and dynamic complexity arising from the influences among endogenous variables. SD models represent systems consisting of causal relationships of variables (the latter are aggregated here and contain relatively few attributes, resulting in low detail complexity). Consequently, dynamic complexity arises due to "non linear, delayed and accumulative/draining causal relationships" (Lane, 2000). Therefore, SD models produce counter-intuitive behaviour.

Let us consider the dynamic complexity resulting from the existence of several nonlinear relationships in a SD model. In this type of system, under a set of conditions, one part of the model becomes more active and under other conditions another part dominates. For illustration purposes, let us consider the well-known fishery example (Morecroft and Robinson, 2005). There are two main non-linear loops in the system, the reinforcing loop of natural fish regeneration and the balancing loop of fish catch depending on the ship fleet size (Figure 2-1). Depending on the size of the fish population in the sea, the dominance of the two loops in the model changes resulting in an s-shaped graph of fish catch. Fish catch initially increases exponentially with the increase of fishing ships due to the fact that the fish regeneration loop dominates. However, the fish population and the catch rate start dropping exponentially after a point where the harvest rate becomes equal to the fish regeneration rate and thus the balancing loop of fish catch becomes more dominant. This explains the collapsing fish population, referred to as the system's puzzling dynamics (Morecroft and Robinson, 2005).

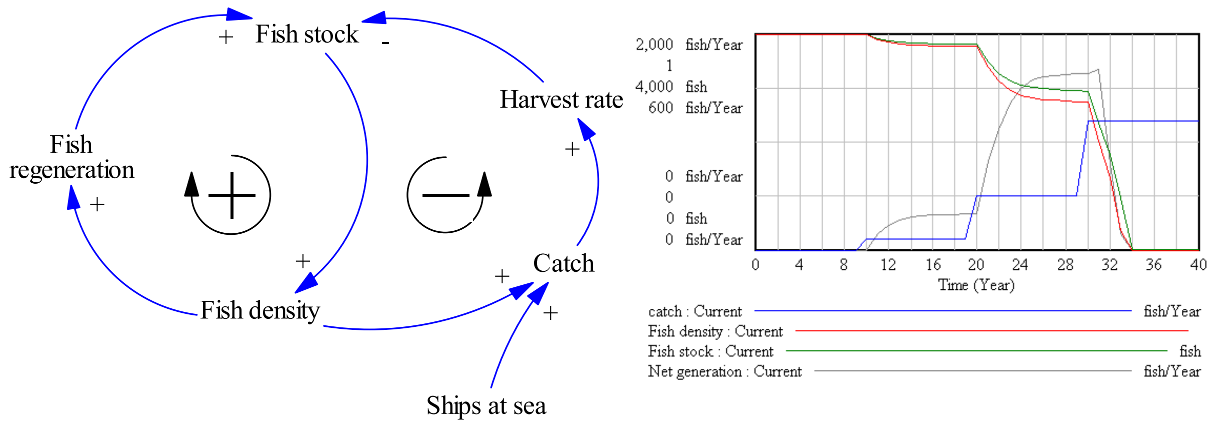


Figure 2-1: Simulation of a harvested fishery with stepwise changes in fleet size (Morecroft and Robinson, 2005)

Taking a philosophical view at model representation, Morecroft and Robinson (2005) in their empirical study of a fishery model, maintain that SD deals with ‘deterministic complexity’, whereas DES with ‘constrained randomness’. While in the SD model the system’s behaviour is predetermined by the feedback structure, the interaction among endogenous, deterministic variables, the system’s future behaviour is unknown to the subjects in the system. In the DES model, system behaviour is affected by “endogenous factors and also by random operational factors”. The future behaviour “is assumed to be partly and significantly a matter of chance” and consequently complexity arises from multiple random processes. This represents different worldviews taken inherently by each approach, which results in the specific modelling practice followed (Morecroft, 2007). While in SD the system behaviour is explained by determining the underlying feedback structure and performance is improved by re-designing polices, in DES the interacting random processes are primarily investigated in order to find alternative ways of improving the stochastic structure of the system or managing the variability better.

2.7.2 Modelling structures

Another practice referred to in the SD field is the use of already existing modelling structures in the representation of decision making. Decision making represents the aggregate judgements of actors in the system and decision rules can be specified as part of the modelling process. Decision makers create a view based on purposive or judgmental information. The well-known 'asset stock adjustment' process is a central structure in the representation of feedback in business and social systems (Sterman, 2000). In DES there are no model building principles to incorporate the decision making processes, but insights from discussions with and observation of decision makers can be incorporated in the model with the use of additional formulae or decision rules. According to Lane (2000), human agents in SD are modelled as bounded rational policy implementers, whereas in DES as decision makers. However, Morecroft and Robinson (2005), maintain that in both simulation approaches decisions are made subject to bounded rationality, in SD taking into account information based on objective evidence (not implied) readily available to actors in the system. In DES decisions are made based on the uncertainty of future random events. For example, in their SD fishery model the decision to buy new ships is made referring to the catch without taking into consideration the fish population or regeneration rate, whereas in the DES model the same decision is made based on total catch in the fishery, making the assumption that the fishery system will continue to demonstrate the same patterns of behaviour in the future.

2.7.3 Other model elements

Another aspect of interest in SD modelling is the handling of delays. Both DES and SD represent delays in social systems. However, this may have different meanings to DES and SD modellers. In DES delays are represented in the form of queues or buffers, where elements or parts of the system wait until the next activity (work-centre or machine) becomes available. In SD delays are represented in the form of the time lag between taking a decision and its effects on the state of a system (Sterman, 2000). As a result, referring back to the information/action/consequences loop (Figure 1-2), decision makers continue to intervene to correct the perceived discrepancies by recruiting people, even after sufficient action has been taken to restore equilibrium in the system. Due to the fact that new recruits undertake training which is an example of a delay, the results of the action taken, that of recruitment, is not instantly obvious in the system. Consequently, there will be more employees than desired, which will influence the occurrence of lay-offs in the future. Therefore, delays can cause instability (overshoot or oscillation) in the system, and a further slow down in the rate of learning (Sterman, 2000). The simplest type of delay is the exponential delay, which is represented by the fraction of the stock level and the length of the delay time.

In both simulation approaches material flows can be incorporated in the model. These are measurable and conserved throughout the system. SD models can also include information flows, which can be part of the feedback loops, whereas, in DES models information flows can be incorporated with the use of priority rules or attributes, but these are not obvious to the users (Mak, 1993).

2.8 Data inputs

DES is generally considered a ‘data-driven’ or ‘data hungry’ approach (Baines et al., 1998; Sweetser, 1999; Rabelo et al., 2005). It is often mentioned that it requires large amounts of quantitative, numerical data and the statistical estimation of model parameters (Brailsford and Hilton, 2001). Whereas statistical estimation of data are less often used in SD modelling (Meadows, 1980). The data required to build a DES model are mainly historical or estimates of the system’s future performance derived mainly from concrete and observable processes. Apart from quantitative data, SD models can incorporate qualitative aspects of behaviour, which while difficult to quantify, significantly affect system performance (Sweetser, 1999). Therefore, SD modellers are considered to be more comfortable with incorporating in their models ‘best guesses - anecdotal data’ (Sweetser, 1999), ‘soft’ variables (Brailsford and Hilton, 2001) or ‘judgmental information’(Lane, 2000). Qualitative variables represent factors for which numerical metrics and data are not available such as goals, perceptions and expectations (Sterman, 2000). For example, the variable ‘hunger’ cannot be mathematically quantified (Brailsford and Hilton, 2001). However, the mathematical relationships between the variables amount of food eaten and hunger can be represented in a graphical form, based on the amount of food in the stomach, the more we eat, the less hungry we feel (ibid).

It is believed that DES is more suitable in modelling ‘hard’ data in great detail, while SD modelling is more appropriate in representing systems at a higher scale involving some level of aggregation. For example, in his study Greasley (2005) found the DES model useful in completing a ‘hard’ technical analysis of the production system, however, in order to deal with the softer issues related to the organisational context of the problem, that of reduced delivery performance, the SD approach was preferred. Furthermore, Rabelo et al. (2005) point out that SD is more suitable in modelling continuous and qualitative parameters, which is the case with top level management decisions, whereas others believe that DES modelling faces challenges in dealing with these sorts of variables, while it is suggested to be better at dealing with a high level of granularity, involving detailed and accurate data (Helal et al., 2007).

DES models usually contain random variables and are stochastic in nature. Randomness is considered an important aspect in DES modelling. It is usually added by incorporating statistical distributions to the events and entities of the model. SD models generally depict deterministic behaviour, where averages of variables are used and, therefore, the aggregate behaviour of the system is depicted. Stochastic features of the system can be added with the use of distributed delays (Brailsford and Hilton, 2001), which can be portrayed through a range of available statistical distributions. However, stochastic elements are rarely investigated by system dynamics modellers who are more interested in feedback dynamics (Morecroft, 2007). The SD paradigm is reluctant to disaggregate the quantities into distributions and therefore, system dynamicists are more likely to ignore randomness (Meadows, 1980).

The type of relationships between variables, represented in the DES and SD models is also considered in the comparison literature. It is mentioned that SD models can represent linear and non-linear functions (Morecroft and Robinson, 2005). Similarly DES models can represent non-linear relationships. However linear relationships are most commonly used. According to Sweetser (1999) and Morecroft and Robinson (2005), DES can also model continuous systems containing feedback structure and non-linear relationships, but this has not been frequently seen in practice. Nonlinear relationships are considered to be an important feature of SD models, which can change the strength of feedback loops depending on the state of the system (Meadows, 1980; Sterman, 2000). An example of a simple nonlinear relationship is the relationship between the inventory level and production rate, an increase in inventory levels will reduce production rate, but the latter can never become zero (no matter how big the excess in inventory is). Another example is the non-linear relationship between the density of the fish in the sea and the catch per ship. When the density is high, the catch per ship is stable, as the density decreases, the catch per ship falls, following an exponential reduction rate, reaching almost 0 (Morecroft, 2007).

Morecroft and Robinson (2005), in their SD and DES fishery models, compared the representation of the growth patterns of the fish stock in the two simulation models. In the SD model, the growth patterns were represented in an S-shaped graph determined by the non-linear function of net fish regeneration depending on fish density. The new fish per year falls as the population density rises and thus reducing the population growth, as the fish stock reaches its maximum sustainable value. In the DES model growth was determined by a linear, but random, function of the number of fish in the sea, limited by a discrete cut-off number of fish that could be sustained in the sea. This structure results in an equivalent s-shaped growth, reached in a non-asymptotic manner towards the allowed limit of fish in the sea, while in the case of the DES model this is not a smooth line, but reaches the limit in a discrete step.

2.9 Model validity

Model validation is considered an important aspect in simulation modelling and in OR more generally (Landry, et al., 1983). A number of techniques are used to test the models created. Overall model validation in DES and SD modelling undertake similar objectives, checking that the model created is sufficiently representative of the real system to answer the specified objectives or questions. With respect to the established model validation practice, DES is considered to take a black-box approach (Lane, 2000). Validation in DES modelling focuses on model outputs, checking that the model is accurate or close to reality for the particular objectives of the study. Whereas, SD modellers take a white-box approach, considering whether the appropriate underlying system structure is incorporated in the model (Lane, 2000). References are made to the ability of the model to represent the reference mode that is the behaviour observed in the real system. Rigorous statistical tests are not the norm, even though SD textbooks suggest the use of a number of statistical tests (Randers, 1980; Sterman, 2000). Model validity is mostly considered qualitatively and informally. Sweetser (1999) mentions that SD models are evaluated on the face validity of models' outputs. References are also made to model utility (Meadows, 1980). This implies that validity is measured by insights or the improved understanding derived from using the models and their outputs.

In terms of the use of models, model validity is considered a measure of the users' confidence in the model (credibility). Credibility is seen in terms of representativeness (Robinson, 2004, p. 231), confidence in the results and confidence in using the model for decision making (Robinson, 2004, p. 214). Randers (1980) rates SD models as highly representative, compared to a predictive model (including DES). It is generally accepted that both simulation approaches are concerned with building models which are representative of reality, providing confidence in the results and in decision-making. Indeed, Akkermans (1995) argues that in most cases DES and SD can represent the real world with equal validity. However, different views still exist. For example, Baines et al. (1998) state that DES models tend to provide a higher level of accuracy and credibility due to on screen animation. Meanwhile, SD models are considered less accurate due to the inherent approximation of treating a product as a flow (Baines et al., 1998).

2.10 Model results and experimentation

The concept of equilibrium and non-equilibrium states is viewed differently in both approaches. SD models are generally interested in equilibrium states (Lane, 2000), and equilibrium states can be achieved by changing the numbers of incoming and outgoing flows. In a DES model a non-equilibrium state is achieved with the inclusion of randomness in the model. If fixed inputs are included then fixed outputs can be derived. In the DES and SD fishery models, Morecroft and Robinson (2005) report that both non-equilibrium models depicted a collapsing fish population, where the SD system behaviour was determined by the change in the number of fishing ships (outflow of fish stock), while in the DES model by the random variation of fish regeneration and catch. In DES practice, however, modellers are interested in getting a steady state model, where the outputs of the model vary according to a specific distribution (Robinson, 2004). A number of DES models (excluding here transient models[5]), start from an empty (unrealistic) state and do not reach a steady state condition until the model is initialised to reach normal steady state conditions. There are two different ways of achieving a steady state in DES. The first way is to run the model for a warm-up period, ensuring that the model reaches a realistic condition. The second approach is to set initial

[5] A transient model is usually a terminating model, where the distribution of the output is constantly changing.

conditions in the model so that the model starts in a realistic condition (Robinson, 2004).

With respect to the results of interest from the simulation models, it is pointed out that SD models provide a full picture of a system in the simulated period (Mak, 1993). Model outputs are read not as much for the quantitative predictions of particular variables as for the qualitative behavioural characteristics of the system (Meadows, 1980). SD models are mainly concerned with policies rather than decisions and model outputs consist of decisions concerned with the qualitative improvements of system performance (Lane, 2000). However, SD models can also provide quantitative output measures, but it is believed that point predictions are rarely made (Sweetser, 1999). In DES modelling emphasis is given to point predictions, with outputs providing statistically valid estimates of the system's performance measures (Sweetser, 1999; Brailsford and Hilton, 2001; Law, 2007). Nevertheless, DES models that involve a qualitative analysis of the problem have been reported (Robinson, 2001; Swenseth, et al., 2002). DES models generally provide a wide range of outputs, principally of a quantitative nature. Additionally, the interpretation of DES model results requires some statistical analysis. The outputs of one simulation run represent only one possible outcome due to the randomness in the model. For this reason, a practice often used in DES is running many iterations of the model with the use of different random number seeds – multiple replications (Pidd, 2004; Robinson, 2004). In order to make an appropriate analysis of the DES output, the model user should have some statistical background (Sweetser, 1999; Brailsford and Hilton, 2001).

When looking at model results, due to the inherent features of the two modelling techniques, different aspects of the model can be picked up by the users. As mentioned in section 2.8 (paragraph 3), DES models contain random variables and are stochastic in nature, while SD systems generally depict deterministic behaviour. Therefore, SD model results are considered as a source of understanding the reasons that cause changes in the system's performance, resulting from counter intuitive effects of the system's structural behaviour (Morecroft and Robinson, 2005). Meanwhile, DES modellers and model users are less interested in the events that actually cause these changes and focus more on the numerical results (Sweetser, 1999).

With respect to the accuracy of model results, it is believed that DES models provide more accurate results than SD models. However, there have been studies that suggest that an SD model can provide equally accurate results (Han et al., 2005).

2.11 Model implementation & learning

Both simulation approaches can be used to understand how systems behave over time (Sweetser, 1999). However, contradictory statements are made regarding the level of understanding that users can gain from using these models. According to

Brailsford and Hilton (2001), DES models are transparent to the clients. Animation and on-screen displays can provide useful insights regarding the model's structure. Lane (2000) argues that while DES models are convincing to the client, users do not necessarily understand the underlying mechanics of the model. On the other hand, Lane (2000) states that SD models are more transparent and compelling to the client. Randers (1980), in a comparison of system dynamics and modelling used for prediction (applicable to DES models), rated SD as having a higher capacity to increase clients' (users') understanding and learning, calling it 'insight generation capacity'. On the other hand, in SD models visual animation is limited, including mainly on screen updates of stocks, graphs and numbers. The user relies to a great extent on graphs and numerical displays (Sweetser, 1999).

Another important factor mentioned in the literature when comparing DES and SD is model usefulness. The concept of learning from using simulation models is widely mentioned in the SD literature (Forrester, 1961; Morecroft and Sterman, 1994; Rouwette, et al., 2002). Business flight simulators are considered to be appropriate 'learning laboratories' that can help managers gain insights regarding their businesses operations. On the contrary, DES models are seen mostly as the domain of simulation experts and are less used as learning tools by non-technical managers (Sweetser, 1999). However, these statements are made by modellers without considering users' opinions about specific models.

Both simulation approaches can be used as tools to facilitate the communication of ideas in group discussions. SD modelling has been traditionally reported as an interactive process involving the modeller and the client, the people who are part of the system under study (Vennix, 1996). Similarly, DES simulation modellers are advised to build their models by interacting with problem owners (Robinson, 2004). In addition, evidence of using DES as a learning tool and in group discussions exists in the literature (Robinson, 2001). Furthermore, animation and graphics facilities provided by DES software are deemed "very useful for communication with clients" (Brailsford and Hilton, 2001).

2.12 Choice of simulation approach

The choice of simulation approach is rarely discussed in published papers. It is often suggested that the choice should be made based on specific criteria. Few studies discuss these criteria, but there is not yet a comprehensive guidance distinguishing specifically the choice between DES and SD models. In their study comparing DES and SD, Brailsford and Hilton (2001) conclude with a tentative list of criteria to assist in the choice between the two simulation approaches. These deal mainly with the type of problem as well as model purpose and requirements. These are: the scope, importance of variability and of tracking individuals, population size, control of queues or rates, timescale and model purpose. Furthermore, Brennan et al.(2006), in their study of models for economic evaluation of health technologies, provide guidance for the choice among a number of modelling approaches, including DES and SD. The authors developed a taxonomy, where models are classified based on specific features (Table 2-2). On the horizontal level, models are distinguished into heterogeneous versus individual level models, which can involve either markovian or non-markovian states and on the vertical level, models are distinguished into non-interaction versus models including interaction. It is suggested from this study that the choice is made based on decision makers' requirements such as output requirements as well as system characteristics involving population size, system complexity, the existence of resource constraints, etc.

Table 2-2: Taxonomy of model structures (Brennan et al., 2006)

			A	B	C	D
			Cohort/aggregate level/counts		Individual level	
			Expected value, continuous state, deterministic	Markovian, discrete state, stochastic	Markovian, discrete state, individuals	Non-Markovian, discrete-state, individuals
1	No interaction allowed	Untimed	Decision tree rollback	Simulated decision tree	Individual sampling model: Simulated patient-level decision tree	
2		Timed	Markov model (evaluated deterministically)	Simulated Markov model	Individual sampling model: Simulated patient-level Markov model (variations as in quadrant below for patient level models with interaction)	
3	Interaction allowed	Discrete time	System dynamics (finite difference equations)	Discrete time Markov chain model	Discrete-time individual event history model	Discrete individual simulation
4		Continuous time	System dynamics (ordinary differential equations)	Continuous time Markov chain model	Continuous time individual event history model	Discrete event simulation

On the other hand, it has been reported that the choice of simulation approach depends on the particular situation and the decision-maker preferences and knowledge rather than the modelling capabilities of either approach or the nature of the problem (Mak, 1993; Sweetser, 1999). From the users' perspective, another view of interest has been expressed by Akkermans (1995), who considered different types of modelling in business (DES, SD and spreadsheets) for real case scenarios. He claimed that as part of the model building process, the choice of modelling approach is not highly important. He also adds that the clients are usually not concerned about the choice of the simulation software used in a modelling project.

Taking into consideration the above, it is important to fully understand the differences between the two specific simulation approaches (DES and SD). From the literature explored here, it is evident that there is a need to go beyond opinions and to get empirically verified claims on the differences between DES and SD.

2.13 Summary of comparison literature

In this chapter, literature on the comparison of DES and SD has been reviewed. The fundamental differences, resulting from the underlying mechanics of each simulation approach are initially discussed. Then the model building process is considered, suggesting that similar stages are followed during DES and SD modelling, but differences exist in the amount of attention paid to various stages. The statements found in the literature are summarised in three separate tables, based on the three specific themes that are identified in sections 2.5 – 2.11. These consist of: model building, modelling philosophy and model use.

- Model building (Table 2-3) deals with the opinions expressed with respect to the model building practice established in the two simulation approaches. These are seen mostly in the form of the approach taken as part of the stages involved in DES and SD modelling.

- Modelling philosophy (Table 2-4) deals with the philosophical aspects involved in DES and SD model representation.
- Model use (Table 2-5) is concerned with the opinions expressed on the comparison of the two modelling approaches with respect to how useful these are deemed from the users' perspective.

There appears to be a general level of agreement on the nature of the differences. However, there exist exceptions and contradictions to these generally expressed opinions, hence, the differences might not be that clear cut as indicated in tables 2-3, 2-4 and 2-5. Even though an attempt has been made to represent the diversity of the opinions expressed, in these tables the generally accepted views are mostly provided. Based on the opinions expressed in the literature the current thesis sets out to empirically validate the accuracy of these statements.

In this chapter the limitations of the existing comparison are also discussed and it is noted that these consist of generally accepted statements, which represent the authors' personal opinions arising from their own area of expertise. Consequently, the need for an empirical study to validate these generally accepted statements is identified. This serves as the focus of the current study. An empirical study is considered important in order to gain an understanding of the two simulation techniques and to provide an evidence base for making comparisons.

Table 2-3: Summary of comparison literature on DES and SD approach to model building.

Model Building	DES	SD
Problem definition		
1. Modelling objectives	*To understand how systems behave over time under different scenarios.*	
2. Model level	*Operational, tactical.*	*Strategic.*
Conceptual modelling		
3. Modelling/ diagrams	*Not a set diagramming method (activity cycle diagrams, process mapping, etc.).*	*Standard diagramming method (causal loop diagrams).*
4. Feedback	*Feedback can be included, but is not explicit.*	*Causal relationships/ feedback structures & delays essential part of model.*
5. System representation	*Analytic view.*	*Holistic view.*
6. Objects/People	*Distinctly represented. Not practical for modelling large populations. History of each entity can be recorded.*	*Indistinguishable, aggregated. Can examine cases with large population size. Aggregate behaviour is examined.*
Model elements/coding		
7. Complexity	*Narrow focus, with great complexity & detail.*	*Wider focus, general and abstract systems.*
8. Modelling structures	*No existing model building structures. Information from observations & discussions with decision makers is added.*	*Standard modelling structures e.g. 'asset-stock-adjustment'.*
9. Material/ information flows	*Mainly material flows. Information flows can be incorporated, but not obvious.*	*Material & information flows can be equally represented.*
10.Delays	*Delays are represented in the form of entities delayed in the system (in queues or buffers)*	*Delays are represented in the form of delayed influence.*
Data inputs		
11.Type of data	*Quantitative. Concrete/observable processes modelled.*	*Quantitative and qualitative. Opinions & best guesses can be incorporated.*
12.Randomness	*Randomness essential part of the model added by statistical distributions.*	*Randomness is not important.*
13.Relationships between variables	*Can represent non-linear functions, but linear relationships are more common.*	*Linear and non-linear relationships modelled.*
Verification & validation		
14.Scope of validation	*To check that the model created is sufficiently representative of the real system* *Concerned with developing models accurate/close to reality for the particular objectives of the study*	*Concerned more with model usefulness*
15.Validation approach	*Emphasis on model outputs – 'black box' approach.*	*Emphasis on internal structure – 'white box' approach.*
Model results & experimentation		
16.Equilibrium or steady state	*Often interested in steady state conditions, derived from randomness in the model.*	*Interested in equilibrium states.*
17.Type of results	*Quantitative, point predictions and optimisation of performance criteria.*	*Qualitative and quantitative results, point predictions rarely made.*
18.Experimentation	*Compares the performance of alternative scenarios.* *'What if' philosophy employed.*	*Study interaction of control policies/ exogenous events & model feedback structure. Can be used as a goal-seeking tool.*

Table 2-4: Summary of comparison literature on DES & SD modelling philosophy.

Modelling Philosophy	DES	SD
1. System behaviour	*Determined by the interaction of random processes & endogenous factors.*	*Determined by internal structure – causal relationships & feedback loops.*
2. Structure	*Open-loop structure.*	*Closed-loop structure.*
3. Human agents	*Both represent rationally bounded human agents* *Arising from unknown stochastic events.*	*Arising from purposive judgmental information.*
4. Behaviour of interest	*Model deals with constrained randomness.*	*Model deals with deterministic complexity.*

Table 2-5: Summary of comparison literature on the use of DES and SD models.

Model Use	DES	SD
Model Understanding		
1. Understanding (parts of) the model	*The client does not understand the underlying mechanics.*	*Models (links & flows) are transparent to the client.*
2. Animation	*Animation and graphic tools help model understanding*	*No animation. Visual display of model aids model understanding.*
Complexity		
3. Level of detail	*Emphasis on detail complexity.*	*Emphasis on dynamic complexity.*
4. Feedback	*Feedback is not explicit*	*Feedback effects are clear to the client.*
Model Validity		
5. Credibility	*Both models are perceived as representative, provide realistic outputs and create confidence in decision making.*	
Model Usefulness		
6. Learning tool	*DES models are less used as learning tools.*	*SD models, so-called 'learning laboratories', enhance users' learning.*
7. Strategic thinking	*DES models are mostly used in solving operational/tactical issues*	*SD models aid strategic thinking.*
8. Communication tool	*Both DES and SD models are seen as good communication tools and facilitate communication with the client*	
Model Results		
9. Nature of results	*DES provides statistically valid estimates of system's performance. Results aid instrumental learning.*	*SD model results provide a full picture of the system. Results aid conceptual learning.*
10. Interpretation of results	*More difficult, requires users to have statistical background.*	*Outputs are easily interpreted, little or no statistical analysis is required.*
11. Results observation	*Randomness/variation of results is explicit.*	*Generally deterministic results, which convey causal relationships between variables.*

Chapter 3: Project scope and methods

3.1 Introduction

In the previous chapter, existing work on the comparison of DES and SD is explored, providing a list of the issues raised in the literature. This chapter sets out the scope of the thesis. It is particularly concerned with expressing the overall aim of the research and the specific research objectives. The hypotheses, in the form of generally accepted views regarding the comparison of DES and SD are stated, based on the findings from the literature review. Furthermore, suitable research methods relevant to each research objective are considered.

3.2 Research aims

The overall aim of this research is to provide an empirical comparison of the two simulation approaches, DES and SD. The literature review has shown that little comparison work exists. The few studies found consist mainly of personal opinions and generally accepted statements based on the authors' personal area of expertise. The opinions expressed are concerned mainly with issues such as: the suitability of each modelling approach, the established modelling practice and philosophy and the perceived model usefulness. However, limited work has been done to confirm, or refute, the opinions expressed in the literature. Albeit, the thesis does not aim to provide answers regarding the choice between DES and SD, it does aim to get an empirical understanding of the differences and similarities of the two simulation approaches. This in turn, can help in providing a common basis of understanding, bringing the two fields closer and opening the route to mutual communication. Therefore, this research contributes towards the comparison literature by providing an empirical basis for comparing the model building process and the use of simulation models in SD and DES. Hence, the objective of this research is to answer the main research question:

"What are the differences and similarities between the DES and SD modelling approaches and the use of DES and SD models?"

The criteria used for the comparison of the two simulation approaches are based on the aspects revealed by the existing literature (chapter 2) and deal specifically with the model building process and model use. The aspects considered are:

1) The model building process followed during DES and SD modelling, regarding:

   - The amount of attention paid to the different stages during modelling
   - The sequence of modelling stages followed
   - The pattern of iteration followed among the different modelling stages

2) Approach to model building taken during the different modelling stages:

- Problem definition
- Conceptual modelling
- Model coding
- Data inputs
- Verification & validation
- Model results & experimentation

3) The use of DES and SD models, with respect to the following criteria:

- Understanding derived from using equivalent DES and SD simulation models
- Perceived complexity of equivalent DES and SD simulation models
- Credibility in using equivalent DES and SD simulation models
- Perceived usefulness of equivalent DES and SD simulation models in terms of learning, strategic thinking and communication of ideas
- Results interpretation of equivalent DES and SD models outputs

Note that in this work ‘equivalent models’ are deemed to consist of a typical DES or SD model of the same problem situation, but not exactly matching models developed using DES and SD.

The current study aims to provide a better understanding of the two modelling approaches, DES and SD, their similarities and differences. The underlying aim is to provide a basis upon which decisions over choice of simulation approach can be made; albeit that this thesis does not go as far as to specifically address this question. Hence, this research benefits practitioners and academics in an effort to understand the similarities and differences between DES and SD modelling from the modeller's and users' point of view.

Next the specific research objectives and research hypotheses are identified (sections 3.3 and 3.4, respectively).

3.3 Research objectives

The research objectives are based on the main aspects identified with respect to the comparison of DES and SD in the literature, involving model building and model use. Modelling philosophy is not specifically considered in this work. In the model building study, discussions on modelling philosophy do not naturally occur in the context of the method chosen, that of Verbal Protocol Analysis (section 3.5.1). Hence modelling philosophy is not the focus of this research. The objectives of the empirical work undertaken in the current thesis are:

1. To determine how different the modelling process followed by DES and SD modellers is.

2. To establish the differences and similarities in the modelling approach taken by DES and SD modellers in each stage of simulation modelling.

3. To assess how different DES and SD models of an equivalent problem are from the users' point of view.

3.4 Research hypotheses

In order to address the research objectives in section 3.3, the research hypotheses are formulated based on the literature discussed in chapter 2.

3.4.1 Hypotheses for the comparison of the DES & SD model building process (objective 1)

The hypotheses regarding the model building process are related to research objective 1, which aims to compare the model building process followed by DES and SD modellers. The assumptions made in these hypotheses (Table 3-1) are based on the literature reviewed (section 2.4.1). It is expected that different levels of attention are paid to the different modelling stages by DES and SD modellers. However, similarities are expected regarding the sequence of stages and iterations that modellers follow during the model building process.

Table 3-1: Hypotheses 1.1 – 1.3 on the comparison of the DES and SD model building process.

Hypothesis 1.1:	DES and SD modellers pay varying attention to the different modelling stages. DES modellers spend most of their time in coding the model on the computer, whereas SD modellers spend most of their time in specifying the structure of the model.
Hypothesis 1.2:	DES and SD modellers follow a similar sequence of modelling stages.
Hypothesis 1.3:	DES and SD modellers follow similarly iterative modelling processes.

3.4.2 Hypotheses for the comparison of the DES & SD modellers' approach to model building (objective 2)

The hypotheses regarding the approach taken to model building are related to research objective 2, which aims to compare the approach taken by DES and SD modellers during the different stages of simulation model building. The assumptions made in these hypotheses (Table 3-2:) are based on the statements found in the comparison literature regarding the aspects involved in the model building process, summarised in table 2-3. Similarities between DES and SD are expected regarding problem structuring and model experimentation, while differences are expected for all other aspects of modelling.

Table 3-2: Hypotheses 2.1 – 2.18 on the comparison of the DES and SD model building approach.

Problem structuring	
Hypothesis 2.1:	Similar modelling objectives are pursued by DES and SD modellers, which are related to understanding how the system behaves over time under different scenarios. (Table 2-3 – line 1)
Hypothesis 2.2:	DES models problems at tactical/operational level, while SD at a strategic level. (Table 2-3 – line 2)

Conceptual modelling

Hypothesis 2.3: There is not a set diagramming method used in DES modelling versus standard diagramming methods (causal loop and stock & flow diagrams) used in SD. (Table 2-3 – line 3)

Hypothesis 2.4: Causal relationships and feedback structures are an essential part in SD models. In DES, while causal relationships and feedback can be included, they are not explicit. (Table 2-3 – line 4)

Hypothesis 2.5: DES takes an analytic view of systems, whereas SD a holistic view. (Table 2-3 – line 5)

Hypothesis 2.6: In DES modelling objects/people are represented as distinct individuals while in SD aggregate system behaviour is examined. (Table 2-3 – line 6)

Model coding

Hypothesis 2.7: DES modelling involves great complexity and detail whereas in SD models tend to be general and abstract representations of the system, taking a wider focus as opposed to a narrow focus. (Table 2-3 – line 7) *(Note this hypothesis considers detail and complexity from the model building point of view)*

Hypothesis 2.8: In DES modelling there are no prior modelling structures for modelling decision-making processes, whereas in SD standard modelling structures exist, such as asset-stock adjustment. However, information from observations and discussions with decision makers (practical structures) can be added in DES models. (Table 2-4 – line 8)

Hypothesis 2.9: DES models represent mainly material flows. Information flows can be incorporated but these are not obvious. In SD modelling both material and information flows are equally represented. (Table 2-3 – line 9)

Hypothesis 2.10: In DES modelling delays are represented in the form of delayed entities in the system, whereas in SD modelling delays are represented in the form of delayed influence or delayed processing of materials. (Table 2-3 – line 10)

Data inputs

Hypothesis 2.11: In DES quantitative data are used which are obtained from concrete/observable processes, whereas in SD modelling quantitative and qualitative data are used, where in the absence of data, opinions and best guesses can be incorporated. (Table 2-3 – line 11)

Hypothesis 2.12: Randomness is an essential part of DES models, whereas in SD modelling it is not as important. (Table 2-3 – line 12)

Hypothesis 2.13: In DES modelling linear relationships are more common. Non-linear relationships are more commonly used in SD modelling. (Table 2-3 – line 13)

Verification and validation

Hypothesis 2.14: DES modellers are concerned with developing an accurate model, close to reality, whereas SD modellers are concerned more with model usefulness. (Table 2-3 – line 14)

Hypothesis 2.15: DES modellers take a 'black box' approach while validating the model, checking mainly the model outputs, while SD modellers take a 'white box' approach, checking the internal structure of the model. (Table 2-3 – line 15)

Model results & experimentation

Hypothesis 2.16: DES modellers are interested in steady state conditions whereas SD modellers are interested in achieving a model in equilibrium. (Table 2-3 – line 16)

Hypothesis 2.17: The results of interest from a simulation model are quantitative point predictions and the optimisation of performance criteria for DES modellers, while SD modellers are interested in quantitative and qualitative results and point predictions are rarely made. (Table 2-3 – line 17)

Hypothesis 2.18: DES and SD modellers use scenarios to the same extent to compare the performance of alternative system configurations. (Table 2-3 – line 18)

3.4.3 Hypotheses for the comparison of DES and SD models use (objective 3)

The hypotheses on the use of DES and SD models are related to research objective 3, which aims to compare DES and SD models from the users' point of view. These hypotheses (Table 3-3) are specified based on the statements found in the comparison literature regarding the use and the learning achieved from DES and SD models and summary table 2.5. Next to each hypothesis, cross-references to the specific lines of the summary table (Table 2-5) are provided. It is expected that users find SD and DES models equally credible for giving answers to a problem situation and equally helpful as communication tools. However, differences in users' opinions about model understanding, model complexity, interpretation of model results and the models' role in learning and strategic thinking are expected to be revealed.

Table 3-3: Hypotheses 3.1 – 3.6 on the comparison of DES and SD modelling from the users' point of view

Hypothesis 3.1: Despite the use of animation, DES models are not transparent to the client, and so difficult to understand compared to SD models, which are transparent to the client. (lines 1 & 2)

Hypothesis 3.2: Different levels of complexity are involved in DES and SD modelling, resulting from the different levels of detail and perceptions of feedback. (lines 3 & 4) *(Note this hypothesis deals with detail and complexity from the model use point of view)*

Hypothesis 3.3: Users find DES and SD models equally credible for giving answers to a problem situation. (line 5)

Hypothesis 3.4: DES and SD models are equally helpful as learning and communication tools. (lines 6 & 8)

Hypothesis 3.5: SD models can aid strategic thinking to a higher extent. (line 7)

Hypothesis 3.6: The distinct nature of results derived from DES and SD models result in differences in learning & interpretation of results.

- DES aids instrumental learning, while SD conceptual learning. (line 9)
- The interpretation of DES model results is more difficult compared to the SD model. (line 10)
- Different aspects of the models are picked up by the users: Randomness is explicit in DES model results compared to the deterministic nature of SD models. (line 11)

3.5 Methodology

After discussing the scope and objectives of the current research, this section explains the methods used to study the comparison of DES and SD model building and model use. Furthermore, justifications are provided about the suitability of the methods chosen and a brief description of their implementation.

3.5.1 Study of model building

Objectives 1 and 2, deal with the comparison of DES and SD model building. In order to understand model building of DES and SD in practice an empirical approach is undertaken. This conveys that the aim of the research is to deal with 'naturally occurring data' (Silverman, 2000), which in the context of this study means the practice followed naturally by modellers during DES and SD modelling. Therefore, a qualitative approach is required. Potential methods considered for this purpose are: interviews with DES and SD modellers or the observation of real life simulation projects involving DES and SD modelling (ethnography) (Berg, 2004; Hair et al., 2007).

Interviews with experienced DES and SD modellers based on preset questions could be a valid method to identify the differences and similarities in model building, where modellers are asked to provide accounts of their experience of model building. However, given that the overall aim of this research is to get beyond opinions, and more specifically research objectives 1 and 2 refer to getting an empirical view about model building, interviews would not be an adequate method for the purposes of this research. The view taken in this study is that modellers' reflections may not reflect correctly the processes followed during model building and this would not represent a full picture of model building.

Ethnographic studies, on the other hand, involve the collection of data by systematically recording observations of people, events or objects (Hair et al., 2007). Observing real life projects can provide a valid view of the model building process and the approach taken, where the researcher attends project meetings and sits by the modellers recording their activities during the project. For a valid comparison it is necessary to have comparable modelling situations, which would require two potential real life modelling projects of equivalent problem situations. Early on during the project it was decided that this would not be feasible.

An appropriate method that mitigates the disadvantages of the two initial methods considered should provide unbiased accounts of the progression of a model building project, where the researcher can control to some extent the data collection process. Verbal Protocol Analysis (VPA) is considered to be one such method. VPA is a qualitative research method that requires the subjects to 'think aloud' when making decisions or judgements during a problem-solving exercise. It relies on the participants' generated verbal protocols in order to understand in detail the mechanisms and the internal structure of cognitive processes that take place (Ericsson and Simon, 1984). Therefore, VPA as a process tracing method provides access to the activities that occur between the onset of a stimulus (case study) and the eventual response to it (model building) (Ericsson and Simon, 1984; Todd and Benbasat, 1987). VPA was originally derived from psychology (Ericsson and Simon, 1984). Willemain (1994; 1995) was the first to use it in Operational Research (OR) to document the thought processes of OR experts while building generic OR models. Building on Willemain's initial work, Powell and Willemain (2007) and Willemain and Powell (2007) used VPA to study the model formulation processes followed by novice modellers in OR, with a view to gaining insights into the best way to teach OR.

It is important to explain why VPA is considered an appropriate method to use in the model building study. Let us imagine a two-axis plan (Figure 3-1), where on the horizontal axis is type of report, personal report (modellers' own reports) versus third person's report (researcher report on observed event). On the vertical axis is displayed the researcher's control on the data collected. The control level varies from minimum control (the researcher observes a real life modelling project, but cannot control the settings or the problem situation modelled and cannot have a complete view of modellers' views or decisions) to maximum control over the data collection process (the researcher determines the data collected and asks relevant questions). Therefore, in this two-axis plan, observation of real life modelling projects and interviews stand diagonally in opposite directions. Around the mid-point, close to the origin position stands Verbal Protocol Analysis, where impartial accounts about aspects of modelling are reported spontaneously by modellers, who undertake a specific modelling exercise. Meanwhile, the researcher controls the collection process to some extent, by choosing the problem situation to be modelled. However, during the modelling exercise the researcher's interaction with the modellers is limited in order to maintain the objectivity of the protocols derived. The researcher does not intervene unless the participants stop talking or require further information. This suggests that VPA makes up for the disadvantages of the originally considered research methods, observation (ethnography) and interviews.

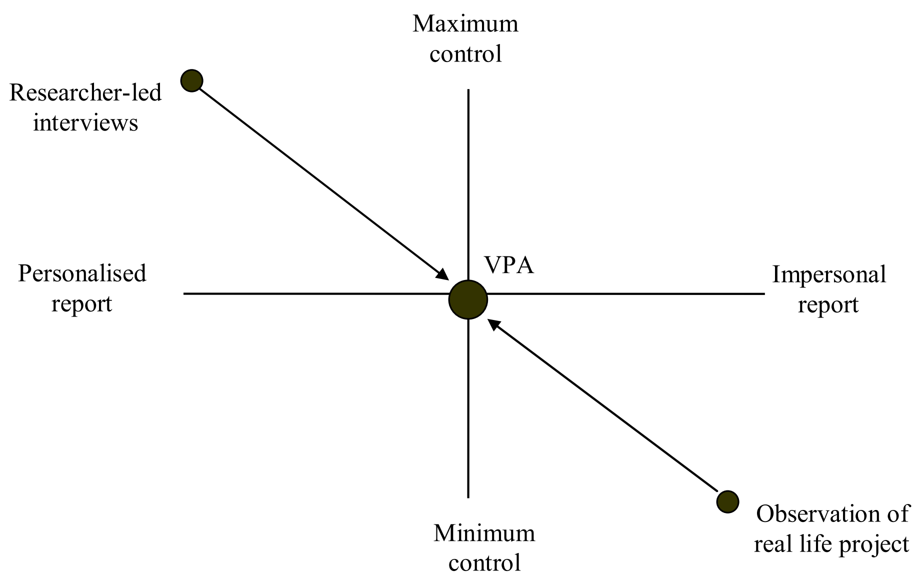


Figure 3-1: Configuration of methods by level of personalisation of reports and researcher control on data collection process.

Limitations of VPA

VPA as an approach has its own limitations. First of all, the artificial setup of the experiments involves some potential disadvantages. The experts are asked to sit alone in the room while working on a modelling exercise for a limited period of time, e.g. one hour. This might not represent the normal environment where the expert works in practice, interacting with the client, colleagues, etc. Furthermore, the absence of a real client, adds more to the artificiality of the experiment. However, to offset this, the subjects solve the exercises in the presence of the researcher. Indeed, at times participants are allowed to ask questions or require clarifications about the scope of the exercise.

In addition, due to the nature of the method, verbal reports may not be able to capture some important data (Willemain, 1995). There are a few factors that can affect the quality of the protocols. An issue often raised in the VPA literature is the fact that experts due to their experience and skill assets, tend to verbalise less compared to novices (Ericsson and Simon, 1984). Moreover, the experts, who might be feeling that they are under observation, may not behave similarly to normal conditions. Being asked to verbalise more than normal, experts' thought processes might be distorted. There might also be a discrepancy between the experts' actual and verbalised thoughts, where instead of thinking aloud they might be only explaining. To overcome this and to ensure that the participating modellers think aloud, short verbalisation exercises, based on Ericsson and Simon (1984) are run at the beginning of each modelling session. Also, a variation in the experts' level of experience (years) and background is targeted (chapter 6).

VPA also involves the researcher's interpretation of the accounts made. The generated verbal protocols are coded and analysed by the researcher, based on their views, beliefs and assumptions, and, therefore, some subjectivity is involved. In this thesis, subjectivity is dealt with by undertaking multiple independent codings and/or involving a third party in the coding process (section 5.5.2).

For the purposes of this study, VPA is considered to be an effective method for the comparison of the DES and SD model building process. It is useful because of the richness of information and the live accounts it provides on the subjects' behaviour during the modelling process. Hence, it is believed that it is possible to gain reliable data from VPA sessions for the purposes of objectives 1 and 2 of this study. However, the protocols obtained from this study will not necessarily reflect a full picture of the two fields of simulation modelling and generalisations will be avoided. Therefore, the limitations must be borne in mind when drawing conclusions from the results of the VPA.

Implementation of the model building study

In the context of the model building study, VPA is used to understand DES and SD modellers' thought processes while undertaking a model building exercise. One-to-one modelling sessions are run in the form of experiments, where DES and SD modellers are provided with the same problem situation - case study (chapter 4) – and asked to build a model based on it. In addition, modellers are asked to think aloud throughout the modelling session. Their accounts while modelling are recorded and transcribed.

With respect to objectives 1 and 2, content analysis of the verbal protocols is undertaken. Content analysis involves analysis of the content or message of written text (Hair et al., 2007). The end result can take the form of quantified qualitative data or interpretation of the text.

With respect to objective 1, which entails the comparison of the modelling process followed by DES and SD modellers, a quantitative analysis of the participants' verbalisations is carried out. This involves dividing the verbal protocols into episodes, which are then coded (given a name based on a coding scheme), according to the modelling stage the modeller's thoughts are most relevant to. The coding scheme and how the coding is done in practice will be explained in more detail in Chapter 5 and the results in chapter 6.

With respect to the research objective 2, qualitative content analysis of the verbal protocols is implemented. For this part of the analysis the coded protocols in the quantitative part of the analysis are re-used, but in this case to give an interpretation of the aspects verbalised by DES and SD modellers on specific themes based on hypotheses 2.1 to 2.18 (section 3.4.2). Text analysis is considered the best means of achieving the second research objective. In this way, inferences can be made about the approach taken by the two groups of modellers as per the statements found in the literature. The design of the model building study will be further discussed in more detail in chapter 5 and the results in chapter 7.

3.5.2 Study of model use

Objective 3 is concerned with assessing how different model users find DES and SD models as decision making tools. More specifically, the aim is to empirically evaluate the statements found in the literature about how different the two simulation techniques are from the users' point of view. Therefore, a survey research is considered appropriate, where participants are asked to provide their opinions about two equivalent DES and SD models. The focus is on collecting quantifiable and standardised data, which can be directly compared for two equivalent simulation models. Therefore, purely qualitative methods, such as participant observation and focus groups, which give emphasis on the contextual understanding of social behaviour (Bryman, 2008), are not considered suitable. Two alternative survey methods are considered, structured interviews with model users and self-completion questionnaire surveys (Buckingham and Saunders, 2004; Bryman, 2008).

Individual interviews involve asking participants specific (structured, semi- or un-structured) questions and receiving detailed answers. Interviews provide an opportunity for detailed investigation of people's personal perspectives about two simulation models (Ritchie, et al., 2003). However, the nature of this study is not so much to gain an individual's perspective than to study users' perceptions regarding two equivalent DES and SD simulation models. Furthermore, as it will be explained in chapter 9, ready access to managers, Executive MBA students attending the module Modelling and Analysis for Management (MAM) was available. Managers, as ultimate users of simulation models are considered a suitable sample. Due to the time constraints involved with the MBA study curriculum, running interviews with each MBA student was not feasible.

On the other hand, a self-completion questionnaire survey would suffice to collect responses with respect to a set of specific questions based on the statements found in the literature. A questionnaire survey is a widely used method in social science research, involving a list of questions, which serve as an instrument for the measurement of data regarding attitudes, opinions, etc (Oppenheim, 1968; Hair et al., 2007; Bryman, 2008). Questionnaire survey ensures some level of standardisation in the research, where all the participants are asked the same questions in the same order. Furthermore, respondents answer at their convenience and the interviewer effect is to some extent eliminated. Social desirability bias, as the tendency of people to provide more socially acceptable answers when the interviewer is present, is also avoided (Bryman, 2008). Hence, a questionnaire survey is considered an appropriate method for the purposes of this study.

Limitations of self-completion questionnaire survey

There are obviously disadvantages in relation to running a questionnaire survey. Most importantly, participants cannot be prompted or probed. For the former (i.e. prompting), there is no one present to help respondents with questions which are difficult to understand and answer. On the other hand, probing is mostly related to open-ended questions, where normally the interviewer would probe with the respondents to elaborate on an answer. The length of a questionnaire is limited and with it the amount of data collected, to avoid ‘respondent fatigue’ (Bryman, 2008). Furthermore, in the absence of an interviewer, salient questions need to be limited, in order to avoid low response rates. The questionnaire needs to be short and the questions simple so that the respondents do not find particular difficulties in answering the questions.

Implementation of the model use study

The questionnaire survey, as part of the model use study, consists of experimental model use sessions. These include group work, problem solving, use of simulation models and checking the results as well as group discussions. Participants are provided with a case study situation (chapter 4) and an equivalent DES or SD model. After working with a simulation model as part of a problem solving exercise, they are asked to record their answers to a list of questions. This part of the study considers the use of models and most importantly testing the statements found in the existing comparison literature based on table 2-5 (chapter 2). The hypotheses that the model use study aims to confirm or otherwise, are 3.1 to 3.6 (Table 3-3). The experimental factor is the simulation model used, where one group uses a DES model and the other an equivalent SD model (chapter 9). Participants’ answers are compared to evaluate the differences in their opinions. The results of the statistical analysis are provided in chapter 10.

3.6 Summary

Drawing upon the literature review discussed in chapter 2, the current chapter presents the overall aim of the study undertaken, which is to compare the DES and SD modelling approach from the model building and model use perspective. Furthermore, three core objectives are set out and the underlying hypotheses articulated based on the existing literature. Then, the chapter describes the methodology and the research methods chosen in order to implement the study. It is concluded that VPA and a questionnaire survey are the most appropriate methods for the comparison of DES and SD model building and model use respectively.

Chapter 4: A prison population case study

4.1 Introduction

The current chapter provides an introduction to the case used for the implementation of the empirical work. First, the choice of the UK prison population case is discussed. Then the case materials used for the two studies (that of model building and model use) are introduced, followed by a justification of why the prison population case study is appropriate for the research undertaken in this thesis.

4.2 Criteria for the choice of the case study

For the implementation of the empirical work consisting of the model building and model use studies (sections 3.5.1 and 3.5.2 respectively), a case on the UK prison population has been chosen. In the model building sessions, involving Verbal Protocol Analysis (VPA), the case study is used as the stimulus for model building. Participants verbalise their thoughts as part of this process. In the second study the survey of model use, DES and SD simulation models are built based on the same case, which are then given to the participants to use. The case study provides model users with background information regarding the problem considered, while the simulation models are the control variable, each given to two different groups of users, who are asked to use the models and make management decisions. While in the first study the case is used as the stimulus for modellers to build a simulation model, in the second study it provides the necessary background information about the simulation models, which then participants use for decision making.

The same case is used for both studies implemented in the thesis. Due to the different aspects compared by the two studies, the integrity of the study is not considered to be affected by the use of the same case. The main criteria considered for the choice of an appropriate case are: simplicity, novelty and the potential to accommodate both DES and SD modelling. For the purposes of the model building study, a suitable case needs to be simple enough to enable the development of a simple model in the limited period of time available. On the other hand, for the model use study, using simple models ensures that users with little or no prior experience of using simulation models can understand and use them for decision-making. From a practical point of view, limited time availability was an issue for the sessions of each study. A time limit of 60 - 90 minutes was set for the modelling sessions involving busy expert modellers, to ensure their participation in the study. Whereas, the model use sessions needed to fit with the MBA course curriculum and requirements for a 1.5 hour long session.

Regarding novelty, the case used in the model building study needs to be equally familiar to both SD and DES modellers, to assure a common starting point for all modellers. For example, the beer game, which considers the bullwhip effect in a supply chain, has been widely modelled in SD. In this case, SD modellers would be far more familiar with building such a model as compared to DES modellers. Using a novel case situation ties-up well with the VPA methodology, ensuring that the participants, without any prior preparations, are actively involved in the modelling exercise and thus providing live accounts of the model building process. With consideration to the survey of model use, the choice of a novel case study and its implementation is appropriate with the view to collecting participants' unbiased opinions on the simulation models used.

Another important criterion is that the case needs to accommodate models from both simulation techniques, but on the other hand it should not lead to the choice of one or another approach. For example, some specific features of each modelling approach the presence of which is required are: randomness in DES vs. deterministic models in SD, the aggregated presentation of entities in SD vs. the individual representation of entities in DES, etc. This is further discussed in section 4.4 regarding the suitability of the case chosen, that of the UK prison population.

4.2.1 Selection of the case

This section describes the process followed for the choice of the case. First, two previously considered problems are referred to, followed by the reasons why these were rejected. Then the selection of the prison population case is described.

Supply chain management was originally chosen as a specific area of application for the empirical work. The supply chain is a topical area and moreover has been extensively modelled in both the DES and SD literature. It was initially expected that a suitable case would be available among the already published case studies on supply chain management. For this reason, a number of case study databases were explored, such as: European Case Clearing House (ecch), Darden case collection, Richard Ivey and the Harvard Business School Cases. The case studies retrieved from these databases were considered to be general in scope, dealing with multiple issues and/or complicated issues which were not suitable for building simple simulation models.

Due to the difficulties encountered in finding a suitable pre-existing case study, the creation of a customised one, tailored to the needs of the research, was considered. An application in the supply chain sector was first attempted and a case of an imaginary doorstep dairy delivery company (Fresh Dairies) was chosen (Appendix A.1). The case study design process was an iterative process revolving between writing the case study and building the equivalent supply chain model based on it. The DES model was built in Simul8 and was found quite complicated, taking long hours of modelling. The complexities arose due to detailed information required for each echelon, i.e. demand, consumption, order processing, etc. Furthermore, as the design of the case progressed, it resembled a generic supply chain, similar to the beer game. As a result, the case of the milk delivery business was discontinued.

In the meantime, the fishery model developed by Morecroft and Robinson (2005) was considered. Their model could serve the basis of a simple model, representing the fish stock in the sea, including the feedback effect of natural fish regeneration and the effect of fishing undertaken by human activity. The DES fishery model created attempts to match the structures and modelling aspects of the problem situation chosen and it is not deemed to be a representative DES model for the purposes of this study because it does not represent individuals. Hence the fishery model was not considered suitable for this study. Some models with a similar structure to the fishery model are generally those that incorporate birth/death processes with inventory, such as: hospital occupancy, prison population, customers of a particular brand/product, workforce recruitment/retention, etc.

After thoughtful consideration, a case study on the UK prison population based on Grove et al. (1998) was chosen. The prison population is a topical subject both in the UK and elsewhere, e.g. (Korporaal et al., 2000). The inherent feedback that exists in the system, with prisoners entering and returning back to prison due to re-offending (recidivism), can be uniquely represented by each simulation approach, DES and SD. While simulation models of the prison population and the criminal justice sector are rare, DES and SD have both been used to model the prison population. DES models of the prison population have been developed by Kwak et al. (1984), Cox et al. (1978), Korporaal et al. (2000). Equivalent SD models of prison population have been developed by Bard (1978) and McKelvie, et al (2007), while the initial UK prison population model by Grove et al. (1998) is a flow model analogous to an SD model. Hence, the prison population case is deemed as an equally familiar application to both DES and SD modellers and also amenable to both modelling approaches. Therefore, the UK prison population case is considered suitable for use in comparing DES and SD model building and model use in the current research.

4.3 The case and the problem situation

In this section the case developed is explained. The UK prison population case is based on the problem of overcrowded prisons, concerning government authorities. The figures and facts used in the case are based on reality, but slightly adapted for the purposes of the research. The case material used for the model building and model use study differs slightly depending on the requirements of each study. However, the basic facts are similar for both studies.

The main characteristic about the behaviour of the prison system is the inherent stability in the prison population numbers. The graph of the prison population size (Figure 4-1), which was also included in the case description, shows a constant prison population in the period between the early 80s until 1993. The increase in prison population levels since 1993 is explained as a result of the introduction of tougher sentences, where the number of offenders sentenced to custody and sentence length, have increased by 83% and 31% respectively.

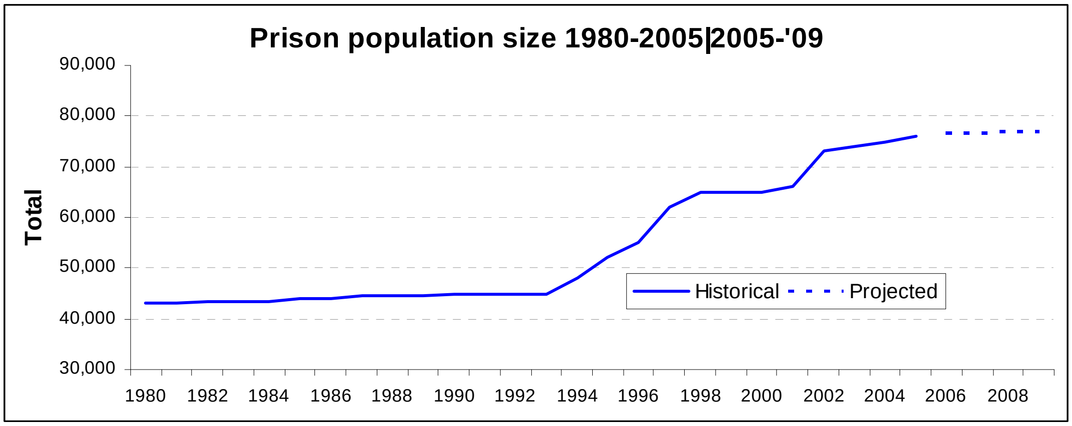


Figure 4-1: Graphical representation of the historical & projected UK prison population size.

The case draws particular attention to the issue of prison overcrowding. The problem is presented as a typical public resource allocation problem with an objective to improve the capacity of the criminal justice system in preventing crime and deterring its repetition. Two alternative scenarios considered by government authorities are suggested, either to increase the current prison capacity along with the introduction of stiffer rules, or to reduce the size of the prison population by introducing alternatives to jail and/or enhancing the social support provided to prisoners.

More specifically, the case considers two types of offenders, petty and serious. At the beginning of 2005, there are 76,000 prisoners in the system, where 50,000 are petty and 26,000 serious offenders. Offenders enter the system as first time offenders and receive a sentence depending on the type of offence. Petty offenders enter the system at a higher rate, due to a higher rate of offending (on average 3,000 people/year vs. 650 people/year for serious offenders), but receive a shorter sentence length (on average 5 years vs. 20 years for serious offenders). After serving time in prison the offenders are released. A proportion of the released prisoners re-offend and go back to jail (recidivists) after on average 2 years. Petty prisoners are more likely to re-offend; 70% re-commit petty crimes and go back to jail and another 3% commit even more serious crimes and are re-convicted as serious offenders. Serious offenders represent a small percentage of the total offender population and have lower rates of recidivism. Only 30% of serious offenders re-offend and go back to jail as serious offenders after 2 years.

The system presented in the current case is a rather simple view of the criminal justice system. Obviously, additional factors that affect the system performance and also more complex relationships can be identified such as the social effects on the number of crimes committed or the number of deaths in prison, etc. However, for the purposes of keeping the case and the associated models simple, these factors were left out of the conceptual model provided, focusing mainly on the key aspects of the problem.

A brief explanation of the case material used for each part of the survey follows in the next sections (4.3.1 and 4.3.2).

4.3.1 The case used in the model building study

The case used in the model building study is provided in appendix A.2. It starts with a brief introduction to the prison population problem, including the main reasons for, and the impacts of, the problem and the two possible scenarios. The task for participating modellers is to create a simulation model to be used as a decision-making tool by policy makers/government officials.

In order to give a common starting basis for all modellers, a simple graphical representation of the model (Figure 4-2) is provided in the case material, as well as a set of data inputs/assumptions (Table 4-1) to be used at the modellers' discretion. The graphical representation is a simple influence diagram, which represents the flow of prisoners, where the arrows are straight lines to avoid any direct association with an SD model. Some data inputs, such as the percentage of petty and serious re-offending were purposely left out of the case study, so that modellers are not lead to a specific model representation. The researcher was prepared to provide these data during the modelling sessions, if the modellers requested them and in the form they asked for.

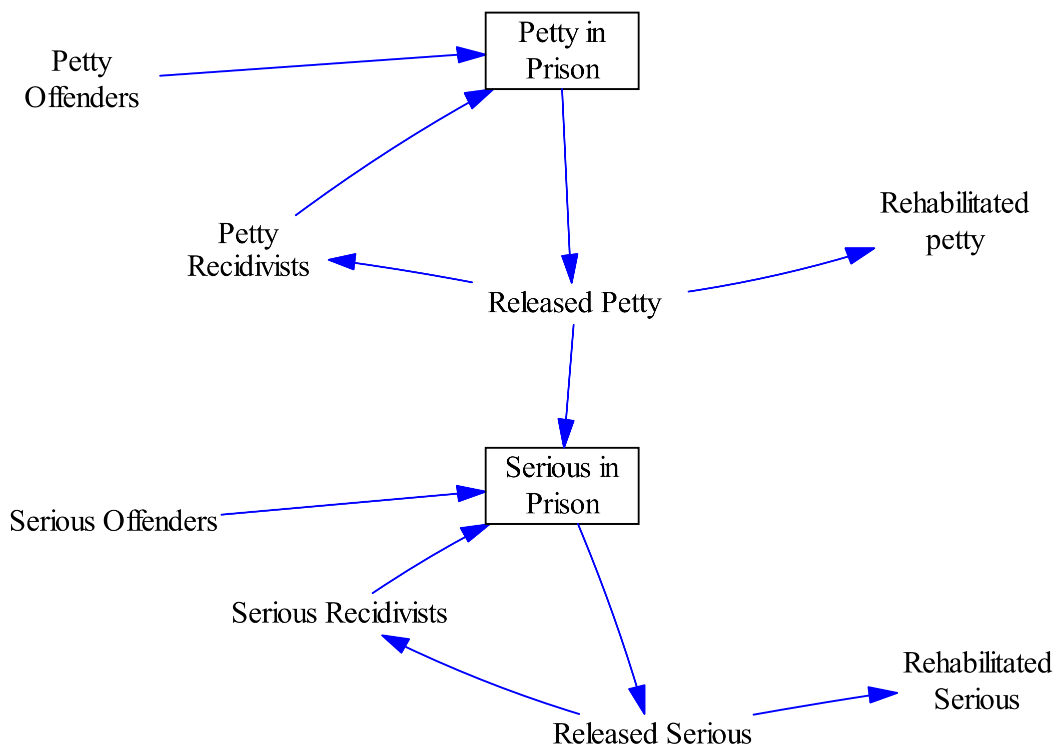


Figure 4-2: Graphical representation of prison system for the model building sessions.

Table 4-1: Initial assumptions/numbers suggested in the case study for model building sessions.

	Petty offenders	**Serious offenders**
Current prison population	50,000 prisoners	26,000 prisoners
Prison admissions	3,000 prisoners/year	650 prisoners/year
Average sentence length	5 years	20 years
Time to recidivist return to jail	2 years	2 years

4.3.2 The case used in the model use study

The same case material is used for the model use study (Appendix A.3), with some differences in presentation. The two possible scenarios available to policy makers are described in more detail, where each includes a list of possible policies that can be implemented in order to solve the existing problem. A more detailed graphical presentation, including all the data inputs used, is also provided in order to help participants' understanding of the model structure (Figure 4-3).

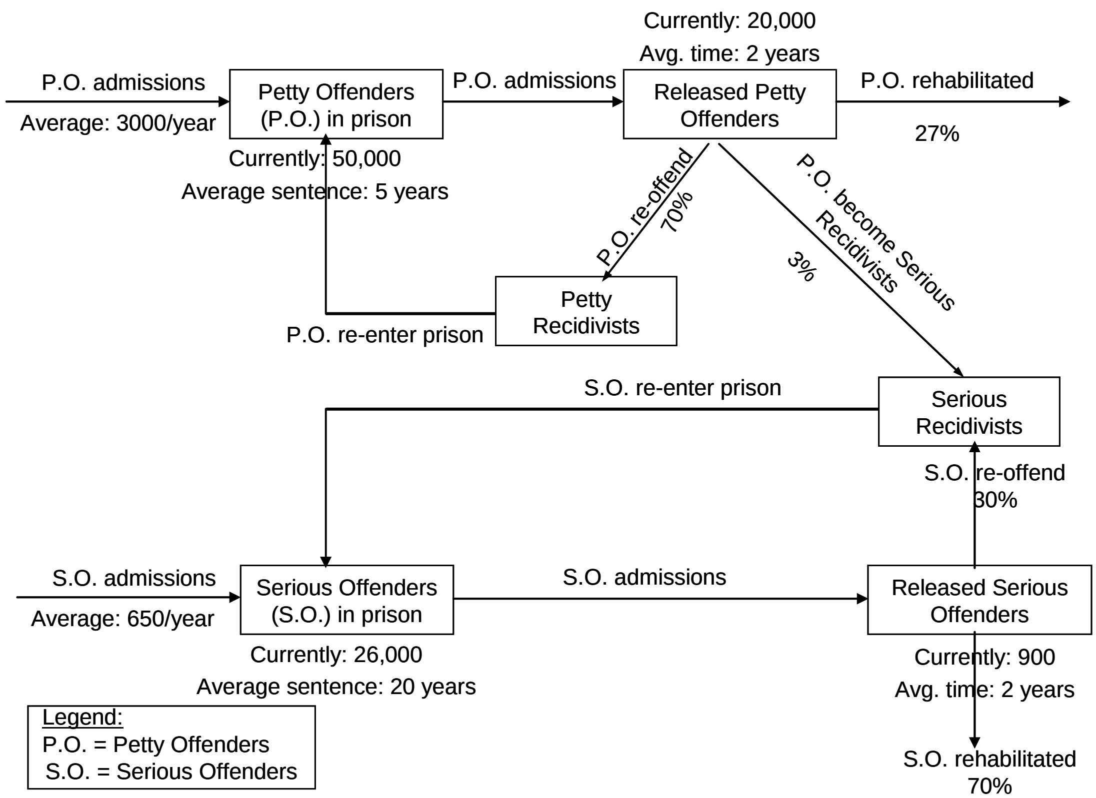


Figure 4-3: Graphical representation of the flow of prisoners in the model use case.

For the choice of the specific policies for each scenario, research was carried out in order to understand the different policies that have been implemented from time to time in different countries around the world. The policies chosen were purposely linked to the experimental factors. Based on their own discretion, the survey participants using the simulation models developed were expected to suggest policy changes after experimenting with different sets of experimental factors. More particularly, the task requires the participants to take the role of a government consulting service and to suggest possible policy changes that would help solve the problem of prison overcrowding.

Budget limitations are also specified for the two possible scenarios with respect to the increase or decrease of prison capacity. Therefore, the task consists of applying relevant policy changes for each of the two scenarios in order to suggest what the best course of action would be for government authorities. The experimental factors with which one can experiment are:

- the number of admissions for petty and serious offenders per year,
- the sentence length for petty and serious offenders (in years),
- the percentage of petty and serious offenders who re-offend and
- the percentage of petty offenders who commit a more serious offence.

There are no right or wrong answers to the problem. The main purpose of the case is to present a multi-faceted issue, which can induce discussions and give the users the liberty to choose their own course of action while using the relevant simulation model.

4.4 Considerations about the suitability of the UK prison population case to the DES and SD modelling approach

The same case with offenders entering the system and flowing throughout the system as prisoners or recidivists can be modelled by both simulation approaches, even though a different approach maybe taken to model the flow of prisoners throughout the model. Considering the existing and generally accepted differences between the two simulation approaches, it is important to point out the reasons why the case is equally suitable for both DES and SD modelling.

Due to the feedback effects inherent in the closed path of prisoners, who come in the system and return back to jail through the recidivism route, the prison system is considered appropriate for both DES and SD modelling. As a matter of fact causal relationships and feedback structures are an essential part of SD models. From a DES point of view, this empirical work aims to reveal how the feedback effects are represented in DES models and thus identify issues related to it.

Individual entities (i.e. prisoners) in SD are aggregated and indistinguishable, while in DES individual prisoners can be modelled and specific attributes can be assigned to them (i.e. we can track how many times a specific prisoner has offended). The large number of people involved in the prison population case study might pose difficulties for the DES approach, but not for SD, which is naturally suited to dealing with large populations.

Regarding time advance, in DES the system is modelled as a series of events, where the simulation clock advances at specific instants of time when a state-change occurs (Pidd, 2004; Robinson, 2004). In SD time advances in equal time steps, Δt, and change progresses through time based on differential equations that describe the relationships in the system. Regardless of the two different time handling approaches, the same system can be modelled, representing either discrete state changes e.g. a prisoner moves from one state to another (DES), or continuous state changes, e.g. annual change in prison population (SD).

Regarding randomness, it is of interest to identify the attitude of modellers and users towards it. While in the case description average data inputs are provided (i.e. average sentence length, prison admissions, etc.), these are expected to be handled differently in DES and SD modelling. In DES random aspects of the problem situation (e.g. the time people spend in prison, which differs from prisoner to prisoner and from prison to prison) are normally represented with the use of statistical distributions. SD models are mainly deterministic and aggregate values are used (i.e. the average time that people spend in prison). Two separate prison models including random processes or deterministic data can be built by using each specific simulation approach.

Given these points and that both approaches have previously been used for modelling the prison population system, the case chosen is considered suitable for both DES and SD modelling. Due to the different ways of modelling different aspects of the prison system, the implementation of the research will reveal the differences and similarities in the model representations of the case created by expert modellers and those created by the researcher.

4.5 Summary

The current chapter discusses the UK prison population case used in the thesis. More specifically, an explanation regarding the choice of the case study is provided followed by the process followed for its design. Then, the customised case studies designed for the two parts of the empirical work are explained. Last but not least, considerations are made about the suitability of modelling the UK prison population case study using the two simulation approaches (DES and SD). We now move on to discuss the design of the model building study.

Chapter 5: Research design I – study of model building

5.1 Introduction

Having initially decided on a qualitative approach and the use of VPA to study DES and SD model building (chapter 3), the next step is the design of the settings of the model building study. The design of the study involves a sequence of iterative processes that evolved as the project progressed.

This chapter describes the main activities undertaken in designing the model building study. These include aspects such as the pilot study, the participant access, the design of the experimental modelling sessions and data analysis.

5.2 Pilot model building study[6]

Before implementing the main modelling sessions a pilot study was run in order to reveal potential problems and assess the suitability of the study design. More specifically, during the pilot study the case study and verbalisation exercises were tested. In addition, the researcher, who was implementing the 'think aloud'

[6] This section is based on the conference paper Tako AA and Robinson S (2008) "Comparison of Model Building in Discrete-Event Simulation and System Dynamics". In the proceedings of the 2008 Simulation Conference (SW08), 1-2 April 2008, pp. 209-218, Worcestershire, UK.

modelling sessions for the first time, gained valuable insights about VPA. Therefore, the pilot study served as good practice for the main model building sessions. Furthermore, during the pilot study the coding scheme was tested and changed a few times. A first attempt to analyse the data gained was made. The initial findings served as a good starting point for the analysis of the real data.

Four modellers were involved in the pilot study, including 2 from a DES background (P1 and P2) and 2 from a SD background (P3 and P4). All 4 participants use simulation modelling (DES or SD) as part of their work and have at least 2-3 years' modelling experience. Their profile was more academic-related. 2 were lecturers in Operational Research, 1 a doctoral student in simulation at Warwick Business School and 1 Senior Research Fellow at the University of Warwick.

The modelling sessions were run at the University of Warwick and the relevant procedures (section 5.4) were followed. Straight after the sessions, the recordings were transcribed and the resulting protocols were coded based on the coding scheme developed at the time. A qualitative analysis was carried out.

During the pilot modelling sessions a number of observations were made. The four participants were also asked to provide opinions and suggestions for the improvement of the exercise as well as the process followed. The main learning points from running the preliminary modelling sessions consisted of the following:

- It was considered necessary for the researcher to intervene during the modelling session to remind the modellers to keep talking, in the case that they remained silent for more than 30 seconds. Especially during the time that modellers build the model on the computer, they tend to concentrate on the technical side and long pauses can result if the researcher does not intervene.
- Being aware of the main requirements of the study for an impartial comparison of DES and SD modelling, it was found important to ensure that the case study material is unbiased towards the DES or the SD approach. This was found to be a sensitive issue because DES and SD modellers need different types of information when modelling. It was, therefore, essential that the information provided in the case study material is generic and that it does not lead the modellers towards DES or SD modelling. Hence it was decided that modellers are provided only with some general data and the researcher asks the modellers to enquire for any additional information during the modelling sessions.
- In addition, it was observed that modellers did not attempt to build a model on the computer when asked to build a simulation model; they tended to work only on the conceptual model. Therefore, it was found that the modellers need to be prompted in order to build a model using the relevant simulation software on the computer.

Another aspect that was systematically checked and revised during the pilot study was the coding scheme. In the first steps, the initial coding scheme was primarily based on the statements found in the comparison literature. However, after consulting with the respective DES and SD literature, the codes were further refined to the final format as displayed in section 5.5.1. A number of initially considered codes were discontinued. For example, a code on modelling practice, considering modellers' thoughts about their modelling preferences, was considered ill-defined and also not relevant to the objectives of the study. It did not represent a specific stage in the modelling process. Another code on model use and model portability in the future was also discontinued, because this was not deemed relevant to the model building process. It resembled issues regarding the use of the developed model, which was not expected to occur during model development. In addition, some other codes were re-defined and re-named. Such was the case with a code on 'model elements', which was later re-named conceptual modelling. Its definition was fine tuned to include only the topics: diagrams, model level, feedback and people. Some other aspects, originally located under model elements, randomness and relationships between variables, were considered more relevant and thus moved to the specifically defined code, that of 'data inputs'.

Furthermore, from the preliminary analysis of the 4 protocols as part of the pilot study it was concluded that DES and SD modelling are quite different from the perspective of the model building process. Hence, the preliminary findings were a good starting point for the analysis of the 10 protocols in the main study.

A summary of the findings from the pilot study follows:

- From the preliminary analysis, it was observed that SD modellers reflected more on the first stages of the modelling process including: problem structuring, conceptual modelling and data inputs. Meanwhile the DES modellers spent more time with model coding and the verification of the model built.
- Overall, the SD modellers took a wider view of the problem situation and looked into other factors that affect the problem, while the DES modellers considered as adequate the definition of the problem already given in the case study.
- Contrary to expectations, both DES and SD modellers rely heavily on data inputs. However, the differences lie in their conception of data inputs. It was difficult to satisfy both groups of modellers with the same type of data. While DES modellers were happy to receive suggestions about some preliminary data, SD modellers were not prepared to accept data that were not supported from research. SD modellers emphasised the need for customised research based on the variables created.
- Differences were also found regarding the verification approach taken. While SD modellers were more concerned with building the right structure of the model, DES modellers focused more on coding correctly and ensuring that the model was working as intended.

Overall, the pilot study provided useful insights in pointing out the strategies for the successful implementation of the main model building study and it helped in choosing the best practice for analysing the verbal data.

5.3 The participants

With respect to the objectives of the model building study, expert modellers from the two fields of simulation modelling (DES and SD) are required. Two types of experts can be identified: academics and practitioners. For the purposes of this study the latter are preferred as practitioners' modelling experiences involve mostly the practical aspect of modelling rather than the theory behind it. In addition, practitioners are deemed a 'cleaner' sample because they are less exposed to academic discussions on the comparison of DES and SD. Therefore, in this research, expert modellers are considered practitioners, as opposed to academics, who use simulation modelling (DES or SD) as part of their work and have at least 4 years of experience in modelling. For accessibility purposes, the sample of the study is confined to expert modellers who live and work in various geographic regions within the United Kingdom.

5.3.1 Participant access

As a starting point, in the process of identifying participants relevant to this study the researcher received useful information and guidance from her supervisor, Professor Stewart Robinson, who is an expert in the DES area and has links with a number of software and consultancy companies in the DES field. Furthermore, the researcher actively attended conferences (the OR Society Simulation Conference in Leamington Spa, March 2006 and International Conference of System Dynamics in Nijmegen, July 2006) and DES or SD-themed events (Operational Research Society Joint Criminal Justice and Simulation Special Interest Group meeting in Harrogate, UK Chapter of System Dynamics Social Gathering in London, February 2007) with the aim to pinpoint potential contacts. Dr. Susan Howick, from the University of Strathclyde, who has also served as president of the UK Chapter of the System Dynamics Society, kindly provided useful contacts with SD modellers too.

After establishing a first contact point, e-mail and telephone contact was maintained between the researcher and potential participants. Even though e-mail contact was preferred in most cases, telephone calls were found more effective in setting an appointment, especially with busy consultants.

Approaching individual modellers, working as part of an organisation can potentially raise a number of issues and considerations (Ritchie and Lewis, 2003). Therefore, patience, respect and sensitivity are considered essential. Indeed a number of issues were encountered during the process of finding potential participants. A summary of the main strategies followed during this process follows:

- **Responsiveness to participant concerns:** The researcher was well prepared and anticipated that questions and concerns would be raised by potential participants. These were tackled with respect towards the individuals, but on the other hand taking into consideration the research objectives and requirements. Attempts to convince modellers to take part in the study were not always fruitful. Some negative replies were received asking for further information. During this process, a judgement call was made as to when to further attempt to convince a modeller to participate in the study or to drop interest. At times it was considered necessary to respect these specific individuals' decisions and to discontinue the approach.
- **Clarity of information.** When communicating with potential participants or dealing with their questions, the researcher was aware that clear information regarding the objectives and purpose of the study should be provided. However, the study objectives were only partially explained by saying that the aim was to study the simulation modelling process. The real objective of the study, that is, to compare DES and SD modelling and the case material were not revealed to the participants prior to the modelling sessions. This is in line with the requirements of the study to avoid biasing the modellers and to ensure that all modelling sessions would have the same starting point.
- **Negotiation.** In cases a need to negotiate with participants or their companies was deemed necessary. In the cases when members of the higher organisational hierarchy were involved during the communication, it was agreed that the findings of the research would be shared with them at later stages of the research.

In the cases, when the modelling sessions were arranged with modellers directly, the modellers were given small 'thank you' treats (e.g. a bottle of wine or a box of chocolate), after the completion of the modelling session.

- **Flexibility.** Flexibility in response to participants' circumstances was considered essential. The researcher travelled to different destinations such as: Glasgow, London, Leeds, etc. to meet up with the modellers. Suggestions were received by some participants to run virtual modelling sessions via web-based conferencing, but for reliability issues and due to limited accessibility of conferencing equipment, this was ruled out in the early stages of the research.

After completing the modelling sessions a good relationship continued to be maintained with all the participants. There have been cases when the researcher has contacted the participants with clarifying questions and these have been promptly answered.

5.3.2 The sample size

As a result of the process undertaken to finding potential participants, 10 simulation experts (5 in DES and 5 in SD) were chosen to participate in the modelling sessions. It should be noted that the DES sample included 8 modellers (3 from the same organisation), but only 5 were selected[7] to ensure an equal sample size from the two groups of modellers and that only one participant was included from any specific

[7] Only one out of the 3 modellers from the same organisation was randomly chosen to be included in the sample.

organisation. The sample size of 10 participants is considered reasonable. Due to the richness of data found in one protocol, VPA samples tend to be small, between 2 to 20 (Todd and Benbasat, 1987). Therefore, considering project timescales, the sample size was limited to 5 participants for each group. The sampling method used is a mixture of convenience and purposive sampling (Silverman, 2000; Ritchie et al., 2003). Some level of diversity among the participants and their profiles was sought, regarding the length of modelling experience, the company they work for and type of simulation software mostly used.

5.4 Model building sessions – Data collection

The modelling sessions were held in an office environment, on a one to one basis. All the sessions were administered by the researcher, who was present throughout the modelling sessions. The subjects involved had access to writing paper and a computer with relevant simulation software (e.g. Simul8, Vensim, Witness, Powersim, iThink/Stella, etc.).

The participants were first provided with the prison population case (Appendix A.2) and were asked to read it through and ask any clarifying questions. Then the researcher read a set of standardised instructions (Appendix B.1), explaining to the subjects the aims of the exercise. In the initial instructions, it was particularly pointed out to the modellers that they were expected to speak their thoughts aloud as if they were talking to themselves and were alone in the room while they are modelling. Furthermore, before starting the modelling exercise, some short warm-up exercises (Appendix B.2) were run with the participants, to aid their verbalisation process and to acquaint them with the experimental situation. 3 out of the 4 warm-up exercises were picked at random for each participating modeller. The modelling exercise asked the subjects to build a model based on the UK prison population case study (chapter 4). The objective of the exercise was to build a simulation model which will be potentially used as a tool for decision-making by government authorities. Starting from common initial settings, the modellers were then left to proceed with modelling whilst verbalising their thoughts.

The researcher sat in the same room, but social interaction with the subjects was limited, to avoid having any effect on the participants' thinking process. She only intervened to encourage subjects to think aloud and to also prompt them to 'keep talking' in the case that participants stopped talking for more than 30 seconds. The researcher was also answering explanatory questions and provided participants with additional data inputs (if they asked for them). She also prompted the subjects to build a model on the computer in case when they did not do so on their own initiative. The sessions lasted approximately 1-1.5 hours and were recorded using a digital voice recorder. These were later transcribed.

The verbalisations made by the 10 expert modellers, including the time length and the number of words verbalised are considered. The pink dots in the graph (Figure 5-1) shows the spread of the time taken to complete the modelling exercise by the

10 (DES and SD) modellers. At first sight, it can be observed that modelling sessions with DES modellers took slightly longer compared to SD modellers, with DES3 and DES5 taking the longest time. Looking into the number of words verbalised, the maroon bars in the graph (Figure 5-1), show that most modellers verbalised at similar levels, except for DES3 and DES5, who articulated more words. Considering the speed of verbalisation (number of words verbalised per minute) by modeller (Figure 5-2), it is obvious that this varies between individual modellers. Three groups can be distinguished. In the first group, DES1, DES2 and SD2 have the lowest speed (ranging between 40 – 55 words/minute). DES4, SD1, SD3, SD4 and SD5 form another group (group 2), where their speed varies between 65 and 85 words per minute. In a separate group (group 3), DES3 and DES5 have the highest speed, ranging between 85 and 90 words/minute. This suggests that DES3 and DES5 tend to naturally speak more compared to other modellers and this will be taken into consideration in the analysis to follow (chapter 6).

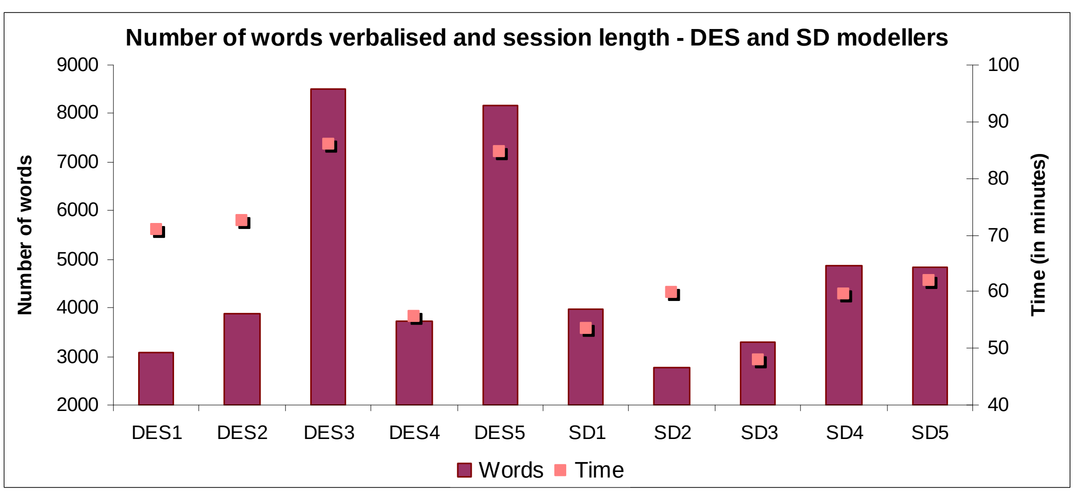


Figure 5-1: Number of words verbalised and modelling time for 5 DES and 5 SD modellers.

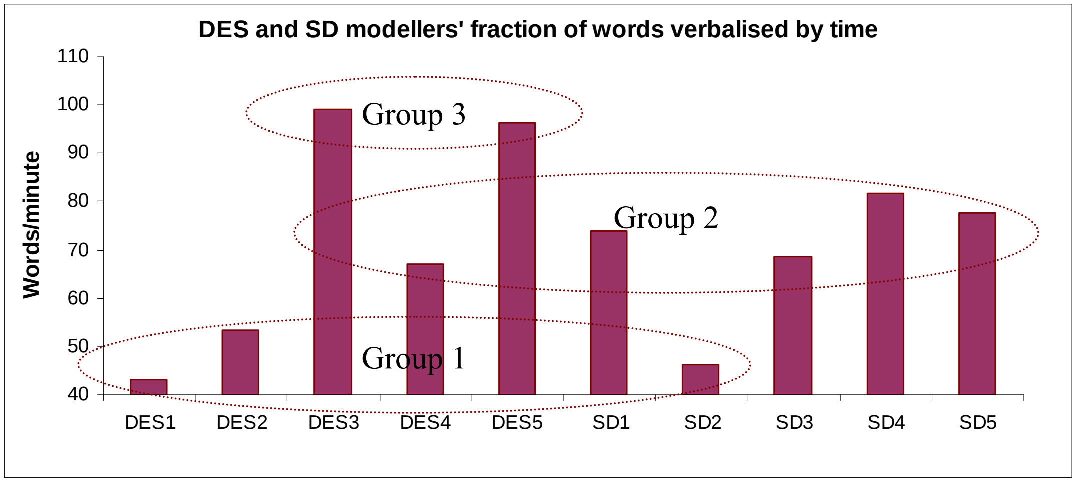


Figure 5-2: Speed of words verbalised per minute by the 5 DES and 5 SD modellers.

5.5 Design of data analysis - Coding the transcripts

The coding of the transcripts is necessary before undertaking a more systematic analysis of the verbal protocols. Therefore, in the next sections, the coding scheme used will be discussed as well as the coding process followed.

5.5.1 The coding scheme

The first step in studying the model development pattern is to set the coding scheme. The coding scheme consists of a list of topics related to the subject of the study. Willemain (1995) calls these the modelling topics. In this study, the coding scheme is devised to identify what expert modellers are thinking about in the context of simulation (DES and SD) modelling. The aim is to design a coding scheme that caters for the model development process followed in both, DES and SD modelling. A few simulation textbooks have been consulted, but there is no definitive list of the steps followed during a typical model building process, although the descriptions of the process are largely similar.

The key stages of a DES simulation study as outlined in Robinson (2004) are:

1. Conceptual modelling – understanding of the problem situation, determining problem objectives, design of the conceptual diagram and data collection.
2. Model coding – converting the conceptual model into a computer model.
3. Experimentation – obtaining results, searching for potential solutions and sensitivity analysis.
4. Implementation – implementing the findings of the simulation model or the simulation model itself and/or implementing the learning or insights gained from the model developed.
5. Verification and validation – actions taken throughout the modelling process to gain confidence in the model and its results.

Sterman (2000) suggests that modelling is a creative process and there is no set procedure followed to ensure successful SD modelling. Individual modellers have different styles and procedures (Sterman, 2000), however, he confirms that all successful modellers follow a disciplined process involving the following activities:

1. Problem articulation: articulating the problem to be addressed, including considerations about the existing issues, problems to be addressed, specifically identifying the purpose of the model, etc.

2. Formulation of dynamic hypothesis: the creation of a hypothesised theory that accounts for the problematic behaviour studied, by creating a conceptual model that explains the dynamics characterising the problem involving feedback and stock and flow structures.
3. Formulation of a simulation model: translating the conceptual diagrams into a fully specified model format, including equations, parameters and initial conditions.
4. Testing: checking the model, its equations and variables, the fit between simulated behaviour and historical system behaviour, sensitivity analysis, extreme conditions test, etc.
5. Policy design and evaluation: using the model to design and evaluate policies for improvement, creating new strategies, structures and decision rules.

Randers (1980, pp. 117 - 139) suggests the following modelling steps in an SD simulation project:

1. Conceptualisation – setting the problem area, definition of questions, description of variables in the model, development of causal diagram.
2. Formulation – developing the simulation model, by determining levels and rates and selection of parameter values.
3. Testing – testing model assumptions.
4. Implementation – testing model behaviour, testing different policies, study insights.

Pidd (2004) recognises three main phases of any generic simulation modelling study, including DES and SD modelling:

- Problem structuring – understanding of the issues to be addressed by the project.
- Modelling – involves 4 main tasks: conceptual model building, computer implementation, validation and experimentation.
- Implementation – the benefits sought from the study, in the form of specific recommendation or improved knowledge and insight.

Based on the above, a coding scheme has been designed, consisting of a new list of modelling topics. This list attempts to cater for both a DES and an SD model building approach and to fit with the statements/assumptions identified in the literature. The modelling topics and their definitions are as follows:

1. Problem structuring: What is the problem? What are the objectives of the project?
2. Conceptual modelling: Is a conceptual diagram drawn? What are the parts of the model? What should be included in the model? How to represent people? What variables are defined?
3. Model coding: What is the modeller entering on the screen? How is the initial condition of the system modelled? What units (time or measuring) are used? Does the modeller refer to documentation? How to model the user interface?

4. Data inputs: Do modellers refer to data inputs? How are the already provided data used? Are modellers interested in randomness? How are missing data derived?
5. Model results & experimentation: What are the results of the model? What sort of results the modeller is interested in? What scenarios are run?
6. Implementation: How will the findings be used? What learning is achieved?
7. Verification &Validation: Is the model working as intended? Are the results correct? How is the model tested? Why is the model not working?

All 7 modelling topics were further defined by specific sub-topics, which helped to maintain the reliability of the coding (Appendix B.3). It should be noted that the coding scheme presented is the final one. During the pilot study (section 5.2) and the coding process, the modelling topics were changed and re-defined a few times. An a priori coding scheme was created in the early stages of the model building study before implementing the modelling sessions. However, this was further enriched and changed to include more sub-topics or to restructure the existing ones. Care was more particularly taken in the definition of the codes, to ensure that the categories are mutually exclusive.

5.5.2 The coding process

The coding process starts with the definition of a coding scheme. As the modelling sessions were completed, the recorded information in a verbal protocol was transcribed. Then the verbal protocols were divided into episodes or 'thought' fragments, where each fragment is the smallest unit of data meaningful to the research context. Then each episode was coded into one of the 7 modelling topics or an 'other' category for verbalisations that were not related to the modelling task. Some episodes, however, referred simultaneously to 2 modelling topics and, therefore, were given two modelling topics. Some examples of episodes by modelling topic are provided below:

Problem structuring:

"*The purpose of the model is to test the strategy …"* (SD1)

"*…it looks like the aim is probably something to do with the reduction of recidivism and the reduction of inflow into the prison system in the first place*." (SD2)

Conceptual modelling:

"*so we need some sort of process for re-offending, which needs to take place every year*." (DES3)

"*So I guess, I would probably actually do it [conceptual diagram] on paper, but for simplicity reasons I will start drawing it in Powersim*." (SD3)

"*So what's going in to serious would be … let's call them new serious…*" (SD4)

Model coding

"*I'm going to set a label in here as well. If they are re-offending I'm going to set minimum wait time label on them, so I'm going to say: set a new label called mmm…*" (DES2) "*therefore the monthly rate is that divided by 12*." (SD4)

Data inputs:

"*Do we have only one piece of information that tells us the period before re-offending is 2 years before re-offend?*" (SD1)

"*Beyond that, we can call it 20% of serious prisoners re-offending for the sake of argument*." (DES4)

Regarding the nature of the coding process followed, a mix of top-down and bottom-up approach to coding was taken (Ericsson and Simon, 1984; Patrick and James, 2004). A theoretical base was already established (the research hypotheses on the comparison of the model building approach – section 3.4.2), which enabled a top-down approach. Throughout the various checks of the coded protocols undertaken, the coding categories were further re-defined through a bottom-up approach. Coding was an iterative process, where the coding scheme was refined as the researcher went through more protocols. This was more prevalent while analysing the protocols obtained from the pilot study (section 5.2), however, even during the coding of the main protocols, some changes were still made.

The transcripts were coded manually using a standard word processor. Initially, it was attempted to use Atlas.ti (a specialised software for the qualitative analysis of textual data), but this was found inconvenient and it was, therefore, decided to use a standard word processor. Automatic coding was not considered appropriate. According to Willemain (1995), the coding process requires attention to the context a phrase is used in and, therefore, subjectivity in the interpretation of the scripts was unavoidable. In order to deal with subjectivity, multiple independent codings were undertaken in two phases. In the first stage, the researcher coded the transcripts twice with a gap of 3 months between codings. Overall, a 93% agreement between the two sets of coding was achieved, which was considered acceptable. The differences were examined and a combined coding was reached. Next, the coded transcripts with the combined codes were further blind checked by a third party, knowledgeable in OR modelling and simulation. In the cases where the coding did not agree, the researcher and the third party discussed the differences and re-examined the episodes to arrive at a consensus coding. Overall, a 90% agreement between the two codings was achieved. A final examination of the coded transcripts was undertaken by the researcher with a view to matching the hypotheses driving the analysis and the codes. Some more changes were made to the codings, but these were fairly minor. The resulting coded protocols are analysed in this thesis.

5.5.3 Data analysis

The data (verbal protocols) obtained from the modelling sessions are analysed from the perspectives of the 1st and 2nd research objectives. The approach of the data analysis process differs based on these specific objectives.

As has already been discussed in chapter 3, objective 1 focuses on the quantitative comparison of the DES and SD model building process, and more specifically on the sequence of events followed. Therefore, the frequency of the occurrence of each modelling topic in each simulation approach is used as the basis for comparison. The main statistical estimates considered are: the number of words and the percentage of respondents' verbalisations per modelling topic[8]. For the comparison of the amount of verbalisations between the DES and SD groups, the Kolmogorov Smirnov test for two independent samples is used. The Kolmogorov-Smirnov test is a nonparametric alternative to the t-test for two independent samples and it is used when there is a reason to believe that the data violates the condition of normality (Sheskin, 2007). The test has a high power efficiency when used for small sample sizes (Siegel and Castellan, 1988). In addition, timeline plots were created, which represent the modelling topics the modellers attend to during the progression of the model building session. These consist of a set of matched timeline plots, including the 7 modelling topics (section 5.5.1). Moreover, the patterns of iterations were examined, by checking the sequence between consecutive episodes, in each individual protocol. Microsoft Excel was used to store and analyse the data. The quantitative analysis aims to explore the accuracy of hypotheses 1.1, 1.2 and 1.3 (section 3.4.1). The analysis is provided in chapter 6.

The analysis regarding the 2nd research objective takes a qualitative approach and, therefore, content (text) analysis is used. This analysis consists of matching the hypotheses 2.1 to 2.18 (section 3.4.2) with the relevant text found in the DES and SD verbal protocols. Inferences about each hypothesis are made based on

[8] (word count on modelling topic / total word count) x 100

indications found in the text. During the study design, and more specifically during the pilot study (section 5.2), it was considered important to assess the viability of studying the 18 research hypotheses taking into consideration the methods and the case study chosen. As a result, the stated hypotheses have been divided into 3 main categories: testable, non-testable and factual. Testable hypotheses (2.1, 2.3 - 2.5, 2.7, 2.8, 2.10 - 2.18) are the ones that the present study can potentially test and the ones, which are further analysed in chapter 7.

The hypotheses that cannot be tested by this study are called non-testable. These are hypotheses: 2.2 and 2.9. The reason for classifying them as non-testable is because the features these hypotheses refer to are based on the nature of the problem modelled (tactical/operational level and modelling of material and information flows) (Table 5-1). Given that only one case study situation is used in this research, these hypotheses cannot be tested by the current study.

Hypothesis 2.6 (Table 5-1) is considered to be a factual hypothesis. The reason being is that it is based on the fundamental modelling characteristics of each modelling approach and the related analysis does not lead to additional knowledge.

The non-testable and factual hypotheses are not pursued in the data analysis carried out in chapter 7. These hypotheses are listed in Table 5-1.

Table 5-1: Non testable and factual research hypotheses not verified in the study

Hypotheses	Reason
Non-testable hypotheses: **Hypothesis 2.2:** DES models problems at tactical/operational level, while SD at a strategic level. **Hypothesis 2.9:** DES models represent mainly material flows. Information flows can be incorporated but these are not obvious. In SD modelling both material and information flows are equally represented.	Based on the nature of the problem
Factual hypothesis: **Hypothesis 2.6:** In DES modelling objects/people are represented as distinct individuals while in SD aggregate system behaviour is examined.	Based on fundamental modelling characteristics of each approach.

5.6 Summary of model building study design

This chapter provides an account of the activities undertaken to design the study of model building, which is aimed at understanding the differences in the DES and SD model building process and modelling approach. The main aspects regarding the study design include: the selection of participating expert modellers, the modelling sessions implemented, the coding scheme and coding process and the strategy followed to analyse the data collected. The pilot study implemented is also presented, which provided some initial insights regarding the two model building approaches and also helped in improving the study design.

Chapter 6: Quantitative analysis of VPA data

6.1 Introduction

Chapter 5 described the design of the model building study for the comparison of DES and SD modelling. This chapter performs a quantitative analysis of the 10 verbal protocols in line with the first research objective. This objective aims to empirically determine the differences in the modelling process followed by DES and SD modellers. The quantitative analysis explores the distribution of attention by modelling topic as well as the sequence of modellers' attention among topics. Therefore, the DES and SD modellers' thinking processes during a model building exercise is tracked with the view to comparing the modelling process followed.

6.2 Participants profile

Before analysing the data obtained from the model building study, it is important to consider the profile of the participating modellers. As has already been mentioned, the choice of the two DES and SD samples was made with the view to including a diversity of participant profiles (section 5.3.2). Thus a mixture of backgrounds in terms of modelling experience, organisations and type of simulation software used, was sought.

The participants who took part in the modelling sessions are modellers with a high level of experience in simulation, most of them holding consultant posts in different organisations. The companies they come from are established simulation software or consultancy companies based in the UK. All participants have completed either doctorates or masters' degrees in engineering, computer science, Operational Research or hold MBA degrees. Their experience in modelling ranges from 4 to 19 years. They have also acquired supplementary OR and simulation training as part of their jobs. They boast an extensive experience of modelling in areas such as: healthcare, criminal justice, food & drinks sector, supply chain, etc. Their names are not disclosed for confidentiality reasons. Hence, they are called by the modelling approach they represent followed by an ordered number from 1 to 5, based on the sequence of interviews. So DES subjects are called DES1, DES2, DES3, DES4 and DES5, while SD modellers are denoted: SD1, SD2, SD3, SD4 and SD5. A summary of all modellers' qualifications and experiences follows. These are also listed in Table 6-1.

- DES1 is a consultant with 9 years experience in modelling. He was initially introduced to simulation modelling at university. In his later career he was involved with various types of simulation modelling, 3-D simulation, i.e. robotics, Quest and C++ and then progressed onto using Witness. He spends almost 90% of the time building simulation models and the rest of the time he spends in communication with the clients.

- DES2 is a simulation consultant with 4 years experience in modelling, using mainly Simul8. Before that he was working on simulation-related projects in an established British University. DES2 spends 85% of the time building simulation models for client projects.
- DES3 has 19 years of experience in simulation modelling, working as a consultant or senior consultant in various organisations. DES3 has also had experience of using other simulation methods (SD or Agent-based modelling), but claims that DES is his area of expertise. He used Flexim for the purposes of the model building exercise. DES3 spends almost 90% of the time building simulation models.
- DES4 is a senior consultant with approximately 8 years' experience, working on a range of projects involving simulation. His experience of simulation software ranges from: WITNESS, Simul8, Arena and ProModel. As part of the exercise, he preferred to use Simul8. DES4 spends almost 80% of the time building simulation models for clients.
- DES5 is a senior modeller working for a consultancy company using simulation modelling. He learned simulation modelling while studying for an undergraduate degree in a highly established British University. He has 4 years of experience in DES modelling and has used a range of software packages, including: Witness and Quest. In the modelling session he used Witness. DES5 spends almost 90% of his time building simulation models.

In the group of SD modellers:

- SD1 has 14 years experience of modelling and runs a consultancy company. For the last 10 years he has been extensively involved in consultancy projects using SD modelling to support strategy in public and private sector organisations. In the modelling sessions SD1 used Stella/iThink.
- SD2 also has a long experience in modelling having spent 16 years working with senior managers and executives on a range of consultancy projects. He uses a range of System Dynamics software, but he highlights that modelling *per se* is not the objective, but rather it is the tool that aids his consultancy activities. SD2 used the Strategy Dynamics software in the modelling session.
- SD3, who used Powersim has 12 years experience in modelling. With background in engineering, SD3 was initially involved in physical modelling. He then undertook a range of OR analyst and consultancy positions. His experience in SD modelling specifically is about 4 to 5 years long, however, he recognises commonalities between SD and the other modelling approaches he undertook in the past.
- SD4 is an independent consultant with over 20 years of experience in SD modelling. He has been involved in a range of projects within the public sector and claims that he spends about 70% of working time on SD simulation modelling. SD4 used Stella/iThink.
- SD5 is a practitioner who started his modelling career studying SD modelling as part of his doctoral degree. Since then he has undertaken a number of SD modelling projects and boasts 10 years experience in SD

modelling. SD5 claims to spend 60% of his time building simulation models. He used Vensim to build the simulation model.

Table 6-1: List of DES and SD modellers' profiles

Modeller	Modelling experience	Simulation software used
DES group		
DES1	9 years	Witness
DES2	4 years	Simul8
DES3	13 years	Flexim
DES4	8 years	Simul8
DES5	4 years	Witness
SD group		
SD1	14 years	Stella/iThink
SD2	16 years	Strategy dynamics
SD3	5 years	Powersim
SD4	20 years	Stella/iThink
SD5	8 years	Vensim

An attempt has been made to achieve a balance between the two groups of modellers. In the search for suitable participants for this study, difficulties were encountered because there is a limited number of expert modellers who specialise in using simulation modelling. Furthermore, their names are not publicly available, unless they were recommended to the researcher by someone who had personal contact with them. Coupled with the unease of some organisations to take part in any survey, there was not much scope to choose among expert modellers who could potentially participate. Table 6-2 shows the tabulation of the participants involved in the study by modelling experience and type of simulation software used. The level of experience has been grouped into junior (4-7 years) and senior (8 years and above) expert modellers. Two participants from each group had 4-7 years simulation modelling experience, and 3 participants with 8 or more years experience in each group (There were 2 participants with 8 and above years experience in Stella in the SD group). Therefore, it is believed that a balance is achieved in the profiles of the DES and SD sample groups.

Table 6-2: DES and SD participants tabulated in groups of experience by software used.

DES modellers	**Type of simulation software**			**Total (5)**
Modelling experience	Witness	Simul8	Other (Flexim)	
Junior (4-7 years)	DES5	DES2	-	2
Senior (8+ years)	DES1	DES4	DES3	3
SD modellers	**Type of simulation software**			**Total (5)**
Modelling experience	Powersim	Stella	Other (Strategy dynamics/ Vensim)	
Junior (4-7 years)	SD3	-	SD5	2
Senior (8+ years)	-	SD1,SD4	SD2	3

Experts in VPA believe that the level of experience affects the amount of verbalisations (Ericsson and Simon, 1984). Therefore, it is considered important to explore how the mixture of the two groups affects their respective verbalisations. Figure 6-1 displays a graphical representation of the amount of verbalisations for all 10 protocols against the modellers' level of experience. A linear regression line fits the data points of the scatter diagram, suggesting that with the increase in experience levels, more experienced modellers tend to verbalise less. 2 outliers have been identified, and this is the case of DES3 and DES5.

When exploring the number of words verbalised and the average number of words articulated per minute for each modeller, the DES and SD groups are not considered very different (section 5.4). However, DES3 and DES5 are found to have the highest values compared to all other modellers, suggesting that they talk fastest

(Figure 5-2). DES3, represents the senior group of modellers (with an experience of 19 years), who contrary to expectations verbalised longer than the average participants, regardless of experience. DES3 encountered technical issues while building the model, which resulted in a longer session, during which time DES3 continued to verbalise about the model and the problems encountered. DES5 represents the group of junior modellers (with 4 years experience), who naturally talks to a high extent (He admitted this at the end of the modelling session). It was observed that DES5 showed an above average interest in the exercise. His session was quite long because he continued modelling until he was confident that a complete model was created. The 2 outliers have not been included in the calculations of the regression line[9].

Due to using only 8 data points the mathematical regression line estimated can be only considered at an exploratory level. It is clear that with the increase in the level of experience, simulation modellers (DES and SD) tend to verbalise less, however, there can be exceptions.

[9] Please note that the 2 protocols identified as outliers when looking at the relationship between the level of experience and amount of verbalisation, does not render them unsuitable for the data analysis carried out next in chapters 6 and 7.

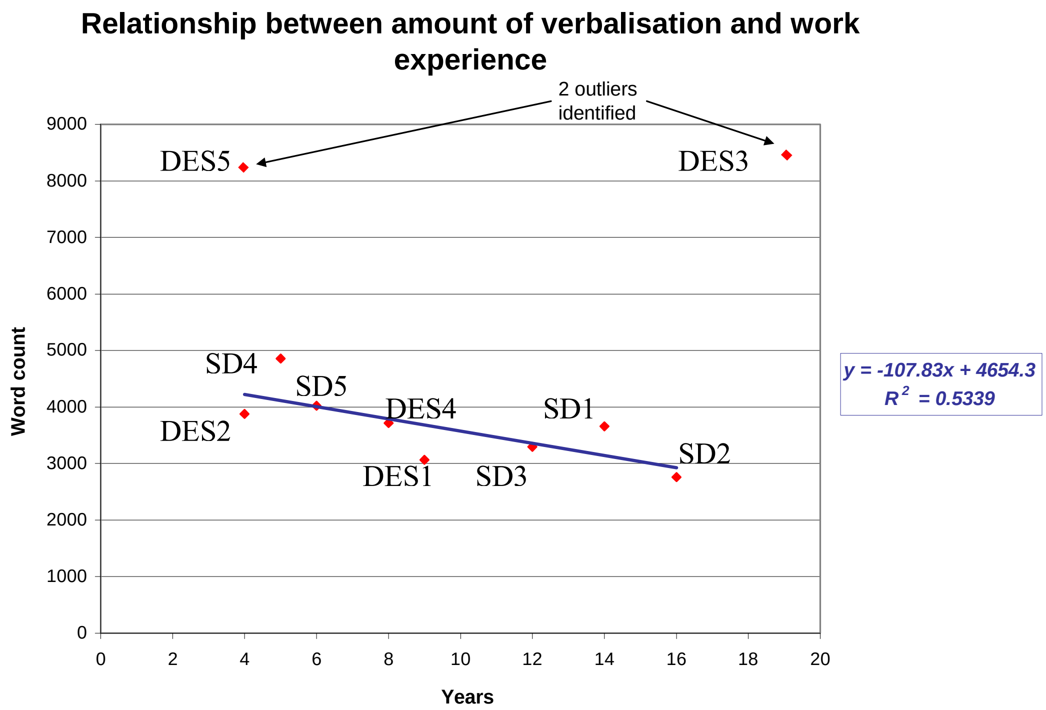


Figure 6-1: Correlation of the amount of verbalisation with length of modelling experience of the participant sample.

In summary, the samples obtained for the two groups are considered reasonably comparable. However, the mixture of the two sample groups should be taken into consideration when comparing the verbalisations of the two groups of modellers. DES3 and DES5 had a natural tendency to speak more than other modellers and, therefore, may well result in a higher level of verbalisations for the DES group. On the other hand, the SD group of modellers had a higher level of experience compared to the DES group of modellers, which could result in less verbalisations.

6.3 Distribution of attention

In this section, the 10 coded verbal protocols are analysed quantitatively in order to explore the distribution of attention by modelling topic. The number of words articulated is considered a suitable measure of the amount of verbalisation by the modellers. In turn, this is used to indicate the spread of modellers' attention to the different modelling topics. This is related to hypothesis 1.1 (section 3.4.1).

Some preliminary observations can be made by comparing the total verbalisations of the 5 DES and 5 SD modellers by modelling topic. The bar chart (Figure 6-2) suggests that DES modellers spend more time with coding the model and testing it, whereas SD modellers spend more time with creating the conceptual model. Furthermore, a considerable difference is observed in the total number of words verbalised, where the 5 DES modellers verbalised 27,352 words in total compared to 18,596 words verbalised by the 5 SD modellers. The difference is 8,756 words, which accounts for 32% of DES modellers' verbalisation. It is, however, possible that the higher total verbalisations result due to the longer sessions and consequently the higher number of words verbalised by DES3 and DES5 (section 6.3.2).

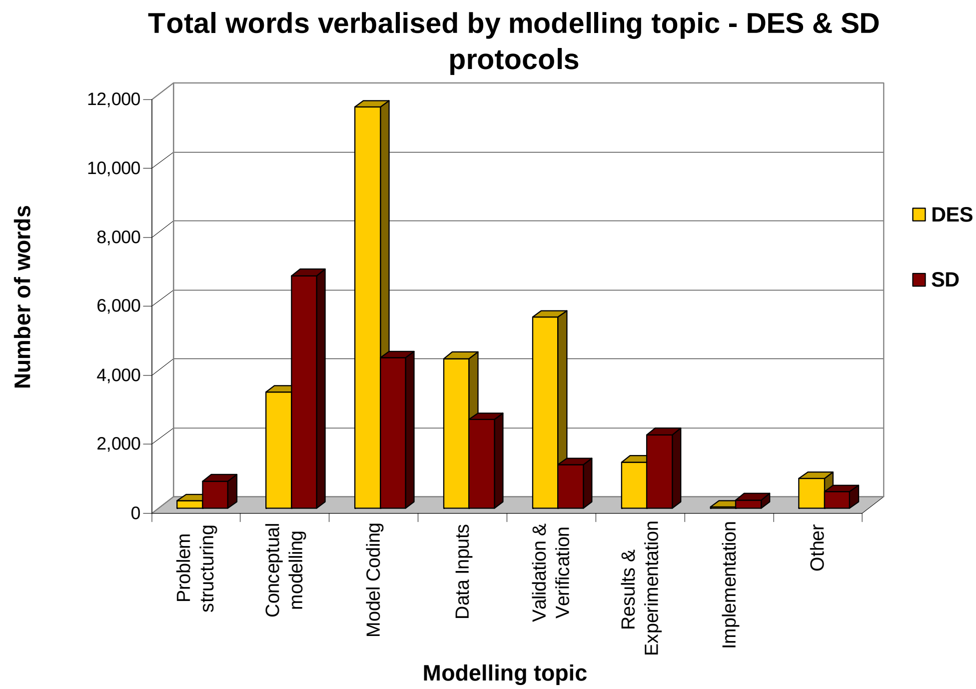


Figure 6-2: Total number of words verbalised by 5 DES and 5 SD modellers by modelling topic.

Next, a comparison of the number of words in the DES and SD protocols is carried out in order to establish whether the verbalisations between the two groups of modellers are significantly different. Table 6-3 shows a detailed comparison, including the respective average word count by modelling topic and the resulting differences for the two groups of modellers. The standard deviation, as a measure of the variation of verbalisations for each modelling topic around the mean is also calculated. The fraction of the standard deviation over the mean is also calculated for the purposes of comparability where the means are very different. In the last column, the extent and sign of the difference between the group means (DES – SD) are also displayed.

Table 6-3: Average number of words and fraction of the standard deviation/mean calculated for each modelling topic for DES and SD protocols.

	DES			SD			Differences in means
	Average words	Standard deviation	Standard deviation/mean	Average words	Standard deviation	Standard deviation/mean	
Problem structuring	44	51	1.2	158	164	1.0	-114
Conceptual modelling	674	656	1.0	1,350	493	0.4	-676
Model Coding	2,327	1,368	0.6	873	386	0.4	1,454
Data Inputs	869	331	0.4	514	191	0.4	355
Verification & Validation	1,109	434	0.4	253	141	0.6	856
Results & Experimentation	268	195	0.7	427	222	0.5	-160
Implementation	8	17	2.2	48	72	1.5	-40
Other	172	79	0.5	96	60	0.6	76
Total protocol	5,470	2,646	0.5	3,719	789	0.2	1,751

Comparing the average number of words in the overall DES and SD protocols an average difference of 1,751 words is identified, suggesting that DES modellers verbalise more than SD modellers. Considering each specific modelling topic, the biggest differences between the average DES and SD protocols (Table 6-3), can be identified with regards to model coding, verification & validation and conceptual modelling. This suggests that DES modellers spend more effort in coding the model in the computer and testing it, while SD modellers spend more effort in conceptualising the mental model. Furthermore, estimations of the standard deviation suggest that there is slightly more variation in the verbalisations of DES protocols compared to SD modellers. The relative standard deviation is consistently greater for DES modellers.

6.3.1 Comparison of the amount of verbalisation

In the box and whiskers plot (Figure 6-3) the amount of verbalisations per modelling topic and for the overall DES and SD protocols is displayed. The biggest differences are observed for the topics: conceptual modelling, model coding, data inputs, verification and validation. A significant difference in the spread of the data is observed for the total number of words verbalised, even though the medians are quite close. Therefore, a statistical test is needed to compare the significance of the differences between DES and SD modellers. The Kolmogorov-Smirnov test, a nonparametric statistical test, is used for the comparison of the word counts verbalised by DES and SD modellers.

The Kolmogorov-Smirnov test is used as a nonparametric alternative to the t-test for two independent samples when it is believed that the hypothesis of normality does not hold (Sheskin, 2007). In this case, only 5 data points (word count for each modeller) are collected from the two groups of modellers (DES and SD). Due to the small sample size and the fact that count data is inherently not normal, the assumption of normality is violated. Indeed, the Lilliefors test for normality (Sheskin, 2007) rejects the hypothesis of normality for all data points (the 7 modelling topics and the total word counts).

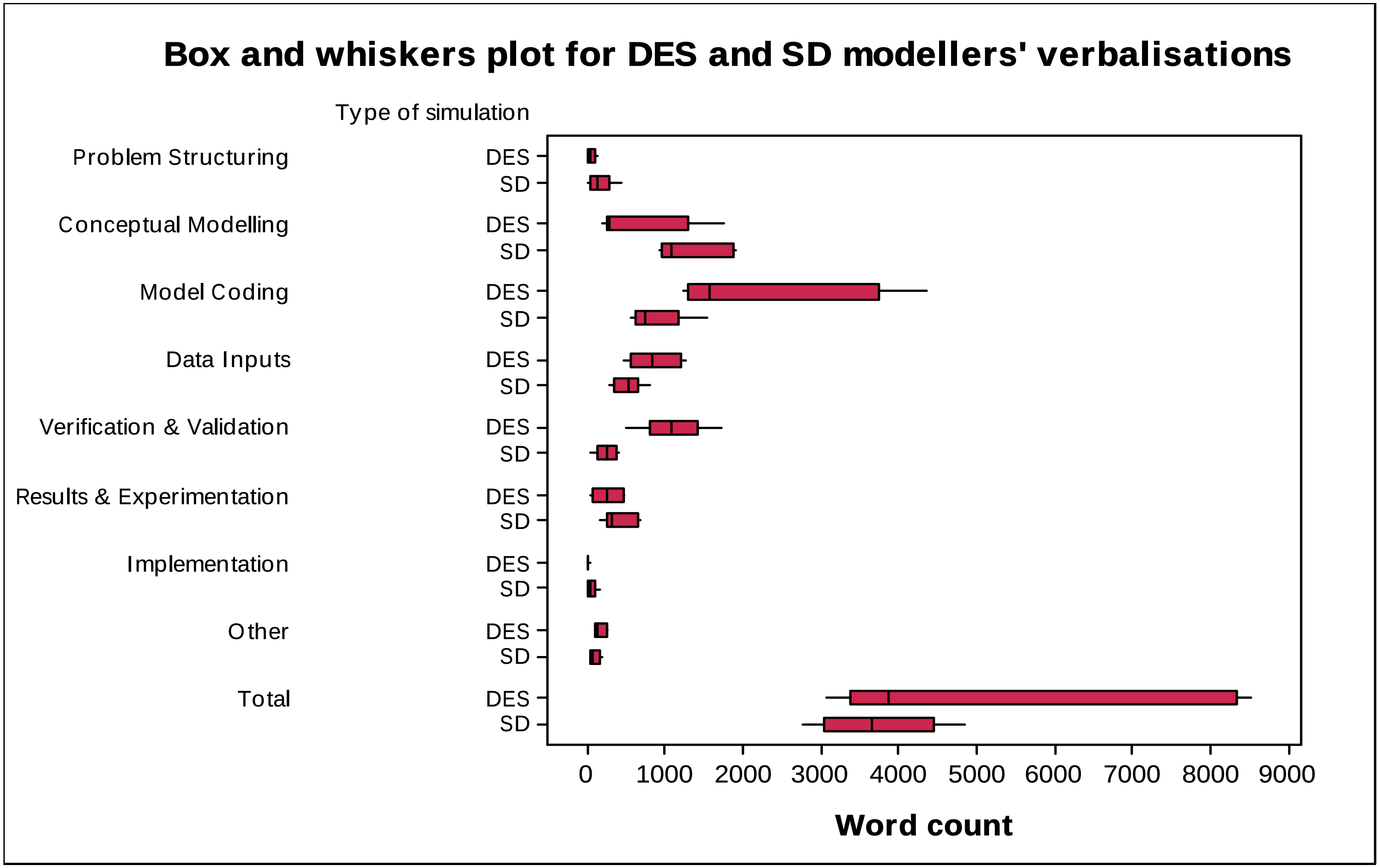


Figure 6-3: Box and whiskers plot of DES and SD modellers' verbalisations by modelling topic.

The null hypothesis for the Kolmogorov-Smirnov test assumes that the verbalisations of the DES modellers follow the same distribution as the verbalisations of the SD modellers. The alternative hypothesis is that they do not come from the same distribution. This test compares the cumulative probability distributions for the DES and SD sample. The test statistic, identified as M, is specified at the point which represents the greatest vertical distance at any point between the two cumulative probability distributions. An example is provided in Figure 6-4. The rest of the tests are provided in appendix C.1 Table 6-4 shows an example of the comparison of the verbalisations on problem structuring. The raw data represent the count of words verbalised by each participant in the group ranked from lowest to highest for the two groups. Note that if the same number of words is verbalised by more than one subject in each group then each value is recorded in the same row. The cumulative proportion associated with the value of each modeller's verbalisations is recorded in columns S1 and S2 for the DES and SD groups respectively. The test statistic (M) is derived for the row that has the largest absolute value of the greatest vertical distance at any point between the two compared probability distributions. This is further compared to the tabled critical value for the Kolmogorov-Smirnov test, which is specified based on the values of the sample sizes. For two independent samples $n_1 = n_2 = 5$ and a two-tailed test at 0.05 level of significance, the tabled critical value is 0.8 (Sheskin, 2007). The null hypothesis is rejected if the test statistic, M, is equal to or greater than 0.8 (the corresponding critical value).

Problem structuring							
	DES			SD		S1-S2	
		S1			S2 (Cum	(Vertical	
	Raw data	(Cumulative		Raw data	distribution)	distance)	
DES2, DES4	0, 0	0.40		-	0	0.40	
	-	0.4	SD4	6	0.2	0.20	
DES5	42	0.6		-	0.2	0.40	
DES1	53	0.8		-	0.2	**0.60**	M
	-	0.8	SD3	67	0.4	0.40	
DES3	125	1	SD5	125	0.6	0.40	
	-	1	SD2	160	0.8	0.20	
	-	1	SD1	431	1	0.00	

Figure 6-4: Example of data verbalisations on problem structuring, illustrating how M, the greatest vertical difference is identified.

Table 6-4: The results of the Kolmogorov-Smirnov test comparing the DES and SD modellers' verbalisations for 7 modelling topics and the total protocols. The significant differences are highlighted, based on the comparison of the greater vertical distance M to the critical value =0.8.

Modelling topic	M	Differences in verbalisations?
Problem structuring	0.6	×
Conceptual modelling	0.8	✓
Model coding	0.8	✓
Data inputs	0.6	×
Verification & Validation	1	✓
Results & experimentation	0.4	×
Implementation	0.4	×
Total protocol	0.4	×

The statistical tests performed indicate significant differences, at a 5% level, in the amount of DES and SD modellers' verbalisations for the three modelling topics: conceptual modelling, model coding and verification & validation. This suggests that DES modellers verbalise more with respect to model coding and verification & validation and thus spend more effort on these modelling topics compared to SD modellers. SD modellers verbalise more on conceptual modelling. For the modelling topics: problem structuring, data inputs, results & experimentation and implementation, the Kolmogorov-Smirnov test suggests that there are not significant differences between the two groups of (DES and SD) modellers. In addition, the total verbalisations of the two groups of modellers are not found to be significantly different.

6.3.2 Comparison of the proportion of attention

The comparison tests carried out in section 6.3.1 compare the absolute values of the verbalisations for each modelling topic between the two groups of modellers and it can be suggested that the results might have been affected by the variation of verbalisations among modellers. For example, DES3 verbalised almost three times the amount of words verbalised by DES1. For the purposes of weighting these numbers with respect to the variation in the amount of verbalisations among modellers, the proportion of the attention distributed among the 7 modelling topics is calculated. The proportion of attention for a specific modelling topic is calculated by dividing the equivalent number of words and the overall number of words verbalised in each protocol, expressed as a percentage. The mean proportions for the 5 DES and 5 SD modellers are displayed in the graph (Figure 6-5).

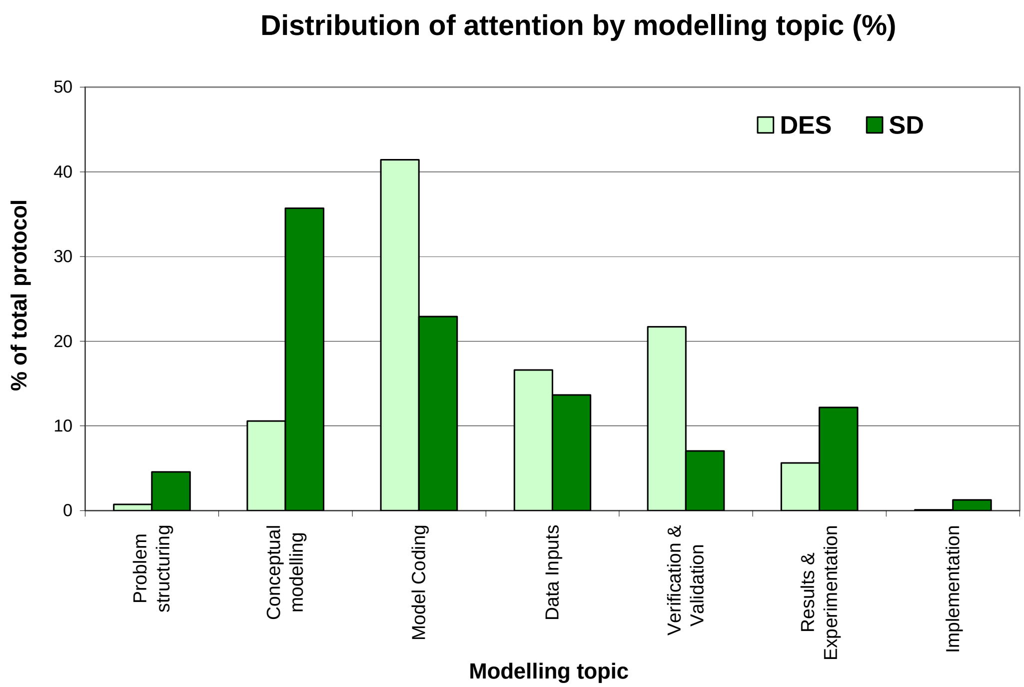


Figure 6-5: Average percentage of attention paid to the 7 modelling topics by DES and SD modellers.

Considering the distribution of attention of DES and SD modellers among the 7 modelling topics (Figure 6-5), the highest differences are observed in the percentage of attention paid to model coding, conceptual modelling and verification & validation. On average, DES modellers devote 41% of attention to model coding, with the next most important topics being verification & validation (22%) and data inputs (17%). While SD modellers, spend on average 36% of their attention on conceptual modelling and 23% of attention on model coding (Table 6-5).

Table 6-5: DES and SD modellers' verbalisations (number of words and proportion of attention) by modelling topic.

	DES		SD	
Modelling topic	**Number of words**	**Proportions**	**Number of words**	**Proportions**
Problem structuring	220	0.8%	789	4.4%
Conceptual modelling	3,371	12.7%	6,751	37.3%
Model Coding	11,636	43.9%	4,366	24.1%
Data Inputs	4,345	16.4%	2,572	14.2%
Validation & Verification	5,544	20.9%	1,266	7.0%
Results & Experimentation	1,338	5.1%	2,136	11.8%
Implementation	38	0.1%	238	1.3%
Total protocol	**26,491**	**100%**	**18,117**	**100%**

It can be therefore, observed that the modellers' distribution of attention to the 7 modelling topics differs depending on whether the modeller is a DES or SD expert (Figure 6-5 and Table 6-5).

To further verify this observation, a chi-square test is performed comparing the distribution of the number of words verbalised among the modelling topics for the two groups of modellers. The chi-square tests whether there are any differences in the distribution of the number of words verbalised among the 7 modelling topics (excluding "Other") for the two groups of modellers. The null hypothesis assumes that the distribution of attention (the number of words) among all modelling topics is equal for the two groups of modellers. The alternative hypothesis assumes that the distribution of attention among modelling topics is different between the two groups of modellers for at least one of the modelling topics. The critical value of the chi-square test at the 5% confidence level and 6 degrees of freedom[10] is: ${\chi^2}_{0.05,6} = 12.59$ (Daniel and Terrell, 1995). From calculations, the chi-square

[10] Given the data of the distribution of attention for the 7 modelling topics (excluding the 8th category 'Other') for the two groups of modellers forms a contingency table, with 7 rows and 2 columns the degrees of freedom is calculated by: (number of rows -1)*(number of columns -1) = (7-1) x (2-1)=6.

value found is 6,892.89 (> 12.59). Therefore, the null hypothesis assuming similar distributions of words is to be rejected and it can be concluded that the distribution of attention among modelling topics is not the same for DES and SD modellers. Indeed in the graphical display (Figure 6-5), it can be observed that the highest differences between the DES and SD modellers groups are seen for the attention paid to conceptual modelling, model coding and verification & validation. It is, therefore, concluded that SD modellers pay the highest attention to conceptual modelling, whereas DES modellers concentrate on model coding and verification & validation.

6.3.3 Comparing DES & SD modellers' distribution of attention with OR experts

The data derived from the modelling sessions with simulation experts carried out in this study could be compared to the findings from Willemain's (1995) study, where 12 general OR expert modellers were involved. The 7 topics used in the current research are quite similar to the ones used by Willemain. Some of the topics are adapted to fit Willemain's 5 modelling topics. His modelling topics are as follows: context (equivalent to problem structuring in the current study), structure (equivalent to conceptual modelling), realisation (includes model coding and data inputs), assessment (verification & validation) and implementation (includes results & experimentation and implementation). A comparison of the bar charts (Figure 6-6) shows that the patterns of attention followed by OR experts and the DES and SD experts participating in the current study differ. OR experts spent more effort on structure, whereas the DES and SD modellers in this study spent more attention on realisation (or model coding and data inputs). This is not surprising, considering that the exercises and the nature of the tasks differed. This study focuses more on the aspect of building simulation models on the computer. In Willemain's study the problems used were less structured and thus the exercises required different levels of attention to the modelling topics.

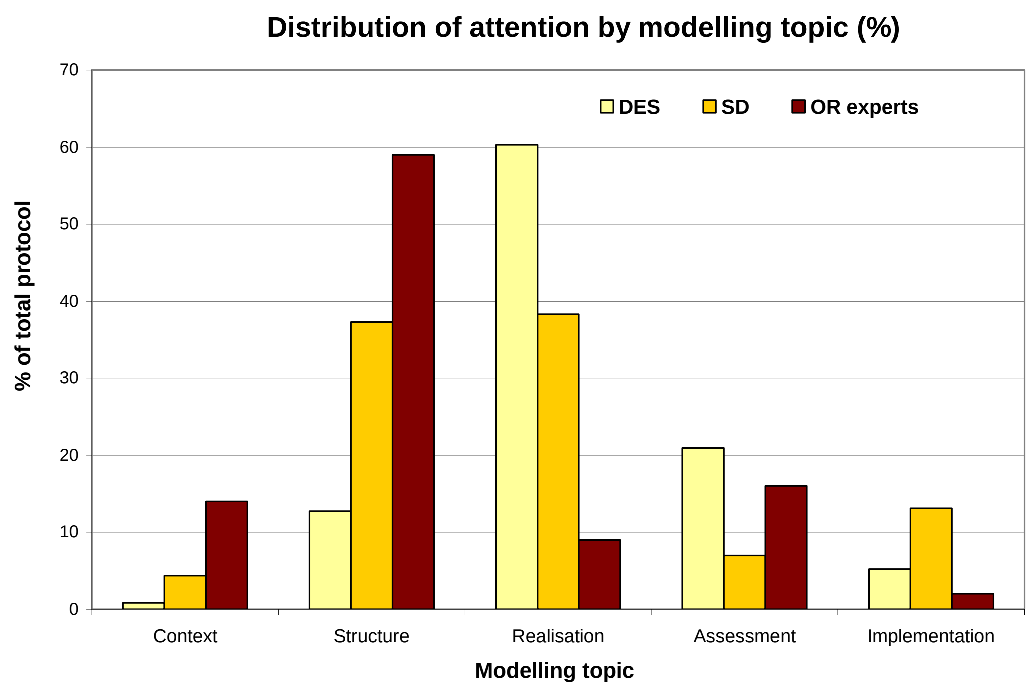


Figure 6-6: Comparing the distribution of attention to 5 modelling topics for DES, SD and Willemain's OR experts.

6.3.4 Summary of findings - comparison of DES & SD modellers' distribution of attention

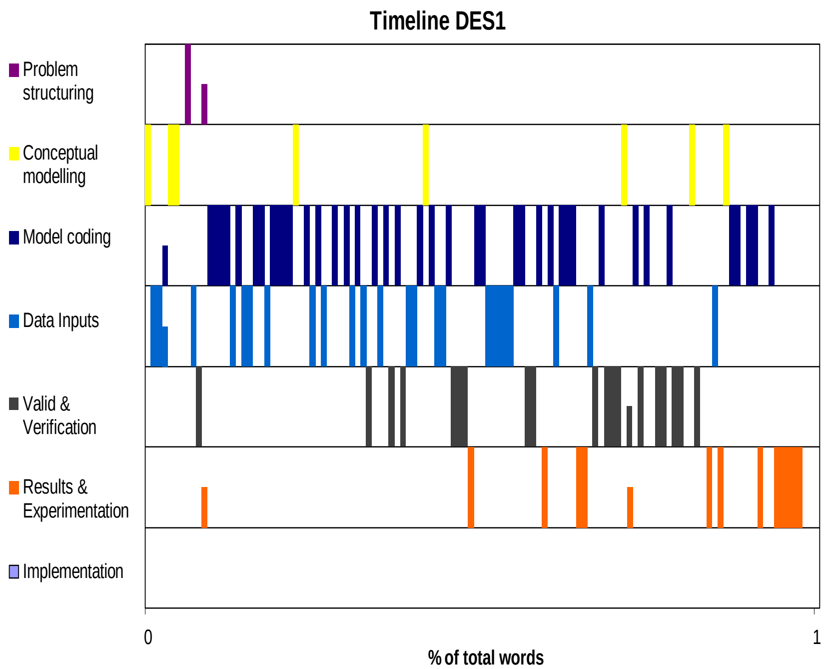
In summary, two types of analyses have been carried out, the comparison of the differences in the number of words verbalised (section 6.3.1) and the comparison of the distribution of attention paid to the 7 modelling topics (section 6.3.2), between DES and SD modellers. In the first case, a nonparametric statistical test is performed, which found that DES and SD modellers verbalise to the same extent. Statistically significant differences were found in their amount of verbalisations for conceptual modelling, model coding and verification & validation. The analysis on the distribution of the proportion of attention revealed that the distribution of attention to the 7 modelling topics indeed differs, and the most significant differences are found for the same topics (conceptual modelling, model coding and verification and validation). These findings suggest that DES modellers spend more attention on coding the model on the computer and verifying or validating it, while SD modellers pay more attention to conceptual modelling. Differences are however, observed between OR experts' and simulation experts' distribution of attention to Willemain's (1995) modelling topics.

6.4 The progression of attention during a simulation modelling task

After comparing the distribution of attention to the different modelling topics for the DES and SD modellers, this section focuses on the progression of modellers' attention during a simulation model building task. The analysis is based on research hypotheses 1.2 and 1.3 (section 3.4.1).

Timeline plots, which show what modellers think about and when during a simulation modelling task (Willemain, 1995; Willemain and Powell, 2006), are used. A timeline plot is created for each of the 10 verbal protocols. It consists of a matched set of 7 timelines showing which of the seven modelling topics the modeller is attending to throughout the duration of the modelling exercise. The vertical axis takes three values, 1 when the specific modelling topic is attended to by the modeller, 0.5 when the modelling topic and another have been attended to and 0 otherwise (when the modelling topic is not mentioned). The horizontal axis represents the proportion of the verbal protocol, from 0 to 1 (100% of the number of words). The proportion of the verbal protocol is counted as the fraction of the cumulative number of words for each consecutive episode over the total number of words in that protocol, expressed as a percentage.

Figures 6-7 and 6-8 show 2 timeline plots each, representative of the 5 DES and 5 SD modellers' protocols respectively. The rest of the timeline plots are provided in appendix C, where the 5 DES protocols are listed in appendix C.2 and the 5 SD protocols in appendix C.3. Observing these plots it is obvious that modellers frequently switched their attention among topics. Similar patterns of behaviour were observed by Willemain (1995) in his study where expert modellers where asked to build models of a generic OR problem. Looking at the overall tendencies in the DES and SD timeline plots, it appears that the DES protocols follow a more linear progression in the sequence of modelling topics. Meanwhile, in the SD protocols, modellers' attention is more scattered throughout the model building session. To test these initial observations, the transition of attention between modelling topics is explored in section 6.4.1.



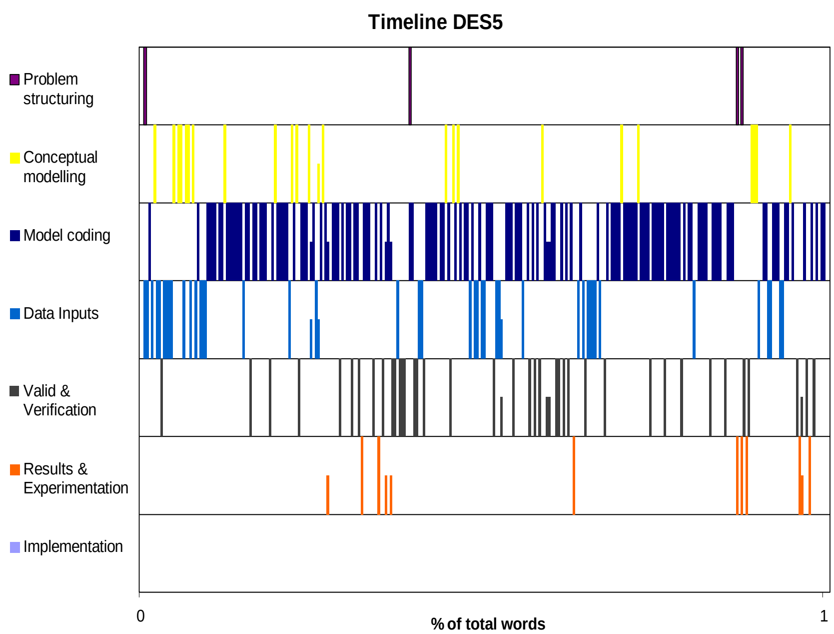


Figure 6-7: Timeline plots for DES1and DES5 verbalisations.

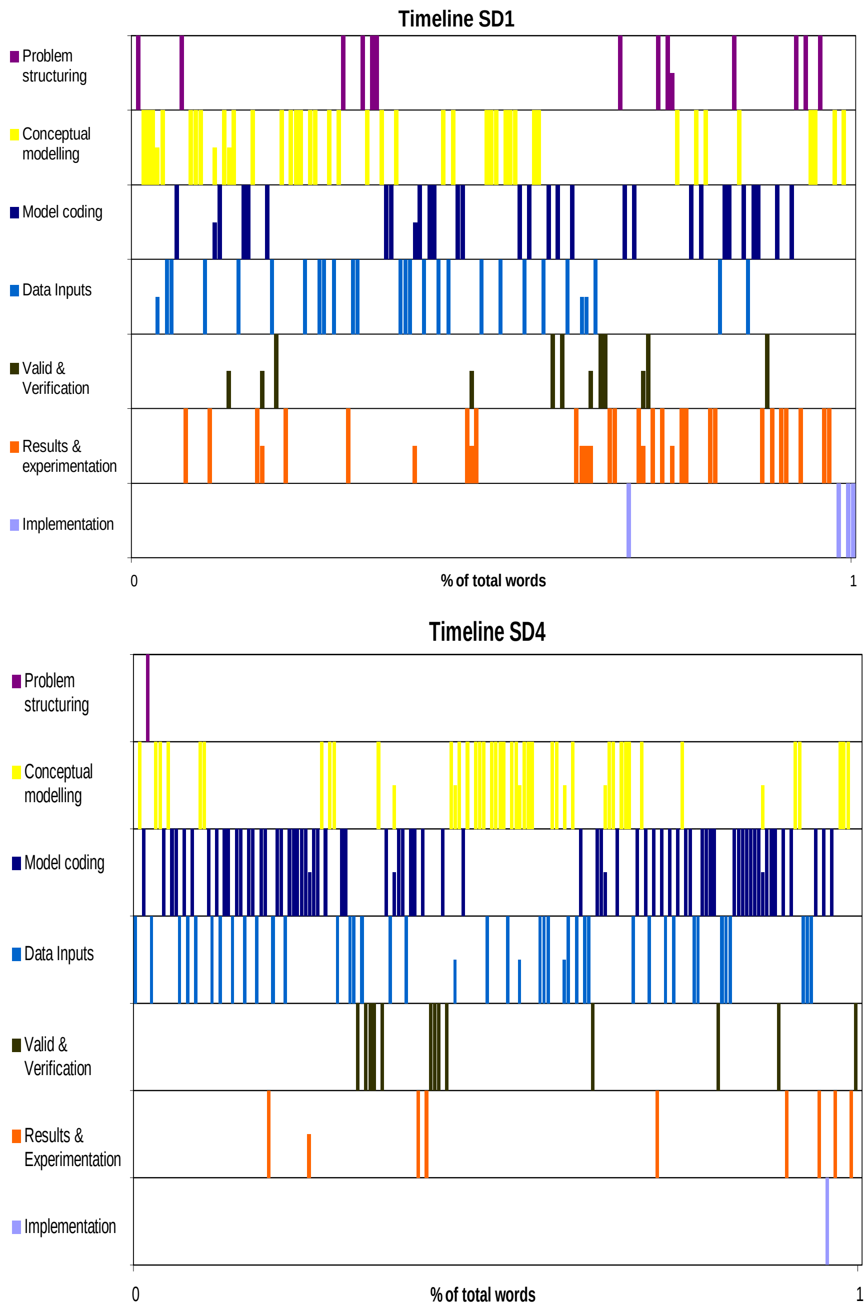


Figure 6-8: Timeline plots for SD1 and SD4 verbalisations.

The DES timeline plots (Figure 6-7 and Appendix C.2) seem to follow a similar pattern. Only DES3 timeline is somewhat different to the other protocols. DES3 attended to results & experimentations and implementation at the beginning of the protocol instead of after model building. It is important to mention that DES3 attempted to go through all the necessary modelling issues, but the model was not finished and, therefore, the thinking process not complete. Had the model been completed, it would have been expected that DES3 would have attended more to results & experimentation and implementation at the end of the protocol. Nevertheless, the protocol was considered appropriate because this still conveys the thinking process followed during a simulation modelling task and, therefore, it is included in the sample.

Looking more closely at the DES timelines (Figure 6-7 and Appendix C.2), some further observations can be made. It is observed that problem structuring is attended to rarely, but most of the times at the beginning of the protocol. Attention to conceptual modelling appears at various points throughout the modelling task. Only DES3 concentrated his/her attention to conceptual modelling more at the beginning of the protocol. Model coding seems to have a high density of bars in all DES timelines, which become more concentrated as the task progresses. Data inputs is a topic well attended to by all DES modellers. It is established that all DES modellers enter the data in the model throughout the first half or during the middle of the modelling task. Verification & validation of the model is attended to after some part of the model has been created and some data entered. Verification & validation can follow two routes. In one case the model is created and data entered and then the model is verified or validated (DES1 and DES2). Alternatively, smaller parts of the model are created and data entered and then the model validated, and then more parts of the model are created and data entered and then again validated, and so on (DES4 and DES5). Results & experimentation as well as implementation have not been attended to extensively by DES modellers. However, these tend to be attended to at the end of the protocol. An exception is observed with DES3 and the reasons for this have been discussed in the previous paragraph.

At first sight, the attention in the SD protocols (Figure 6-8 and Appendix C.3) does not appear to follow a specific pattern. SD modellers' attention to various modelling topics is spread throughout the modelling task. Each modelling topic is now discussed. Problem structuring was attended to at different points during the protocol by each SD modeller. SD1 attended to problem structuring at various points at the beginning, middle and end of the protocol, SD2, SD3 and SD4 very sparingly at the beginning or end of the protocol and SD5 in the middle of the protocol. In most SD protocols, with the exception of SD4, the density of conceptual modelling bars is higher in the first half of the protocol. This suggests that for SD modellers conceptual modelling is an important modelling task that anticipates coding of the model on the computer. Model coding appears at various times throughout the protocol for SD1, SD2 and SD3, while for SD4 concentration on model coding appears in two parts. Conceptual modelling is attended to in the first half of the protocol, followed by model coding and then attention turns back to conceptual modelling and then a concentrated attention to model coding follows in the second half of the protocol. SD5 attends to model coding only at the end of the protocol, after being prompted by the researcher. SD5 was reluctant to create a model on the computer commenting that it does not involve great modelling capabilities, *"feeding the model into the computer [coding the model] is more or less a mechanical task"* (SD5). In most SD protocols, data inputs is attended to in the first half of the task. Verification & validation is attended to less and most of the time it follows model coding and data inputs or conceptual modelling and model coding. Results & experimentation is mostly attended to at the end of the protocol, with the exception of SD5, who attended to it in the middle of the protocol. Implementation was attended to by only 3 modellers and this was mainly at the end of the protocol.

Let us consider how the iterative modelling process is visually represented in the DES and SD literature respectively (Figures 6.9 and 6.10). Sterman (2000) represents the stages of the simulation model in an iterative cycle (Figure 6-9), where the star in the centre of the diagram implies the interconnections between the different steps. In a diagrammatic representation of the DES modelling process (Figure 6-10), a similarly iterative cycle between the different modelling steps is represented. The main difference that can be identified is related to the stage of verification and validation, which is positioned in the centre of Figure 6-10. Robinson (2004) points out that during a DES modelling study at least one aspect of verification or validation is performed in parallel to any of the other steps of a simulation study. The two graphs represent DES and SD modelling as iterative modelling processes. Indeed the timeline plots (Figures 6-7 and 6-8).suggest that there is iteration in the model building process followed by both DES and SD modellers. However, different patterns of iteration are observed for the two groups of modellers, where the SD modellers' attention seems to be more scattered among modelling topics compared to the DES modellers, whose attention progresses in a more linear pattern. The pattern of iterations will be further explored in the next section.

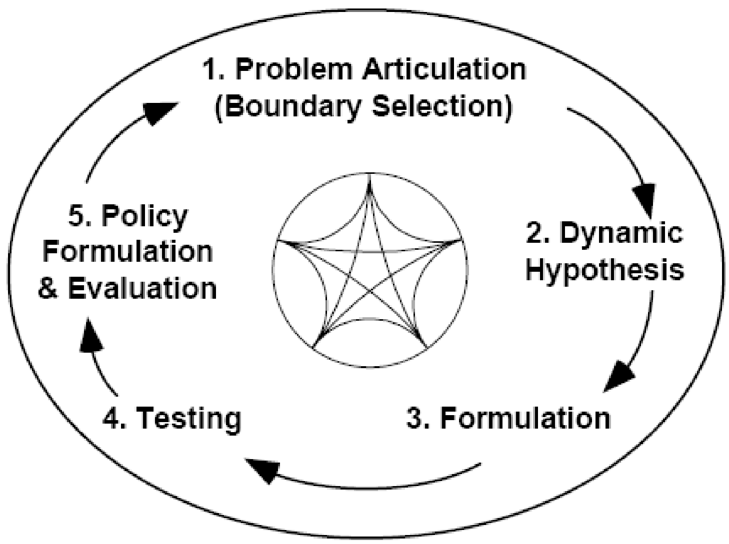


Figure 6-9: SD modelling as an iterative process (Sterman, 2000, pp. 87)

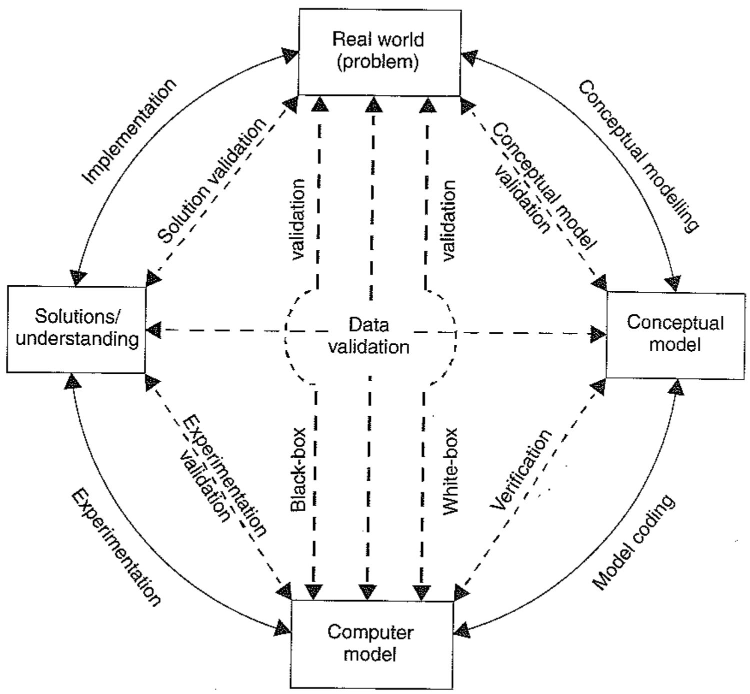


Figure 6-10: Key stages in DES modelling, as an iterative process, with verification & validation in the centre (Robinson, 2004).

6.4.1 Transition matrices

In this section the transition matrices for the DES and SD protocols are explored. The transition matrices are used with a view to summarise and to further understand the differences observed in the DES and SD modellers' timeline plots. A transition matrix represents the cross-tabulation of the sequence of attention between successive pairs of episodes in a protocol. The total number of transitions occurring in the combined DES and SD protocols is displayed (Figure 6-11). It is observed that DES and SD modellers switched their attention from one topic to another almost to the same extent (505 times for DES modellers and 507 times for SD modellers).

Transition matrix - Total DES

	PS	CM	MC	DI	V&V	R&E	Impl	Totals
PS	0	1	2	4	0	3	0	10
CM	2	0	42	21	6	6	0	77
MC	1	37	0	75	62	10	0	185
DI	0	21	70	0	15	2	0	108
V&V	2	11	58	9	0	10	0	90
R&E	3	4	12	5	7	0	2	33
Impl	1	0	1	0	0	0	0	2
Totals	9	74	185	114	90	31	2	505

Topic transition - Total SD

	PS	CM	MC	DI	V&V	R&E	Impl	Totals
PS	0	8	1	4	0	8	0	21
CM	6	0	44	41	7	16	2	116
MC	4	37	0	46	22	24	2	135
DI	3	40	50	0	8	10	0	111
V&V	1	10	17	11	0	9	0	48
R&E	5	21	21	13	11	0	1	72
Impl	0	1	2	0	0	1	0	4
Totals	19	117	135	115	48	68	5	507

Yellow	0
Light cyan	1--10
Light blue	11--20
Purple	21--40
Dark blue	40+

Figure 6-11: Comparative view of the transition matrices for the combined DES and SD protocols, where each cell has been colour-coded depending on the number of transitions.

In order to explore the dominance of modellers' thinking, the cells in the transition matrices have been highlighted according to the number of transitions counted. The darkest colours, in this case purple and dark blue represent the transitions that occur most frequently. Observations of the DES modellers' transition matrix consist of the following:

- Model coding is the topic DES modellers return to most often (185), especially after data inputs, verification & validation and conceptual modelling.
- Attention on conceptual modelling is mostly followed by model coding and to a smaller extent by data inputs.

- Model coding is partly followed by data inputs, verification & validation and to a smaller extent by conceptual modelling.
- Data inputs is mostly followed by model coding.
- DES modellers return to verification and validation after thinking about model coding and data inputs.
- The loop of modelling topics that DES modellers alternate mostly includes: conceptual modelling, model coding, data inputs and verification & validation. This is shown by the blue highlighted cells in the DES protocols table of topic transitions (Figure 6-11).
- The dominant loops in DES thinking consist of transitions between model coding and data inputs. This is shown by the highest numbers of transitions, 75 between model coding and data inputs and 70 for the opposite (transitions from data inputs to model coding).

On the other hand, observations for the SD modellers' transition matrix (Figure 6-11) consist of the following:

- Model coding is the topic SD modellers return to most often (135), followed closely by conceptual modelling (117) and data inputs (115).
- SD modellers' attention on conceptual modelling is followed by model coding or data inputs.
- Attention on model coding is mostly followed by data inputs.
- Whereas attention on data inputs is mostly followed by model coding and conceptual modelling.

- SD modellers return to verification & validation after attending to model coding and results & experimentation.
- SD modellers alternate mostly in a loop between conceptual modelling, model coding and data inputs. These transitions determine the dominant loops in their thinking process. The pattern is captured by the blue highlighted cells in the bottom table of the SD topic transition matrix (Figure 6-11).
- Compared to the DES transition matrix, the values in the SD matrix are less extreme, suggesting that in SD modellers' dominant thinking loop, the transitions are more spread among the modelling topics.

Overall, comparing the two transition matrices, the pattern of the transitions for SD modellers' attention follows a more horizontal progression, while in the case of DES modellers this pattern follows a diagonal progression towards the right-hand bottom end of the matrix. DES modellers' loops of thinking suggest that they switch their thoughts from conceptual modelling to model coding and then to data inputs or verification & validation, returning back to model coding. SD modellers follow a similar loop, with the difference that verification & validation is not included in the dominant loop. Therefore, it can be concluded that an indication of a more linear progression from topic to topic is observed in the DES transition matrix compared to the SD one.

This indication of linearity is further verified using the total number of transitions of attention for the parallel linear strips in the two transition matrices (DES and SD). The cells in the transition matrices in figure 6-9 have been colour-coded, where one cell follows the next cell down on the right (Figure 6-12). So 8 linear strips with different colours have been created, for which the total number of transitions is counted in table 6-6. In the case of absolute linear thinking, it would be expected that all transitions would be concentrated in the blue strip. However, this is not the case with any of the DES or SD protocols. Nevertheless, the total number of transitions for all 8 linear strips provides an indication of the extent of linearity involved. The further away a strip is from the central strip, the fewer transitions are expected in order to convey linearity. The total numbers of transitions are compared for the DES and SD protocols (Table 6-6).

Transition matrix - Total DES

	PS	CM	MC	DI	V&V	R&E	Impl
PS	0	1	2	4	0	3	0
CM	2	0	42	21	6	6	0
MC	1	37	0	75	62	10	0
DI	0	21	70	0	15	2	0
V&V	2	11	58	9	0	10	0
R&E	3	4	12	5	7	0	2
Impl	1	0	1	0	0	0	0
Totals	9	74	185	114	90	31	2

Topic transition - Total SD

	PS	CM	MC	DI	V&V	R&E	Impl
PS	0	8	1	4	0	8	0
CM	6	0	44	41	7	16	2
MC	4	37	0	46	22	24	2
DI	3	40	50	0	8	10	0
V&V	1	10	17	11	0	9	0
R&E	5	21	21	13	11	0	1
Impl	0	1	2	0	0	1	0
Totals	19	117	135	115	48	68	5

Figure 6-12: Colour code linear strips of DES and SD transition matrices

Table 6-6: Total number of transitions per linear strip in the DES and SD transition matrices

Colour code	DES	SD
	145	116
	125	116
	87	74
	85	74
	20	**35**
	23	**34**
	6	**18**
	7	**24**

Comparing the total numbers of transitions for the 8 linear strips of the DES and SD modellers (Table 6-6), it is observed that for DES modellers the higher numbers are at the top of the table, representing the most central strips (Figure 6-12). This implies that the transition of attention for the DES modellers focuses mainly in the most central strips in the matrix. Whereas, for SD modellers it is observed that the higher totals are found at the bottom of the table, representing the furthest away from the centre strips (Figure 6-12). The highest totals in the further away strips suggest that the SD modellers switched their attention in a more vertical pattern compared to the DES modellers. It is noted that an equal total number of transitions among topics has been found for DES and SD modellers. It is therefore, concluded that DES modellers’ attention progresses relatively more linearly among modelling topics compared to that of SD modellers.

6.5 Summary of the findings from the quantitative analysis

In the current chapter a quantitative analysis of the 10 DES and SD protocols is carried out. A summary of the findings follows:

1. From the comparison of DES and SD modellers attention to various modelling topics (6.3) it is found that:
   - DES modellers focus significantly more on model coding and verification & validation of the model.
   - Whereas SD modellers concentrate their attention mostly on conceptual modelling.
   - DES and SD modellers spend almost the same level of attention on problem structuring, data inputs, results & experimentation and implementation.

2. From the analysis of the DES and SD timeline plots and transition matrices (section 6.4), it is concluded that DES modellers progress more linearly through modelling topics compared to SD modellers. Some further observations follow:
   - The dominant loops in DES thinking consist of transitions between model coding and data inputs. In SD thinking the dominant loops consist of transitions between conceptual modelling, model coding and data inputs.
   - Considering the number of transitions in the DES and SD modellers' dominant loop of thinking, the values in the SD matrix are less extreme and more evenly distributed among modelling topics.

Problem structuring

- When starting a simulation task, DES and SD modellers do not necessarily start by thinking about the objectives of the modelling task. Comparatively, SD modellers have the tendency to return to problem structuring throughout the modelling task, whereas DES modellers refer to it more at the beginning of a modelling task.

Conceptual modelling

- Attention to conceptual modelling is scattered throughout the modelling task for DES modellers, while for SD modellers' attention to conceptual modelling is higher in the first or middle stages of the task.

Model coding

- Model coding is the topic to which DES and SD modellers return to most often. However, a higher number of transitions was observed for DES modellers (185 compared to 135 transitions for SD modellers).
- After thinking about model coding, DES modellers focus on entering the data inputs and on checking that the model is correct, whereas SD modellers switch their attention to various topics, but mostly to data inputs.

Data inputs

- Data inputs is a topic well attended to by most DES and SD modellers, mainly in the first half of the modelling task.
- Data inputs complements model coding in DES modelling, while in SD modelling it complements model coding and conceptual modelling.

Verification & validation

- DES modellers are more interested in verifying and validating the model code and data inputs, while SD modellers validate mostly the model code and the results of the model.

Results & experimentation

- Almost all DES and SD modellers shifted their attention to results & experimentation towards the end of the modelling task, apart from DES3 and SD5[11].
- DES modellers consider the results and the scenarios to be tested in the model after creating and verifying the model code, while SD modellers do so after conceptualising and coding the model.

Implementation

- The attention paid to implementation is significantly low compared to the previous modelling topics. If the focus of the exercise was different,

[11] DES3 and SD5 did not develop a complete model on the computer and therefore references to results and experimentation do not occur in the respective protocols. The references made by DES3 at the beginning of the exercise and by SD5 in the middle of the exercise, represent mainly thoughts made about the results of interest or the scenarios to be built in the model.

then more relevant conclusions could be drawn with respect to this modelling topic.

3. The modelling process followed by DES and SD modellers is an iterative one. However the cyclicality of thinking is more obvious in the SD protocols (section 6.4.1).

   - DES modellers start most often with conceptual modelling or model coding, where the parts to be included in the model are considered and modelled at the same time. In the first half of the modelling task, a first model is created, where modellers alternate mostly between conceptual modelling, model coding and data inputs. In the second half of the modelling task, the model code is tested in conjunction with verification & validation. In addition, results & experimentation, though not often referred to, is considered towards the end of the modelling task.

   - SD modellers start most often with conceptual modelling and problem structuring. A first model is created in the first half of the model, where modellers alternated most often between model coding, data inputs and results and experimentation. In the second half of the modelling task, more concepts are added into the conceptual model and computer model, where modellers' thoughts alternate between conceptual modelling, model coding and verification & validation. The few implementation episodes are considered towards the end of the modelling task.

These findings are further discussed in support or otherwise of hypotheses 1.1 - 1.3, in section 8.2.

Chapter 7: Qualitative analysis of VPA data

7.1 Introduction

In this chapter a qualitative analysis of the 10 verbal protocols with respect to the second research objective is presented. According to this objective, the aim is to establish the differences and similarities in the modelling approach taken by DES and SD modellers in each stage of simulation modelling. The analysis consists mainly of text analysis of the transcribed protocols. The chapter is built around the 7 modelling topics, related to the research hypotheses derived from the literature (section 3.4.2). A separate description of DES and SD modellers' thoughts relevant to each testable hypothesis is initially provided, which is then followed by a summary of the comparison of the two groups of modellers. Some additional observations relevant to the comparison of DES and SD model building have emerged, which are discussed under the relevant modelling topics.

7.2 Problem structuring

Problem structuring is related to the thoughts expressed by participant modellers regarding their understanding of the problem and the objectives of the modelling exercise. This is linked to research hypothesis 2.1 (section 3.4.2):

DES modellers on problem structuring

Only some DES modellers (3 out of 5) attend to problem structuring. A possible explanation for this could be that the problem described in the case study was perceived as well-defined by DES modellers and, therefore, they did not consider it necessary to further explore the objectives. DES modellers think about the objectives of the model in direct connection with the outputs of interest.

> "*...need to look at, find out what the actual results need to be.*" (DES1, episode 8)

The output of interest for DES modellers is in most cases the size of the prison population. DES1 also considers the occupancy of prison facilities. Most DES modellers think about the model as a tool for estimation.

> *"...want to look at the amount [number of people] in prison ..."* (DES1, episode 8)
>
> *"...we are looking at occupancy*" (DES1, episode 11)
>
> "*...all about the throughput of prisoners in the prison system..."* (DES3, episode 1)

Rarely do DES modellers relate the objectives of the model to the comparison of scenarios for various policies. Indeed only DES3 (out of the 3 DES modellers, who attend to problem structuring) refers to the comparison of scenarios for different policies.

There is a general feeling that the problem situation is more or less clear to most DES modellers and, therefore, only 2 (out of 5) DES modellers attempt to understand the problem and the issues involved. However, these do not involve an in-depth consideration of the issues involved.

> *"How I understand the situation."* (DES5, episode 2)

SD modellers on problem structuring

Almost all SD modellers think about the primary objective of creating a model to 'estimate' or 'project' the size of the prison population.

> "*...forecasting how it [prison population number] is going to proceed in the future according to the input decisions"* (SD2, episode 3)
>
> "*...estimate of what happens to the prison population*" (SD5, episode 29)

However, SD modellers take a wider perspective with respect to the objectives of the model. This becomes clear from the model objectives mentioned throughout the protocols, such as: reduction of criminal acts, solving the problem of prison overcrowding, the total cost incurred by the prison system to the society, etc. The objectives mentioned in the SD protocols, consider the issues involved not only inside prisons, but they also consider the wider consequences to the society or community. Therefore, it can be concluded that SD modellers consider that the issue lies beyond projecting the size of the prison population and solving the problem of prison overcrowding. The fact that SD modellers consider the interlinked objectives of reducing the incidence of crime and the safety of the community conveys a broader perception of the problem.

> "*It highlights the difference between solving the prison overcrowding population problem and actually what we want to do is reduce crime*". (SD1, episode 142)
>
> "*We could track the serious prisoner released in the community…"* (SD2, episode 108)

Testing scenarios or strategies comes as an inherent objective to SD modellers. Similarly, their thoughts about testing various scenarios to solve the problem of prison overcrowding look also into the wider implications of a specific scenario.

> *"… I would want to explore the extent to which particularly some of the solutions are about the community.*" (SD1, episode 49)

The objectives of the model are related to scenario testing and studying their implications. Some SD modellers (2) imply that the scope of the model is determined by the policies or scenarios considered.

> "*Those strategies themselves sometimes influence the way with which you build the model*." (SD1, episode 52)

Comparative view of DES and SD modellers' perspective on problem structuring

Some overall comments on DES and SD modellers' thoughts related to problem structuring for a model of the prison population follow.

- DES modellers think about problem structuring to a smaller extent. 2 out of 5 DES modellers (DES2 and DES4) do not mention the objectives of the problem during the model building exercise.
- DES and SD modellers think about the model as a tool that projects or estimates the output of interest, i.e. the total number of people in prison.
- Most SD modellers take a broader scope of the problem, looking into further issues that cause prison overcrowding. 3 out of 5 SD modellers (SD1, SD2 and SD5) take a more holistic approach towards the problem.
- SD modellers relate the objectives of the model with testing policies, while DES modellers rarely bring up the idea of policies as part of problem structuring. They most often consider problem structuring in conjunction with the outputs of interest.

7.3 Conceptual modelling

Quotes on Conceptual Modelling deal with the process of designing or thinking about the conceptual model. The aspects analysed here are related to research hypotheses 2.3, 2.4 and 2.5 (section 3.4.2) and these include DES and SD modellers' views related to: conceptual diagrams, feedback and systems' view. These will be discussed next under the respective headings.

7.3.1 Diagrams

The current sub-section considers the diagrams built by DES and SD modellers, and the relevant views expressed during the protocols.

DES modellers on diagrams

Only 2 DES modellers considered drawing a conceptual diagram. DES1 drew a very basic conceptual model, which did not resemble a specific DES diagramming method, but was in the form of notes. DES1 also noted down the numbers and attributes of the individuals in the system. In addition, DES1 made some notes regarding the variables that could be varied in relation with the possible scenarios.

Compared to all DES and SD modellers, DES3 thought more about drawing the conceptual diagram, who dedicated 8.3% of the script on conceptual modelling to creating the conceptual diagram. DES3 considered using either an activity cycle diagram, which he called "*state change diagram*" or a process flow diagram, mentioned as "*a series of processes and queues*", and decided to use a process flow diagram. He also expressed the opinion that drawing diagrams helps the modeller to understand the problem situation to be modelled and it eases the model building process.

"*...it's always good to draw these diagrams and I think ... to make things clear in your mind."* (DES3, episode 181)

"...makes it easy really". (DES3, episode 185)

However, 3 out of 5 DES modellers (DES2, DES4 and DES5), did not draw a conceptual diagram on paper, but DES4 mentioned that he would prefer to do so if he had more time available. According to DES4, drawing the conceptual diagram helps in finding "…*suitable names to give my parameters and variables…*" DES5 was keen to create the structure of the model while coding the model on the computer:

"*So what I am going to do here is just build the basic flow of the model"* (DES5, episode 25)

"*I think the first thing I am going to do is try to build the process up*" (DES5, episode 30).

SD modellers on diagrams

The approach of SD modellers to diagrams differed. Some SD modellers (SD1, SD2 and SD5) built a conceptual diagram on paper before building a model on the computer. The diagrams created consist mainly of a list of the basic stocks (i.e. petty criminals, serious criminals, jail capacity and recidivists) to be included in the model. They do not resemble a standard SD diagram, apart from that created by SD5, who created a stock and flow diagram on paper first and pointed out that the conceptual model is the most important stage of the modelling task. Drawing a conceptual model is considered as a means that helps the modeller to think about the main issues to be involved in the model. He was even reluctant to create a simulation model, mentioning that all the conceptual issues need to be resolved first before building a model on the computer.

Most SD modellers conceptualised at the same time as building the main stocks and flows on the computer. Furthermore, some modellers (SD3 and SD4) did not create a conceptual diagram, despite the fact that they might do so in practice:

> "*I would normally sketch it on a piece of paper first, but I might just try building it straight away in Powersim and see where I get to*." (SD3, episode 1)

The main reason for this could be the limited time available. On the other hand, SD4 did not refer to conceptual diagrams in his verbalisation. Considering his long experience in SD modelling, it could be suggested that given the simplicity of the exercise, he did not feel the need to create a conceptual diagram before building the model on the computer. It could also be suggested that this is related to the type of software used. Some software, such as Stella/iThink, seemed to be easier to use and it is probably more convenient to develop a conceptual diagram straight on the computer.

Comparative view on DES and SD diagrams

Some overall conclusions regarding DES and SD diagramming methods follow:

- Overall, the creation of diagrams was not a priority for most DES and SD modellers. Due to the nature of the task, the participants were provided with a basic diagram in order to help their structured thinking. Therefore, while the modellers had already been given a basic conceptual diagram of the model, one can observe a more or less similar reaction by most DES and SD modellers.
- Some modellers (DES2, DES5 and SD4) did not mention the word 'diagram' at all while some (DES1, DES4, SD1, SD2 and SD3) thought only briefly about drawing a conceptual model on paper first. Only DES3 and SD5 thought more seriously about creating a conceptual diagram. SD5 considered drawing a diagram of the conceptual model as the most important task in a modelling exercise.
- The diagrams created generally do not resemble any formal DES- or SD-like conceptual diagrams. The exceptions were DES3 (who created a DES process flow diagram) and SD5 (who created a stock and flow diagram).
- Most DES and SD modellers conceptualise at the same time as coding the model on the computer.

7.3.2 Feedback

In this section the DES and SD scripts have been reviewed with a view to identifying the modellers' thoughts on causal relationships and feedback structures in connection with research hypothesis 2.4 (section 3.4.2). Typical words that give an indication of thoughts on causal relationships are: '*influence', 'impact', 'effect of …on…', etc.* The consideration of causal relationships leads subsequently to the identification of feedback structures in the model.

DES modellers on feedback

Reviewing the DES scripts, no evidence is found to reveal that the DES modellers considered any relevant causal relationships and feedback structures present in the prison population system. This does not imply that the DES models created do not include causal relationships and feedback effects (i.e. the return of recidivists in prison), but rather that the DES modellers were not interested in identifying or were not explicitly aware of the existence of feedback effects. They were more interested in laying out the flow of processes in the model without explicitly considering the effects between the processes or parts of the model.

SD modellers on feedback

In SD modelling feedback is considered an important part of the model structure. During the modelling sessions it was observed that most SD modellers (SD1, SD2, SD3 and SD5) built the conceptual model by considering the feedback effects in the system. The actual word 'feedback' has been mentioned by only 2 modellers (SD1 and SD3). However, it can be observed that SD2 and SD5 considered the feedback effects present in the model. SD3 and SD5 also suggested further feedback effects, such as the deterrence effect of sentencing or offending that could be further introduced into the model. Only SD4 did not make any references to feedback effects.

Some examples of causal relationships and feedback effects mentioned by the SD modellers are:

- Effect of recidivism on the number of people in prison (SD1 and SD3)

  "…if we reduced the re-offending rate, then what impact would that have on the prison population." (SD3, episode 82)

- Effect of sentence length on people in prison (SD3)

  "…to extend the sentencing what implication would that have on the prison population?" (SD3, episode 95)

- Effect of sentence length on re-offending (SD1, and SD3)

  "*…play with the sentencing, so petty criminals stay for a shorter period, which would have **feedback** on re-offending…"* (SD1, episodes 147 and 148)

"*... if I increase the sentence length, what would the impact be on re-offending...?*" (SD3, episode 98)

- The deterrent effect of imprisonment on commitment of crime and on re-offending (SD2, SD3):

  "You could think about ... if the sentences were longer then maybe there is some sort of deterrence effect, that would mean that the re-offending would be lower..." (SD3, episode 103)

- Effect of money spent by the society on the degree of overcrowding (SD5)

  "...amount of money being spent by society ... see the effect on the degree of overcrowding..." (SD5, episode 72)

- Impact of degree of overcrowding on offending or re-offending (SD2 & SD5)

  "...the degree of overcrowding which has an effect on the amount of crime committed and the population ..." (SD5, episode 72)

  "*...overcrowding will increase recidivism...*" (SD2, episode 96)

  "*...degree of overcrowding ...has an impact on the recidivism rate*" (SD5, episode 70).

Comparative view of DES and SD modellers on feedback

While conceptualising and experimenting with the model, most SD modellers naturally thought about the causal relationships and the feedback effects present in the model. The feedback effects were less often identified, however, most SD modellers considered the effect of a variable on another, which is the basis that leads to the identification of feedback effects. On the other hand, the DES modellers did not consider any equivalent structures. While conceptualising, DES modellers were most interested in setting up the sequential flow of events in the model rather than considering the effects among them.

7.3.3 Systems' thinking

Modellers' thoughts on systems' thinking are explored considering the approach taken during modelling, with respect to hypothesis 2.5 (section 3.4.2). Two extreme concepts with regards to systems' thinking are distinguished: analytic and holistic. In this context, analytic thinking is viewed as an approach taken where attention is focused on the elements, which are considered to be independent of the context. In holistic thinking, elements are perceived to be interconnected, focusing on the context and the interrelationships between the elements and the context (Checkland, 1999). Related to this distinction, systems' thinking is explored looking at the perception of modellers about the model and parts of it. When looking for holistic systems' thinking, the use of the word 'system' is identified. The text was also analysed looking for cues that identify instances of thoughts related to 'system archetypes' (Senge, 1990). System archetypes consist of a way of thinking which includes insights from and understanding of the wider environment within which a system exists. Senge (1990) suggests that when one thinks in terms of systems archetypes, then systems thinking is actively used to understand and reveal the reality.

"*Learning to see the structures within which we operate begins a process of freeing ourselves from previously unseen forces and ultimately mastering the ability to work with them and change them*" (Senge, 1990, p. 64).

DES modellers on systems' view

In the DES protocols a rather analytic perspective of the prison population system is identifiable. It is evident that DES modellers conceptualised the prison model by thinking about representing specific activities/processes and the individuals or so-called entities part of the system, and their detailed characteristics in the model. This is established from the way they progressed during the modelling task. Each event or part of the system was modelled one after the other, without considerations of the inter-relationships between them. As has already been established as part of the discussion on problem structuring (section 7.2), the DES modellers did not consider the wider environment in which the prison population exists, nor did they think about other issues that might be related to the prison system.

> "*…so we need some sort of process for re-offending, which needs to take place every year.*" (DES3, episode 65)

> "*...we also need to think about the jailing that is going to take place throughout the duration of the model." (*DES3, episode 116)
>
> "*...I am now going to create a machine called M arrivals ..."* – (DES5, episode 59)
>
> *"...what we need to think about going through the system is, the actual entities themselves."* (DES3, episode 14)
>
> "*Now obviously each prisoner is different, each prisoner is going to have a series of characteristics or attributes or labels ..."* (DES3, episode 15)

Most DES modellers extensively used the word 'system', but this was not related to systems' thinking. The main context it was used in is an analytic one, related to a prison population system made up of a set of individuals and activities. Some examples are: "*joins the system/ coming into the system"* (DES1), "*people in the system*" (DES1)*, "... [prisoners] go through the system"* (DES3), *"... we have parts entering into the system*" (DES5), etc.

SD modellers on systems' view

In the group of SD modellers, a more holistic approach is observed compared to the DES modellers. A systems' thinking perspective is revealed by some SD modellers, who thought about the environment or the context the prison model is part of. Particularly, SD1 and SD4, who use iThink, thought about the prison model in connection to the environment or system it is part of. They considered the boundaries to the prison system, by separating the prison sub-system from the community or society, but overall considered it as part of the larger society system.

> "*I need to think about whether the prison is in effect a sector within a larger model.*" (SD1, episode 14)
>
> *"…they [recidivists] are positioned outside of prison because they are not part of prison population...*" (SD4, episode 121)

Considerations were also made with respect to the effect that solutions to the prison population problem can have on the wider system (society). For example, SD5 expressed an interest to know the impact of various variables on society. This is an indication that SD5 looked into the problem from a broader perspective, referring to the effect that the prison model has on society. Additionally, SD5 explicitly mentioned that he is looking at the bigger picture.

> *"… the problem is what happens to the society, which is better for society"* (SD5, episode 29)

Another indication of holistic thinking is the fact that most SD modellers thought about the interrelationships between variables in the system. For example SD3, referred to the existence of interconnectedness in the model when considering the effects of scenarios in the model. This has already been discussed in section 7.2.

Furthermore, the SD modellers expressed an interest in extending the model with the introduction of new variables or effects in the model, which were not included in the original conceptual model provided. Such is the case of the variables: money spent by society (SD5), damage to society (SD5), deterrent effect (SD2, SD3 and SD5), death rates (SD4), etc. These examples suggest that the SD modellers thought outside the box, considering the effects that changes to the prison model can have on society.

Furthermore, most SD modellers mentioned the word 'system' or equivalent expressions. SD1 mentioned the word 'system' 8 times whereas SD4 did not mention it at all, apart from referring to the word 'sector' (a term used in iThink) 3 times. SD2 referred to the word 'system' only while thinking about the overall system's objectives:

> *"... it looks like the aim is ... the reduction of inflow into the prison system*" (SD2, episode 1)

SD3 used the word 'system' on 7 occasions, mainly in the context of prisoners, entering, leaving or being in the system. SD5 did not mention the word 'system' during the modelling task, the words model or prison, were used to convey a similar concept, i.e.:

> *"... people who are coming into the prison population"* (SD5, episode 16)

Another strategy followed to identify the existence of systems' thinking in this research is to look for references to some existing system archetypes as suggested by Senge (1990). Some relevant system structures that would be relevant to the UK prison population would be the ‘balancing system’ or the ‘shifting the burden’ archetypes. Only SD1 made some comments relevant to analysing system archetypes. This suggests that not all the SD modellers thought about system archetypes and considered the deeper reasons why some effects occur. This is nevertheless an indication that some SD modellers tend to get beyond the structure of the model when discussing the issues involved in a problem situation.

Referring to the ‘balancing system’ archetype, a balanced system is portrayed bearing in mind the existence of feedback in the system when taking corrective action. Some evidence of referring to the balancing system archetype was found in SD1’s protocol. For example, he was aware of the existence of a delay in the system:

> *“…What I would expect the model to tell me is, because we are going to have a delay in the system.”* (SD1, episode 56)
>
> *“…It will be the time delay that affects that [the variable petty rehabilitated].”* (SD1, episode 85)

The ‘shifting the burden’ archetype is also highlighted in the case study description, where previous actions of increasing prison sentences are mentioned and the subsequent increase in the numbers in prison, resulting in overcrowded prisons and the related problems. For example, SD1 talked about the counter-intuitive behaviour of the model, referring to the issues highlighted in the case-study that “prisons serve as a school for crime”, suggesting that by serving time in prison, petty prisoners become better criminals (converting from petty to serious prisoners). Therefore, SD1 implied that changing the sentence length is probably not a long-term solution to the problem, and made considerations about the longer term factors that would solve the situation:

> *"I think what the model highlighted, is that you need to make it less of a school for crime. What that means is that even for people who are staying for a relatively short period they need to get serious about what they have done, etc."* (SD1, episode 151)

Comparative perspective of DES and SD modellers on systems' thinking

From the observations made regarding DES and SD modellers' approach towards system's thinking, differences are found. These are:

- The DES modellers go through an analytic thinking process when modelling the prison population model. Furthermore, they did not consider the wider issues involved and their inclusion in the model. The DES modellers' attention focused mostly on the individual parts of the model, without considering the wider environmental or social factors that affect the prison system.

- In contrast, different aspects of systems' thinking were identified in the SD protocols. The SD modellers thought about the context and the wider environment the prison model is part of. They also made considerations about the interrelationships between variables as well as the effects of prison on society, which does not occur in the DES protocols.

- Not all the SD modellers conceptualised ‘calling’ system archetypes, however the fact that one SD modeller considered them, reveals some level of systems’ thinking involved. Whereas, a lack of mentioning system archetypes is observed in all DES protocols.

7.4 Model coding

Verbalisations on model coding include articulated thoughts in relation to building the model on the computer. These are analysed from the perspective of the level of detail and complexity involved in the model (Hypothesis 2.7), the use of modelling structures (hypothesis 2.8) and modelling of delays (hypothesis 2.10). These are analysed in the following sub-sections.

7.4.1 Level of detail & complexity

In this sub-section, model coding protocols are reviewed in order to identify the level of detail and complexity involved in the modellers’ thinking process. Hence in these episodes evidence is sought to identify a tendency to add further detail in the model or to broaden the scope already provided. The protocols are also examined to identify expressions of unhappiness (or the opposite) with the scope provided as part of the case study description. Some examples of references to detailed information are considerations of: entities, attributes, how processes work, data required, etc.

DES modellers on level of detail & complexity

It is observed that DES modellers were overall happy with the data and the structure of the model given in the description of the exercise. Most DES modellers followed this structure while building the model. Only DES4 considered a slightly different structure to the one already provided, by changing the routing of serious offenders to include the re-offence as petty. DES2 changed the structure of the model, by making the prison a common facility that housed both petty and serious criminals and used attributes to track the type offender (petty or serious). Other modellers considered changes in the model, mainly in the form of adding further detail. For example, DES4 considered extending the model by adding detail to it such as: including the number of prisoners coming from different geographic regions, the cost of living in different regions, transportation costs, etc., with the intention of optimising the system or minimising the costs to the prison system.

Another aspect of detailed thinking is conveyed by the fact that DES modellers think about the individual characteristics of each prisoner. They show an interest to give each prisoner characteristics such as: type of sentence, sentence length (or sentence date and date entering prison) and a re-offending probability. Depending on the software used, attributes (DES1 and DES5) or labels (DES2, DES3 and

DES4) are used to distinguish petty and serious offenders. Some examples of how the labels work and their role in the model follow:

> *"…will put some labels on them. So labels, I am going to have one label called 'offender type' and it is going to be set to, so petty = 1 and serious = 2"* (DES2, episode 19)
>
> *"… tag those ones [offenders] which are re-offending…"* (DES3, episode 64)
>
> *"…stamp the sentence on them [prisoners] …"* (DES5, episode 62).

In addition, some DES modellers express a preference to using attributes or labels. This serves as an indication of a detailed way of thinking. It is revealed that using attributes or labels is further related to model verification.

> "*I like stamping attributes when I build models. I like to stamp attributes so then you could trace people through the system, if you need to debug it.*" (DES5, episode 154)
>
> "*… every time a prisoner turns up at the door I will stamp them with this and this is how long you will be in prison for …*" (DES5, episode 71).

Most DES modellers expressed a preference in representing each single individual prisoner. However, due to the complexity arising because of the large numbers in the system, most DES modellers considered aggregating the numbers. This was achieved by scaling down the numbers by a fraction of 10 or 100. For example a population size of 76,000 prisoners becomes 760. There were, however, cases where DES modellers expressed unhappiness with scaling down the numbers. Such is the case of DES1, who felt that this would undermine the accuracy of the model:

"*Could take the reduced size of it [meaning the 76,000 total prisoners and 50,000 petty prisoners], to 7,600 and 5,000, but I'm not happy about doing that*." (DES1, episode 26)

The choice of the time unit is an indication of the level of detail involved in the model. Most DES modellers (DES1, DES2, DES3 and DES4) used days as a time unit. DES5 set the time unit to be in minutes. A time unit of minutes or days for a model where the time-frame could consist of some tens of years indicates to some extent detailed thinking. The main consideration about choosing a time unit is how it is going to affect model run time:

"*I have just managed to speed up Witness so I think we will be alright in minutes."* (DES5, episode 87)

Interestingly, the DES modellers were aware of the level of detail involved due to the choice of the time units.

"*Because I've done it in minutes, it's a quite a granular model …*" (DES5, episode 282)

With regards to the existence of complexity in the model, DES modellers refer to it sparingly. For example, DES2 referred to the complexity arising in the model due to the use of high volume, where people with similar characteristics are grouped into batches. Another aspect of complexity arising in the system is related to the large population and the time that it takes the model to handle the large population. Most

DES modellers commented on this, however, they found one way or another to deal with it.

- DES1 encountered problems with fitting the amount of people in the prison buffer, because the software (Witness) has a default limit on buffer capacity (Witness has a maximum buffer capacity of 32,767 or 65,535 depending on the release), which cannot be changed. For this reason DES1 decided to reduce the size of the population in the system, by reducing it by a 10^{th}, but expressed unhappiness with doing that (explained above).
- For DES2, the large population affected the speed of the model and used the 'high volume' option, which aggregates entities in Simul8 to deal with the high number of prisoners in the system.
- DES3 scaled down by 100. *"...scale down by, a factor of 100, just for the purpose of putting this together ..."*
- DES5 also encountered a problem with coping with the large prison population.

  "*A limitation of Witness is that we can't have a buffer greater than 10,000 [the exact number is 32,767] ...*" (DES5, episode 120)

  "*I thought Witness had bigger buffer sizes than that but it's not too much of a problem."* (DES5, episode 201)

  Nevertheless, DES5 thought of an original way to deal with this limitation. He tried *"... a different approach ... what I need to do is re-design how I want to do that*" (DES5, episode 124) and solved the problem by giving the buffer *"... an instance [replications] of 10, so basically I have given it* *[buffer prison] a capacity of 100,000 …"* (DES5, episode 126). The same applies when DES5 tried to set up the initial number of people in buffers.

> "*So, this is the problem with Witness. It's a bit awkward when you need to initialise buffers*." (DES5, episode 199)

He decided to create a dummy machine that pulls people from the entry point into prison. This dummy machine *"…at time 0 is … going to cycle through 70,000 people, stamp them with an attribute and shove them in the buffer and it then won't operate again.*" (DES5, episode 204)

Furthermore, by adding coding routines or conditional coding, DES modellers added some level of complexity in the model. They use Actions language in Witness (DES1 and DES5), C++ in Flexim (DES3) or Visual Logic in Simul8 (DES2 and DES4). The purpose of using coding routines is explained with some examples:

- DES1 created functions, consisting of some lines of code to set or change the attributes of people in the system, to count the number of people in prison, etc.
- DES2 used Visual Logic and labels to control the routing of people in the system.
- DES3 also created conditional coding to set the release of prisoners after finishing their sentence.

- DES4 created a conditional-structured code using randomness to determine prisoners who are rehabilitated and the ones who re-offend, after serving their prison sentence.
- Similarly DES5 created conditional coding: *"...whether they re-offend or not and if it's a yes, then what's the re-offence is it serious or petty?"* (DES5, episode 142)

With the addition of conditional coding, more complexity was added into the model. While the calculations are handled inside the model code, the logic becomes more difficult to follow mentally without the aid of the computer. This is nicely pointed out in the following episode:

> "*There's a lot of stuff going on behind the scenes …*" (DES5, episode 284).

Another aspect that adds to complexity in DES models is related to the problems encountered in making the model work or finding ways to represent specific parts of the model. In some cases modellers thought 'outside the box' and found new ways of dealing with the problems encountered. However, in other cases these problems persisted. Some examples follow:

- DES3 finds it difficult to identify the mistake that is causing the model to stop. The problem was not apparent to him and he was unable to find the mistake after a few checks. Therefore, in these situations it is mentioned that it is more productive to discuss it with and ask for the help of a colleague, rather than spend fruitless time working on it.

- DES5 encounters difficulty with the 'Init' actions, because the variables and the code were set up in different environments/folders inside Witness. In this case, DES5 mentions that the software is not flexible and changes are difficult to make:

  "*…once you have set it up at the beginning, you don't have a chance to change that much*" (DES5, episode 85)

- Furthermore, some complexity is involved in the DES software in relation to the different types of menus available. For example, modellers needed to be aware of the functionality of different menus in Witness. Coupled with the fact that the logic of the software is hidden behind the scenes, if one is not careful, mistakes could be made and it would be more difficult to reveal them or to find the reasons for their occurrence.

  "*Witness, it has this order of executions. If I was to put that in actions on finish, it might do it once, whereas actions on Output it might do it 3 times so sometimes you have to think about how Witness would execute that.*" (DES5, episode 242)

SD modellers on level of detail and complexity

Most SD modellers were happy with the scope of the model as set out in the case study description. However, in some instances, some modellers (SD4 and SD5) considered changing the structure of the model. SD4 talked about an alternative way of modelling the prison system which would include two separate processes, that of committing crimes and that of imprisonment. SD4 was also interested in including the population at large, which can evolve into different states or categories of people and the model would be looking into the lifecycle of a prisoner going from one stock or state to the other. This is a somewhat different configuration of the model, where the population at large and recidivists (released re-offenders) are part of the same pool of people. SD5 also considered a different structure of the model regarding people in prison. Recidivists, who have already been to prison before were separated into different stocks from prisoners who go in prison for the first time.

SD modellers also added further factors in the model scope or considered to do so as part of further considerations at the end of the exercise. Some examples consisted of adding in the model factors such as: released prisoners in society, the deterrence effect, the inclusion of death rates, general impact on society, recidivist factor, disposition petty and serious, etc. The latter 2 variables represented auxiliary variables that calculate the overall number of released prisoners, before routing them into rehabilitation or the re-offending route, etc.

> *"...overcrowding is going to drive by some sort of measure ... guess it could be Proportion of recidivism. ... I will call it 'recidivist factor'.* (SD2, episode 90)
>
> "*We could track the serious prisoner released in the community, here we've got them going out into the ether...*" (SD2, episode 108)

"*... if the sentences were longer then maybe there is some sort of deterrence effect that would mean that the re-offending would be lower…"* (SD3, episode 103)

"*…I suppose what you might want to introduce would be death rate, so some people die inside."* (SD4, episode 107)

SD modellers were happy to look at aggregate numbers of people in the system. This is evidence of high level thinking. Furthermore, SD modellers were aware of this fact too.

> *"... we are just tracking a mass of petty prisoners, we are not tracking individuals…* (SD3, episode 31)

However, on one occasion, SD4 considered the route that a specific individual takes in the model:

> "*…If you think in terms of an individual [serious prisoner] then I go into prison and then I have 30% chance of coming back into prison. And then I'm released … this fraction of people [serious prisoners] being released, will go back in within 2 years."* (SD4, episode 176)

In addition, the choice of the time units showed a higher level of thinking compared with the case of the DES models. SD1, SD2 and SD5 use years as a time unit, whereas SD3 and SD4 use months.

With respect to the complexity involved in the model, some dynamic complexity was identified resulting from cause and effect relationships, which are not obvious prior to creating the model. For example, when creating the part of the model including people in prison and sentence length, some modellers expected to achieve a model in equilibrium. However, this was not achieved unless the flow of re-offenders was included. This was not clear and straightforward to the modellers.

Furthermore, no particular modelling issues were reported by the SD modellers with respect to building the model or regarding the large size of the prison population. The fact that less attention was paid to model coding by the SD modellers compared to the DES modellers (section 6.3), conveys that fewer problems were encountered and, therefore, less detail complexity was involved in the SD models.

Comparative perspective of the DES and SD modellers on level of detail and complexity

Comparing the DES and SD modellers' thinking related to the level of detail and complexity, some differences are identified. In summary, these are:

- Most DES and SD modellers were happy with the structure and data given in the case study description. However, a tendency to either change the structure of the model (SD4 and SD5) or suggest additions to the model such as new structures or variables (SD1, SD2 and SD3) was common in most

SD scripts. On the contrary, the DES modellers followed the structure given in the case study with higher fidelity and suggested the addition of more detailed parts or information in the model, with regards to potential improvements to the model.

- All the DES modellers thought about creating labels and attributes which were given to each individual in the system. This implies a detailed way of thinking compared to most SD modellers who were happy to look at aggregate numbers or people in the system.
- The DES modellers encountered more problems in making the model work or finding out how to represent specific parts of the model, whereas the SD modellers did not find many difficulties in this respect. They were more relaxed with the use of the software. The DES modellers also commented on the problems encountered with the large population in the prison system, whereas the SD modellers did not raise any relevant issues.
- During the model building process, detailed complexity was obvious in the DES modellers' thinking. The inclusion of conditional coding and specialised functions in the DES model added to the complexity and, therefore, the difficulties encountered during modelling. On the contrary, in the SD protocols no specific indications of detailed complexity were identified.
- Furthermore, some evidence of dynamic complexity was present in the SD protocols. This is evident, in the case when modellers considered the stock of people in prison, which in order to reach an equilibrium level, the flow of

re-offenders needs to be taken into consideration in the model. Due to the simplicity of the model, very few examples of dynamic complexity were present, however, some level of dynamic complexity was identified in the SD protocols, while this was not the case with the DES protocols.

7.4.2 Modelling structures

With regards to hypothesis 2.8 (section 3.4.2), the DES and SD scripts are reviewed looking for evidence where modellers refer to or use already existing modelling structures. Two types of modelling structures are distinguished: theoretical structures, for example the ‘asset stock adjustment’ structure found in the SD literature, and practical structures derived from discussions with the client or with people who are knowledgeable about the system being modelled.

DES modellers on modelling structures

There were no references made to specific modelling structures or to previously built models. In addition, it is observed that some DES modellers, while talking about the people in the system, referred to them as passive entities which are controlled by machines. Conditional coding was often used, where routing of people in the system is controlled by commands written in the form of code. Some examples referring to people in the system follow:

> *“... we need the work centre to pull them out and release them.”* (DES2, episode 27)

"... another work centre that forces them back into prison ..." (DES2, episode 32)

"... a machine here, that can pull from the released queue and put them back into prison ..." (DES3, episode 66)

Some examples of conditional coding are:

"So if offender type is 1, so if it's petty offender then I sample the route to see whether they are going to be rehabilitated or not and if they will re-offend then I use that distribution to decide whether they are serious or petty and then I just need to use that label route, so 1 is going to be rehabilitated and 2 is going to be re-offenders." (DES2, episode 53)

"...So the push logic on there, I need to do this conditionally. So conditional port. So ... so if it's a serious offender then push to port 1, so will go back into the serious queue, or the other possibility is this ratio of turning into a serious offender from a petty offender." (DES3, episode 251)

"...for petty ...we look at the crime, there's 2 rates here, there's a bit [code] whether they re-offend or not and if it's a yes, then what's the re-offence, is it serious or petty?" (DES5, episode 142)

In one case, DES1 used a structure, where the level of overcrowding is determined by comparing the number of people in prison and the target number, to represent the available places in prison. This kind of structure is more often used in SD modelling. However, it is observed that the outcome of this structure, which is the level of overcrowding, was not then fed back into the model as it would be the case in SD. It was instead used as a type of model outcome to show the levels of prison overcrowding at the end of the model run.

SD modellers on modelling structures

System dynamics modelling often refers to already existing structures either from SD modelling or other disciplines. Some structures found in the SD protocols are those borrowed from other disciplines or from previously built models and existing modelling structures. These are next outlined.

To start with, SD1 uses structures borrowed from other disciplines when modelling criminals entering prison from the society. He brings into the picture the interrelated social issues that influence the occurrence of crime and referred to already existing structures from what he calls 'social anthropology':

> *"Within the community there is a propensity to commit crime amongst ..., and one could identify within the community. Just as with an awful lot, i.e. if saying, if people are obese they are more likely to suffer from diabetes, this is where social anthropology comes in. If people are on drugs, then they are more likely to commit crimes. So if you know the number of people who are on drugs then you have an indication, how many are criminals in the population"* (SD1, episode 50)
>
> *"What's happening to the population to make it to have a greater or less propensity to commit crime, etc."* (SD1, episode 51)

Similarly, SD5 referred to economic and psychological structures present in the prison system.

> *"...what is the economic impact and psychological impact of the crimes committed by these 4 categories of people who are outside the prison?"* (SD5, episode 30)

In addition, there were examples of the SD modellers referring to previously built models. For example, SD1 by mistake referred to admissions in hospital instead of admissions in prison. This was an evidence of referring to already existing structures coming from previously completed (healthcare) projects. This also shows that SD1 identified similarities between the prison population model and healthcare models (the aspect of knowledge transfer will be further discussed in section 10.7). SD5 also borrowed the term of diminishing returns on investment from economics. SD3 too, suggested that psychological factors could be added into the model and referred to previously built models including psychological factors, such as "*...models where customers see advertising and tend to buy things..."* or models depicting promotions, etc.

Furthermore, SD4 referred to a previously built model relevant to the prison population model. Of course these models were more complicated and larger, including hundreds of variables and looking at different and more specific aspects of the criminal justice system. However, SD4 made mental references to the structures involved in these previous models. Such is the example of thinking about a different structure of the model, where in an alternative model, SD4 was keen to include the population at large in the model and as a consequence recidivists would flow back into these population pools. Therefore, this demonstrates another example of using previous modelling structures.

An existing SD modelling structure is that of decision making, which was represented as typical decision rules of the actors in the system (Morecroft and Robinson, 2005), where people were considered active agents in the system. In the prison population, no such structures were observed due to the simplicity of the model. However, observations can be made regarding the way people were handled in the model. For example, in the case of SD2 people were referred to as active agents, who decide what they do after being released:

> "*…alternate route, is the one that criminals leave prison, they can decide either to be rehabilitated and join general population and stand the same chance to become criminals again if they are properly rehabilitated or they become recidivists and they are presumably committing crimes out there…"* *(*SD2, episode 29).

Another simple example found where SD modellers referred to existing modelling structures is related to modelling the prisoners entering the system. While in DES, people physically enter prison through an entry point, in SD quite often there is a pool of population already in the model from which then a proportion is taken to represent the admissions in prison. Most SD modellers used the structure suggested in the case study description, where numbers were assigned to the flow of admissions to prison. However, some seemed to be more comfortable with having a structure where this flow was driven by a stock representing the population at large. Some examples that identify this are:

> "*…Then there might be a case of having a pool of the population and a crime rate of serious, rather than just having numbers going in."* (SD3, episode 118)
>
> *"...people who are capable of offending, again you have capable of offending petty crimes and capable of offending serious crimes."* (SD5, episode 4)

Creating a common feedback structure can also be considered an existing SD modelling structure. Some SD modellers (SD2 and SD5) created a feedback structure where the degree of overcrowding is determined by comparing the prison capacity with the existing prison population. The value of the degree of overcrowding was then fed back into the model. SD2 introduced a relationship between the degree of overcrowding and the recidivist factor, i.e. overcrowding affects recidivism or the proportion re-offending. SD5 also suggested a similar structure, where:

> *"… the degree of overcrowding has an effect on the amount of crime committed…"* (SD5, episode 72) or

> *"…the degree of overcrowding should have an impact on the sentencing rate…"* (SD5, episode 68)

SD5 also included in this feedback structure the effect of the interaction between people in prison and re-offending, where the different types of people a prisoner meets in prison influences his/her behaviour after being released:

> *"…degree of overcrowding not only has an impact on sentencing rate but also has an impact on the recidivism rate [equivalent to proportion re-offending]"* (SD5, episode 70).

Comparison of DES and SD modellers on the use of modelling structures

From the evidence reported, it can be claimed that differences exist in DES and SD modellers' approaches to using modelling structures during model building.

- The DES modellers did not refer to any specific modelling structures during modelling, while SD modellers refer to already existing structures, such as structures borrowed from other disciplines, structures borrowed from previously built models or already established SD modelling structures.
- The DES modellers viewed people as passive elements of the model, whereas the SD modellers viewed people as active agents who decide their movement in the model based on the underlying structure.

7.4.3 Representation of delays

The analysis that follows considers DES and SD modellers' references to the representation of delays in the model. This is related to hypothesis 2.10 (section 3.4.2)

DES modellers on delays

In the protocols analyzed, it is observed that all DES modellers were aware of the existence of delays in the model. DES modellers mentioned delays as the time that individuals spend in queues or buffers. For example, DES1 modelled prison as buffers *"time buffers with maximum delay"* and similarly recidivists (released prisoners in society) waiting to re-offend. DES2 used 'storage bins' (buffers) to represent prison and used the sentence length as the minimum wait time. This represents the minimum time that an entity stays in that storage bin. In the same way DES4, when modelling re-offenders considered that they "*…are going around in a 2 year delay loop…"* (DES4, episode 99) or *"…I will put them in a 2 year minimal delay buffer…"* (DES4, episode 56) and for petty offenders entering prison he said they *"…will be going to a buffer to delay them [petty offenders] over 5 years…"* (DES4, episode 14). DES5 referred to delays in the same concept. Sentence time was used as the delay variable.

> *"… shove [send] them [petty prisoners] in a buffer which has a minimum delay time of 5 years…"* (DES5, episode 17) or
> "*I will choose minimum delay time for buffer prison, to represent time that prisoners are kept in prison*" (DES5, episode 78).

SD modellers on delays

It is believed that in SD delays are used to represent randomness (stochastic features) in the model (Coyle, 1985; Brailsford and Hilton, 2001). From the analysis of the protocols, it was found that only two SD modellers (SD1 and SD3) were aware of and talked about the existence of delays in the model. Interestingly, 3 SD modellers (SD2, SD4 and SD5) did not refer at all to the existence of delays in the prison system. SD1 thought about the existence of a delay in the system caused by the time that prisoners take to come out of prison. It should be noted that 4 SD modellers (apart from SD3) modelled delays in the form of a simple 1st order delay, where the variable released prisoners was defined as a fraction of the number of people already in prison divided by the time in prison. Only SD3 modelled delays using the embedded 'pipeline delay' function in Powersim.

> "*…what I want to do is a delay; I'll do it as a simple pipeline delay of the released petty offenders."* (SD3, episode 39)

It is observed that the use of delays was not related in any case to the existence of randomness in the model.

Comparison of DES and SD modellers views on delays

- From the examination of the protocols on the representation of delays in the model, it is identified that all DES modellers were aware of their existence in the model, while not all SD modellers referred to the concept of delays.

- The DES modellers represented delays in the form of entities (individuals), delayed/kept in the buffers or queues for the specified delay (i.e. sentence) time.
- Most SD modellers, modelled delays using a simple first order delay, and only one modeller used the embedded delay functions included in the menu options. However, the use of delays was not thought of as a way of allowing randomness in the system.

7.4.4 Further observations on model coding

Use of parameters to create the user interface

DES modellers showed a tendency to use Excel to present the inputs and outputs to the model. Obviously, this was considered as an add-on, because most of the modellers did not have the time to do this as part of the modelling exercise. However, all DES modellers mentioned that they would build an Excel interface and that they follow this practice when working on client projects. The idea expressed here shows that the model is more or less used as a black box, where clients interact through the inputs and options available in the Excel environment, rather than interacting with the structure of the model.

> *"…make the model itself almost just like a press go, so all the clever stuff is in the model, but our customers all they need to worry about is the simple front end."* (DES1, episode 119)

DES modellers created variables rather than entering hard numbers in the structure of the model. Some of these variables were: re-offending percentages, inter-arrival time of petty and serious, and sentence length. Some DES modellers thought more actively about which parameters should be flexible, to enable the set up of the user interface. DES3 and DES5 mentioned that they would choose the parameters that change in the real system:

> "*...now I could vary this [inter-arrival rate for serious offenders] at a later date. I can put that in as a table ...*" (DES3, episode 140)
>
> "*So if government statistics came out and the numbers have changed, you could still use this model without having to phone me up and ask me to change the numbers which I had hardwired in.*" (DES5, episode 147)

SD modellers also thought about the user interface, apart from SD2 and SD5 who did not refer to it at all. The creation of the interface was more flexible in SD software, where any of the variables, apart from the stocks, can be reflected in the user interface. SD1 and SD4, who used iThink, created sliders in the interface layer. SD3 on the other hand, who used Powersim, decided on the variables to be included in the interface and set them up as permanent because at the end of the simulation run they can be set back to their default value, ready to be changed for the next run. The choice of the variables for the user interface was usually made based on the modellers' considerations about the variables that policy makers would like to change.

"*When I set these to permanent I should be able to just restore permanent variables and install these to their initial default figures*" (SD3, episode 96)

"…who is going to be using the model and what kind of inputs they might want to vary and what kind of numerical scales would make sense to them…" (SD4, episode 41)

Obviously, for SD modellers the interface was an important part of the model building process.

"*… if we are doing a consultancy project … allow the client to enter the policy to see the outcomes. It is actually the major part of the model rather than the actual building of the model."* (SD3, episode 85)

"*…depending on who the model is for, if it is policy makers then you would want a nice interface…"* (SD4, episode 169)

7.5 Data inputs

In this section, the DES and SD protocols are explored to find evidence related to the type of data used (hypothesis 2.11), representation of randomness (hypothesis 2.12) and the type of relationships between variables (hypothesis 2.13). These are analysed in the next subsections.

7.5.1 Type of data

In this section DES and SD protocols are explored, looking for references regarding the type of data modellers used when creating their models. This relates to hypothesis 2.11 (section 3.4.2). Quantitative (in the form of numbers) and qualitative (in the form of graphical displays) data were provided in the case study. Some data (i.e. percentage of re-offenders) were purposefully left out with a view to observing modellers' reactions to missing data and to also avoid leading the modellers to a specific representation (section 4.3.1). The data are categorised into quantitative vs. qualitative data, concrete versus abstract data.

DES modellers on type of data

All DES modellers referred only to the numerical data provided in the case study description. They were happy to use the data given such as: sentence length, number of offenders (petty and serious) entering prison, the number of prisoners already in the system, time that re-offenders stay in society before re-offending. With regards to the arrival of offenders in prison, DES software require data in the form of inter-arrivals or time between arrivals. The DES modellers were happy to make the necessary calculations in order to adapt the numbers given to the format required by the software. For example, DES2 calculates first the number of petty and serious offenders entering per day.

> *"So the inter-arrival rate is 3,000 divided by 365 days for petty offenders and 650 divided by 365 days for serious offenders. So 3,000 divided by 365 is 8.2, so that's 8 petty offenders will arrive every day. 650 divided by 365 is 1.78. So that means that … 1.78 serious offenders [will arrive] per day.*
> (DES2, episodes 7 and 8)

Then DES2 finds the inter-arrival rates, by dividing 1 by the number of people entering per day:

> *"...so if we have 8 arriving in a day, so we have 1 divided by 8 as inter-arrival time. Inter-arrival time of 0.125 and this one is petty offenders and this one is serious, that's 1 divided by 1.78. So petty offenders' inter-arrival time is 0.125 [days] ... and serious offenders is 0.5 [days]..."* (DES2, episodes 15, 16 and 17)

Most DES modellers looked for the missing data on re-offending, which were purposefully left out from the case study description. When establishing that these data were missing, the DES modellers' reactions differed. DES1, DES2 and DES3 requested to be supplied with the numerical data in the form of percentages. Some examples are presented below:

> *"...of the released petty people, do we have a percentage of how many become serious?"* (DES1, episode 63)
>
> *"But we don't have any specific data on the probability of these [probability to re-offend]..."* (DES3, episode 177)

After establishing the numbers DES1, DES2 and DES3 were looking for, the researcher provided them with the data required. On the other hand, DES4 and DES5 did not raise any specific issues regarding the missing data and were happy to make their own guesses or assumptions.

> *"... the number of recidivists, there is ... no numbers here, but we can make an assumption on that [chance of serious offenders re-offending]"* (DES4, episode 49)
>
> *"... it is fair to say at a guess that that [percent re-offend serious] would be a lower percentage"* (DES4, episode 72)
>
> *"So what I am going to do is set up some input data to understand what the [re-offending] rates are going to be."* (DES5, episode 22)

Furthermore, on one occasion a DES modeller (DES4) made assumptions about the existing number of offenders in society waiting to re-offend, with the intention of setting a model in a steady state. Therefore, the number of re-offending serious was set up based on calculations of the initial number of people in prison multiplied by the percentage that re-offend.

> *"...so 10% of these people [serious offenders in prison], 26,000 initially, divided by 20 equals to 1,300 coming out each year and 130 of those [serious released] will re-offend each year over 2 years which is 260 in there, in that buffer there [recidivists]."* (DES4, episode 107)

Overall, the provided data were an important starting point for the DES modellers. When making considerations about the model scope, the factors and the level of detail to include in the model, some DES modellers made decisions based on the available data. They conceptualised the model and referred to parts of it based on the data inputs. This gives an indication that the DES modellers specified the model based on the data provided. This is of course against advice given in textbooks on DES modelling, where modellers are advised not to let the data drive the model (Pidd, 2004; Robinson, 2004). For example, DES4 decided not to consider the different offending rates around the country and the geographic representation of the prison system due to the fact that specific data had not been provided.

> *"…so we are just looking at the total prison population as a total …because we haven't got the data on offending rates in different areas of the country*" (DES 4, episodes 5 and 6)

Furthermore, DES5 gave more concrete evidence of this phenomenon:

> *"Work with the data backwards really, kind of I'm going to map out the numbers, how they are allocated and try and understand what assumptions need to be made in the model at each stage."* (DES5, episode 2)

Modellers' reactions to missing data depended on each individual modeller. There was, however, an indication that DES modellers did not totally rely on the data when building a simulation model. 2 out of the 5 DES modellers were happy to make their own assumptions about the missing data.

SD modellers on type of data

SD modellers referred to the numerical data provided in the case study, such as number of offenders in prison, offenders entering prison, sentence length and time for recidivists to return to jail, and used them without encountering any difficulties.

Contrary to initial expectations, the SD modellers emphasised the need for data collection or so-called ‘data mining’:

> *“… if you want to explore two alternatives, you should have available some data mining. What the consequences are of pursuing those different alternatives.”* (SD5, episode 3)

In addition, 3 out of 5 SD modellers (SD1, SD2 and SD4) referred to the data provided in the graph representing the total prison population over time (Figure 4-1). These references are classified as qualitative data. The attention these modellers paid to qualitative data is relatively small. Nevertheless, this shows that some attention was paid to the qualitative aspects of the data provided. SD1, SD2 and SD4 verbalised on qualitative data, respectively 7.4%, 4% and 1% of the total words verbalised on data inputs. Their attention to qualitative data includes mainly references made to the graph provided. Some examples follow:

> *“…when I was reading the first part I was thinking, that’s the historic trend that you have laid out…”* (SD1, episode 9)
>
> *“The information we got here is largely based on the total prison size and that has been built over time”* (SD2, episode 2)
>
> *“…the graph you’ve got here is over 30 years…”* (SD4, episode 5)

In some instances, the SD modellers referred to soft variables such as the psychological factors involved in the prison population model. Only 2 out of the 5 SD modellers referred to these. SD3 only mentioned that these factors could also be considered in the model, while SD5 included these factors into the model. SD5 referred to the fact that the sentence length can affect prisoners' behaviour in prison or their criminal career in the future.

> *"...the interesting question is that if the petty crimes people [petty criminals] spend a long time in jail before coming out, are they going to offend more, are they going to become more violent and what are they going to do?"* (SD5, episode 5)

With regards to the missing data on petty and serious re-offending, only SD3 was willing to make assumptions, whereas SD1, SD2 and SD4 wanted to be provided with these data. This shows some level of unwillingness to make assumptions regarding missing data by SD modellers. SD5 also required to be provided with further data feeding into his conceptual model before starting to build the model on the computer.

> *"...it [re-offend petty] is not there, there was some figure wasn't it? No. OK. That's a figure that we can investigate anyway. So let's say, it's about 20, let's say its 30% or 0.3, commit petty crimes."* (SD3, episode 35)
>
> *"But we don't know what proportion of people [return back to jail]..."* (SD1, episode 41)
>
> *"...and we don't know what the fraction of petty recidivism is."* (SD2, episode 50)

"I think what is missing from here is that … but we don't know what proportion of petty criminals are recidivists. It won't be all of them. Just to confirm that there is no proportions given." (SD4, episode 102)

"*…what data I need, is what is the economic impact and psychological impact of the crimes committed by these 4 categories of people who are outside the prison. You have to research it."* (SD5, episode 30)

"I know that defining the exact mathematical forms is actually a big research project…" (SD5, episode 76)

Furthermore, SD modellers showed an interest in creating models in equilibrium. Therefore, they calculated the initial number of recidivists in society, so that a system in equilibrium was achieved. For example, SD1 calculated the initial number in the stock so that an equilibrium between the inflows and outflows was achieved. SD2, SD3 and SD4 also adapted the data inputs in order to achieve an equilibrium state. Interestingly, SD2 used the term 'steady state' to refer to the well-known SD term of equilibrium, but this is not the same as in DES terminology. SD modellers' approach to equilibrium will be further considered in section 7.7.1.

"What we gonnna do is initialize that stock [serious recidivists]. 7,000. So if we …, just putting the right values, which is, 7,000 for a steady state." and then verifies the result by: "*OK. That looks steady state. That looks steady state*." (SD2, episodes 75, 76 and 77)

> "*So I suppose, for that to be stable, we can increase this proportion here which is 1%.*" (SD3, episode 79) *"... Yes there we are. That's pretty stable. OK.*" (SD3, episode 80)

Comparative view of DES and SD modellers on type of data

From the analysis made, it is evident that a few similarities and differences exist between the DES and SD modellers involved in the modelling sessions.

- For the missing data of petty and serious offenders re-offending, only 2 DES and 1 SD modellers were happy to make guesses.
- Both DES and SD calculated some missing data to enter in the model. However, the nature of the calculations differed depending on the software requirements. Some DES modellers attempted to calculate the inter-arrival times for prisoners entering jail, while the SD modellers tried to set up the initial number of prisoners in the stocks of petty and serious recidivists.
- When entering in the model the number of recidivists already in society, the SD modellers set up these data so that equilibrium would be achieved. One DES modeller (DES4) also made similar considerations, which suggests that this feature did not characterise the SD modellers only.
- 3 DES and 4 SD modellers clearly pointed out that they needed to be provided with relevant data inputs. The SD modellers expected to be provided with data based either on concrete/observable processes or abstract

aspects of the model. DES modellers, on the other hand, referred mainly to data based on specific/concrete processes already mentioned in the case study.

- Apart from the numerical data, the SD modellers referred to the graphical display in the case study, while none of the DES modellers considered the graph.
- With regards to qualitative data, these were only mentioned by two SD modellers as compared to none of the DES modellers.
- It was observed that the DES modellers built a model based on the data provided, whereas the SD modellers preferred to build a conceptual model first and then ask the client to provide them with the data required.

7.5.2 Randomness

Information regarding the randomness in the model was intentionally left out of the case study description. In accordance with hypothesis 2.12, evidence of DES or SD modellers referring to variability or randomness was sought. It was observed that all DES modellers thought about adding variability into the model using statistical distributions or conditional coding. The DES modellers were not satisfied with using average numbers. On the other hand, the SD modellers did not think about randomness as it is not mentioned at all in the SD protocols. It is believed that in SD, randomness is modelled via delays. However, the analysis reveals that the use of delays was not necessarily associated with the representation of randomness for all the SD modellers (section 7.4.3). Due to the lack of references to randomness by the SD modellers, the analysis that follows focuses on the 5 DES modellers only.

> *"Well it says on average, on average it says, it doesn't really tell me anything."* (DES3, episode 90)

DES modellers on randomness

Some DES modellers (DES1, DES3 and DES5) started off with fixed numbers and added statistical distributions later on. Eventually, all the DES modellers added randomness in the model in one way or another. The following statement stresses the importance of randomness for the DES modellers:

> "*I am not quite a fan of fixed times, I like to have distributions in there …"* (DES5, episode 187).

The DES modellers added randomness to their models by either entering statistical distributions or using conditional coding, which included random numbers in order to determine the routing of individuals in the model (DES3, DES4 and DES5). The DES modellers' considerations with regards to choosing statistical distributions differed. Some modellers (DES3 and DES5) directly requested to be provided with some information regarding the behaviour of the variables or the shape of distributions:

> *"… but ultimately I won't put a distribution in there unless someone is going to say, here's your data."* (DES5, episode 188)

Therefore, they were provided with some information regarding the type of distribution to use. The researcher suggested the use of the negative exponential distribution for arrivals in prison, and the Erlang for time spent in prison and time recidivists spend in society before re-offending. When using a statistical distribution, the DES modellers tended to request further information regarding the parameters involved. For example, for the Erlang distribution information regarding the k parameter was requested.

There were also cases when DES modellers made their own assumptions about the type of distributions used. For example, in the absence of data, DES1 and DES4 considered using the uniform distribution for the inter-arrival of prisoners. As for sentence length, DES3 considered using the uniform distribution, while DES1 the negative exponential. DES3, who used Flexim, considered using a discrete distribution (the Bernoulli distribution) to set the re-offence probability for petty and serious offenders. DES2 represented the same aspect using probability profiles (empirical distribution) in Simul8. DES2 distinguished petty prisoners into re-offenders/recidivists and rehabilitated people, giving respectively 73% and 27% chance for each. A similar procedure was followed for serious offenders, using different percentages (30% and 70%). Petty recidivists were distinguished into petty and serious recidivists in a similar way.

However, in a real life project the DES modellers would be more cautious about the choice of statistical distributions. The client would be the first point of contact to get relevant information about the distribution of variables in the real system:

> *"…clearly we are not going to have a constant in there and that would be worth reviewing with the customer."* (DES3, episode 131)
>
> *"…go back to the client and say right we need an arrival distribution. Do you have an arrival distribution so that we can then translate that into a model distribution?"* (DES5, episode 13)

Furthermore, the choice of distribution was normally based on the type or number of parameters required. Distributions that do not have requirements for many parameters or that use average values as parameters, were preferred. For example, on one occasion the normal distribution was avoided due to its requirements for the value of the standard deviation:

> *"I'll go for normal distribution actually. No we won't, it has a standard deviation. Stick with the negative exponential."* (DES1, episode 37)

Comparison of DES and SD modellers on randomness

- As expected, the DES modellers considered randomness as an important aspect of DES modelling, while the SD modellers were not concerned about the aspect of variability in the data.

- The DES modellers' reactions to the missing information on the variability of the data in the model differed. Some modellers required to be provided with this information, while some were happy to make their own assumptions.
- The information requested by some DES modellers included the behaviour of the variables or the shape of the distribution (parameters of the statistical distribution), which is equivalent to requesting relevant information from the client in a real life project. In general, the DES modellers mentioned that in real life projects they are cautious about the choice of statistical distributions. The choice of statistical distributions was implied to be a decision based on the existing data and discussions with the client.
- Other DES modellers chose to use some distributions, based mainly on the type or number of parameters required. The uniform, negative exponential or probability profiles (empirical distributions) were most preferred. The normal distribution was avoided due to the fact that the standard deviation, which was not provided in the case study, is required.

7.5.3 Relationships between variables

With regards to the type of relationships between variables, the DES and SD protocols are reviewed with the view to identifying linear and non-linear relationships (hypothesis 2.13). It should be noted that most relationships were already provided in the case study, which mostly included linear ones, with a limited scope for non-linear relationships. A relationship between variables is considered linear where the rate equations are weighted sums of the rate variables (Sterman, 2000, p. 264). Whereas in non-linear relationships the increase or decrease of fractional rates change with time and these changes are not constant as in the case of linear relationships. In non-linear relationships, a diminishing effect is present, implying that the increase/decrease cannot go on forever (Sterman, 2000).

It was observed that all DES and SD modellers considered linear relationships between the variables in the model. In the case of DES1, for example, the variable released prisoners who re-offend as serious was defined by a linear equation, as a proportion of the number of petty in prison.

> "*…so a percentage will go into the prison buffer for petty and serious*" (DES1, episode 5).

Similarly, SD modellers set up petty and serious recidivists in the form of a linear equation, where the number of petty or serious prisoners in prison was divided by the sentence length and multiplied by the proportion of re-offending. Compared to DES modellers, most SD modellers were more familiar with using the word 'proportion', while some SD modellers (SD1 and SD4) used 'proportion' and 'percentage' inter-changeably. Obviously, both these terms played the same role in the equations.

> *"...petty criminals average sentence multiplied by the proportion of re-offending"* (SD2, episode 54)

"...serious offenders, average length serious sentence divided by the proportion serious re-offending" (SD2, episode 59)

"...the rate of released is going to be the number of prisoners in there [prison] divided by the average sentence length..." (SD3, episode 29)

The value of prison occupancy was also measured using a linear relationship by DES and SD modellers. SD modellers called it degree of overcrowding, which conveys the same idea.

"*...occupancy value must be number of people in prison divided by the ideal capacity*" (DES1, episode 79).

"Total prison population ... connect that to prison capacity" (SD2, episode 82)

"This capacity can be compared with the existing prison population. And that gives the degree of overcrowding." (SD5, episode 67)

Nevertheless, some examples of non-linear relationships between variables can be identified in some of the SD protocols. Non-linear relationships were not very frequently mentioned, however, it is worth observing that the SD modellers thought about them as opposed to the DES modellers who did not refer to them at all. For example, SD2 set the relationship between re-offending and recidivism, calling it 'recidivist factor', as a non-linear function. SD5 also considered the relationship between the amount of money invested on reducing crime and the number of rehabilitated prisoners, as a non-linear function.

> *"...overcrowding is going to drive by some sort of measure. ... guess it could be Proportion of recidivism. So ...I will call it "recidivist factor".*
> *"I will make it a graph [an S-shaped graph]. That would be a function of overcrowding, which will drive the amount of re-offending, the proportion re-offending."* (SD2, episodes 90 and 91)
> *"So now, what you find is maybe you can easily take the first 10 persons to come this route [be rehabilitated] without spending too much money, but if you want 90% of them to come this way, maybe it will cost you huge amounts of money, so that's the research you have to do. Figure out how much it costs to have one person go this way [the rehabilitation route] instead of going this way [the re-offend route]."* (SD5, episode 44)

Summary of the comparison of the DES and SD modellers' use of relationships between variables

From the analysis made, similarities are revealed between DES and SD modellers with respect to the use of linear relationships. Differences are detected with regards to the use of non-linear relationships.

- It is clear that both the DES and SD modellers thought in terms of linear relationships between variables.
- Interestingly, even though non-linear relationships were not mentioned in the case study, SD modellers referred to some non-linear functions. This was not the case with DES modellers.

7.6 Verification & validation (V&V)

Verification & validation is an important stage of the model building process in DES and SD modelling. As part of this stage, the various activities attended to whilst building simulation models have been grouped into 3 categories. Verification consists of checking the consistency of the computer programme, by making sure that the computer model is free of programming errors (Pala et al., 1999). On the other hand, validation is defined by the same authors as the process of checking whether the model is an accurate, good enough representation of reality and checking the correspondence with the real system. Validation can be divided into white- and black-box validation. White-box validation deals with checking that parts of the model are true to the real world (Robinson, 2004, pp. 215). As opposed to verification, white-box validation is not performed by the modeller alone, but it requires the involvement of those knowledgeable about the real system. It should be noted, that in the model building exercise implemented in the thesis, there were no real world clients involved and, therefore, the modellers had to accept that the conceptual model is relevant to the real world. As a result, a clear-cut distinction between verification and white-box validation is not considered appropriate. Referring back to the two types of validation, white-box validation is normally performed throughout model coding, while black-box validation requires a completed model. Black-box validation compares the model to the real world, checking it provides accurate outputs (Pidd, 2004; Robinson, 2004). In SD terms this is equivalent to behaviour validity tests, that is, checking that the model is capable of producing acceptable output behaviour (Pala et al., 1999).

From the quantitative analysis of the scripts, it has already been established that DES modellers pay more attention to model verification and validation compared to SD modellers (section 6.3). In the following sections, the DES (section 7.6.1) and SD (section 7.6.2) protocols are analysed regarding verification and validation. At the end, a summary (section 7.6.3) of the key differences found in the DES and SD modellers' approach to verification and validation is presented, drawing mainly on the aspects modellers give most emphasis to (hypothesis 2.14) and on the amount of attention paid to black-box or white-box validation by the DES and SD modellers (hypothesis 2.15).

It should be noted that even though a specific definition of the different types of verification and validation was set before coding the protocols, the distinction between them is not clear cut. As has already been explained in section 5.5.2, this has been dealt with by coding the scripts twice after a period of 3 months and also by checking the codings with a third party. In this way a more objective coding is derived. Due to the nature of the exercise, the distinction especially, between verification and white-box validation is not considered appropriate and these are, therefore, considered as one category.

7.6.1 DES modellers on verification and validation

While building a simulation model, DES modellers check the model to ensure that it depicts the intended behaviour. Below, comments are made about the two categories of V&V identified in the DES protocols, that of verification and white-box validation and that of black-box validation.

Verification and white-box validation

The DES modellers spent a significant amount of time with verification and white-box validation. As part of verification the modellers were expected to check that the code works correctly. As part of white box validation, it was expected that the modellers would check that parts of the model behaved as intended (i.e. similar to the real system). Access to real world information was not available in the modelling sessions, so it is accepted that white-box validation is implemented based on the information provided in the case study material. In most cases the DES modellers performed verification and white-box validation at the same time and there was little scope for distinguishing them. Two types of checks were performed, visual checks and numerical tests. As part of visual checks, the DES modellers confirmed that a specific part of the model was working or that the right type of prisoners were going to the right place. Numerical tests performed ensured that parts of the model exhibit the correct behaviour, by using time series graphs or numerical outputs. These included checking various aspects of the model, such as: *"do we get the expected prisoner arrivals in the system?"*, "*are petty and serious* *offenders in prison represented correctly?*", "*is the release activity working as intended?", "does the routing of released prisoners provide the expected results?*", etc.

In the instances when the model was not working correctly, the DES modellers went back to the code to check and find out the reason/s this was happening. DES modellers often checked the code while model coding to ensure that no mistakes were made as they went along.

> *"I'm gonna look at what we've actually got, to figure out why is that not working. It's definitely got values coming in there. Why is it not doing the calculations? … Ah, this should be multiplied by 100. A simple mistake"* (DES1, episode 83).
>
> *"Let's just check that this is OK, this is just for single people, the quantity is 1, so when they get back in there they just go back in as one person. Check that for serious."* (DES2, episode 97)
>
> *"Maybe the best thing to do is to display the ports on the screen so that we can see what's going on. If we just blow the connector size that might give me a clue as to what's happening."* (DES3, episode 266)

The DES modellers added various facilities while coding to enable the verification and white-box validation of the model. Some examples are: giving different colours to petty and serious offenders or creating a code to ensure that a specific condition (e.g. when someone gets released, the model is stopped) is taking place or by closing different routes in the model, etc.

> *"I can check to make sure that when I have got new arrivals coming in, they are actually the right colour."* (DES3, episode 127)
>
> *"…let's just check it out. … Stop when you are due to be released. Your release date is one thousand and …"* (DES3, episode 239)

Furthermore, the DES modellers tracked the individual people to ensure that the model was working as intended. In this way the characteristics of each individual prisoner were checked to ensure that the routing to re-offenders or non-re-offenders (people being rehabilitated) worked properly:

> *"Let's have a look at bad guys [serious offenders]…So he is going to re-offend. And he is also going to re-offend. Anyone who is not going to re-offend? No."* (DES3, episode 225)
>
> *"Let's check that because I don't want to get this wrong. So he is going to re-offend, he is going to re-offend and … [repeated another 4 times]. Goodness me. Hooray! Somebody who is not going to re-offend! Happy with that."* (DES3, episode 226)
>
> *"So this will give me if all things are right, the correct flow for the prisoners jailed that weren't in there to start with."* (DES3, episode 261)
>
> *"Over the course of the year just checking that I have the right number of parts arriving. OK, I have."* (DES4, episode 29)

Due to the fact that the codes are embedded into the internal structure of the model, the mistakes were not always obvious. More specifically, DES3 found it difficult to identify the problem encountered with the model. This did not allow the continuation of model coding. In this case, it was suggested that the best practice was to verify the logic and the code of the model with colleagues or people knowledgeable about the software. This conveys that on one hand, powerful software such as Flexim can provide very useful facilities, e.g. 3-D animation, etc, but on the other hand, the modeller might spend significant time on understanding the code and identifying the mistakes made.

> *"No it's not really working. I think I must have done something wrong. I don't know what."* (DES3, episode 285)
>
> *"…what I will do now in reality is ask one of my colleagues why it doesn't work because I have spent enough time on it…The most productive way forward is to ask somebody why this is getting stuck."* (DES3, episode 291)

Black-box validation

The DES modellers showed an interest in checking the overall behaviour of the model, ensuring that model outputs match those of the real system (Robinson, 2004). The central output used by DES modellers to validate the model is the total prison population, checking that there are 76,000 offenders in prison. Equivalently, the same aspect is validated by checking that prison occupancy/utilisation reaches 120% (this is only the case with DES1), based on overcrowding data provided in the case study description (Appendix A.2). Most of the time, black box validation was performed when a complete version of the model had been created. Therefore, DES3, who did not complete the model, did less of this type of validation, but complemented it with longer attention paid to verification and white-box validation.

> *"…still below what it should get. We should be getting 76,000 there [total prison population]…"* (DES1, episode 94)
>
> *"…so at the moment it should give us 120 [occupancy] when it's full."* (DES1, episode 80)
>
> *"So I need to check that at the end of my warm-up time I have 76,000. Yeah, it is. And that's 265 rehabilitated, which maybe OK"* (DES2, episode 118)

In contrast to other DES modellers, DES5 inspected the outputs of the model, checking whether the model achieves a steady state, meaning that the model reaches a realistic condition. He was mainly observing the graph of the petty and serious prisoners, checking that the lines become flat with variation.

> *"It looks like it gets to a steady state, 57,000 now. Oh and now we are going up again. I will let it run on and we'll see if it gets to a steady state*." (DES5, episode 276)

7.6.2 SD modellers on verification & validation

In this section, the SD protocols are analysed in order to identify the objectives pursued as part of verification and validation by the SD modellers. SD literature suggests that confidence in the model is achieved as a result of structural and behavioural tests (Lane, 1995; Pala et al., 1999).

Verification and white box validation

The SD modellers checked the model and specific parts of the model, but this was less frequent than for the DES modellers. These checks included checking the equations included in the model, the behaviour of separate structures or parts of the model, identification of the inflows, outflows and the stocks' behaviour, and checking that an equilibrium state can be achieved. In a few instances, the SD modellers checked the conceptual model against the model on the computer to ensure that this was correct. Characteristically, SD5 checked the computer model systematically against the conceptual model on paper to ensure that all the parts were included in the model:

> *"I would like to just double check on the other two stocks [petty and serious recidivists] that they are not behaving in some sort of strange way."* (SD1, episode 100)
>
> *"...looking at the figures, I suppose 50,000, so on average with a 5 year sentence, I suppose according to the model anyway, we are releasing 10,000 per year and we are only bringing back in 3,000 new offenders. So that's why it is declining."* (SD3, episode 73)

> "*I'm just checking here that what I've done so far seems to work…*" (SD4, episode 58)
>
> *"Just check if that's right. So we've got 54, 54 a month coming in [the stock petty in prison]. That's 650 divided by 12, we've got 62 a month coming out. OK. So that's interesting."* (SD4, episode 59)
>
> *"…serious prisoners it should have 4 inputs and 2 outputs, so 1, 2, 3, 4, 5, 6…"* (SD5, episode 95)

Other techniques used by one SD modeller (SD3) to validate the results of the model were the use of the 'extreme conditions' test or the 'integration error' (Sterman, 2000), where different time steps were used to check the effect on the final model outcomes.

Black box validation

When validating the model in the context of black box validation, the SD modellers concentrated mostly on making sure that an equilibrium state is achieved, this provides confidence that the model outputs are correct. In the cases when the model behaviour did not conform to equilibrium, the SD modellers considered checking or validating the model with someone who is knowledgeable about the situation or the real system.

> *"We've created an output now that provides a perfect equilibrium on petty criminals in prison"* (SD1, episode 97)

> *"That seems a steady state. It seems OK for petty criminals."* (SD2, episode 58)
>
> *"Just to see what we've got in terms of that. Let's just run the model again. Nice straight lines. I'm pleased with that."* (SD4, episode 160)
>
> *"I think at this point I would say to the people I was building the model for, I would expect that number to be flat or maybe slightly increasing, why do you think it is reducing?"* (SD4, episode 60)
>
> *"…the fact that these lines are reasonably straight would re-assure me that we've got the general information correct."* (SD4, episode 179)

Most SD modellers used the graphs to validate the outputs of the model. This is relevant to performing behaviour validity tests. The SD modellers were keen to ensure that the model reproduced the reference data provided in the case study description. Emphasis was given to the capacity of the model to represent the appropriate pattern of behaviour.

> *"I don't like the shape of that graph so I am going to double check that I've got it right."* (SD4, episode 75)

The accuracy of the model was mentioned by one SD modeller, who implied that the accuracy of the model is not the objective of the modelling exercise. Due to the large number of assumptions made, the model created did not serve the purpose of providing realistic results. This suggests that the modeller was more concerned with issues related to the structure of the model rather than the predictive ability of the model.

> *"There is so many assumptions that the accuracy of the model is not significant."* (SD3, episode 107)

7.6.3 Comparative view of DES and SD modellers on verification & validation

Based on the analysis performed, differences in the verification and validation (V&V) approach taken by DES and SD modellers have been identified. In addition, the attention paid to V&V by DES and SD modellers has been divided into the two categories: verification & white box validation and black-box validation. Table 7-1 depicts the number of words verbalised by DES and SD modellers and the respective proportions by each of the two V&V categories. The box and whiskers plot (Figure 7-1) shows the spread of the number of words verbalised for the two V&V categories. Based on the box and whiskers plot, it is identified that the DES and the SD modellers pay a different level of attention to the 2 V&V categories (verification & white box validation and black box validation). In addition, it is observed that attention among the SD modellers has a bigger spread. Using the chi-square test it is confirmed that the distribution of attention among the two categories differs among the DES and SD modellers at a 5% level (χ^2=311.36[12] compared to the critical chi-square value for 1 degree of freedom = 3.841 (Sheskin, 2007)). The

[12] Yates correction for 2x2 tables has been applied for the calculations of the χ^2.

figures (Table 7-1) suggest that the SD modellers paid more attention to black-box validation, that is checking the model results, compared to the DES modellers. Meanwhile the DES modellers paid more attention to verification & white-box validation.

Table 7-1: Verification & Validation – division into two categories comparing the DES and SD modellers' verbalisations (number of words and proportions) on verification & white-box validation and black-box validation

	DES modellers		SD modellers	
	Number of words	Proportions	Number of words	Proportions
Verification & White-box validation	3143.0	60%	858.0	83%
Black-box validation	2097.0	40%	176.7	17%
Total on V&V	**5240.0**	**100%**	**1034.7**	**100%**

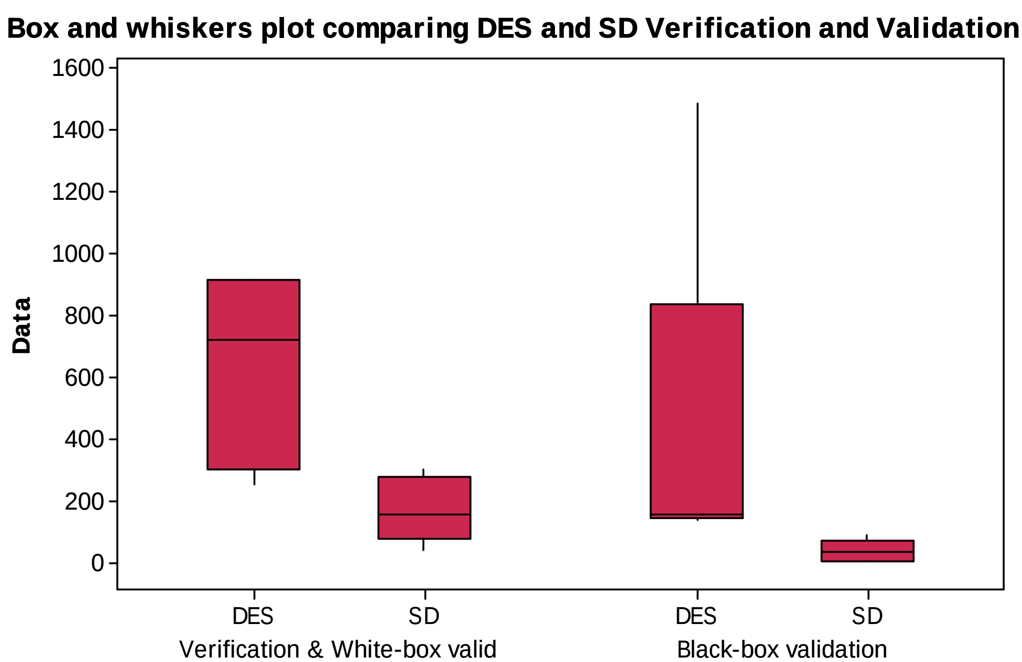


Figure 7-1: Box and whiskers plot comparing DES and SD modellers' distribution of attention between verification & white-box validation and black-box validation.

The following key differences between the DES and SD modellers' V&V process undertaken are identified:

- The DES modellers spent significant effort in model verification and white-box validation. They were keen on checking the code on the computer and that parts of the model were working as intended. It is suggested that the greater attention is a result of the greater complexity involved in most DES software, where the mistakes are not always obvious and, therefore, it takes DES modellers longer to identify. On the other hand, the SD modellers engaged themselves in less detailed verification and white box validation activities. Coupled with the fact that the SD modellers did not identify many mistakes in the code, they spent less effort implementing these tests.
- Regarding the approach taken to V&V, it was found that the SD modellers engaged relatively more in black-box validation as compared to the DES modellers. The SD modellers systematically checked the final results of their models against the real life system. The difference lies in the fact that the SD modellers show more interest in predicting the pattern of systems' behaviour against the real life system, rather than point predictions. Meanwhile DES modellers mainly checked the point predictions, that the outputs derived were accurate, compared to the real system. Furthermore, it was found that the SD modellers checked that these outputs represent a system in equilibrium. SD modellers gained confidence in the outputs of the model when ensuring that these reach an equilibrium state.

- Interestingly, no evidence was found to suggest that the SD modellers referred to model usefulness with respect to its purpose (Forrester and Senge, 1980). One plausible reason that explains this is that in the modelling exercise a simple model is involved. As a result, there was not much scope for the modellers to consider model utility, in the form of: thinking how the model will be used, testing parameters, or performing sensitivity analysis. Furthermore, the fact that there was not a real client has to some extent reduced the scope for model validation.

7.7 Model results & experimentation

The findings presented in this section deal with aspects of the verbal protocols referring to discussions about the results of the model and the development of scenarios. It should be noted that the exercise did not necessarily include much experimentation, but this was discussed by some DES and SD modellers in the protocols. In addition, the time limits for the completion of the exercise restricted modellers' thoughts on experimentation. In the analysis that follows the protocols are explored regarding the achievement of a steady state or equilibrium (hypothesis 2.16- section 3.4.2), model results (hypothesis 2.17- section 3.4.2) and experimentation (hypothesis 2.18 - section 3.4.2).

7.7.1 Achieving a model in a steady or in equilibrium state?

After building the simulation model on the computer, usually DES modellers take an interest in creating a model in a steady state. Meanwhile SD modellers think about achieving an equilibrium state. The DES and SD protocols are analysed to explore whether these aspects are approached in a different or similar way by DES and SD modellers, based on hypothesis 2.16.

DES modellers on achieving a steady state model

The 5 DES protocols are reviewed to count for the number of times the keyword 'steady state' appears in DES modellers' verbalisations. In turn, this implies the attention paid by DES modellers in achieving a model in a steady state condition. Only DES5 mentioned the word '*steady state*', which appeared 10 times in the protocol. No other DES modeller did. On one occasion, DES2 referred to the behaviour of the variable prison population, by observing the graph to identify the "*numbers levelling off*". The activity of checking the graph lines becoming flat can be considered part of dealing with the initialisation bias.

> *"It would be interesting to see if it gets to a steady state."* (DES5, episode 253)

There are two main ways of handling the initialisation bias in DES, either to run the model for a warm-up time before collecting results or to set initial conditions in the model (Robinson, 2004). On this note, all the DES modellers (apart from DES3) thought about setting a warm-up time, but only DES2 finally used a warm-up time to initialise the model. DES1, DES3, DES4 and DES5 set the initial conditions in the model by entering prisoners in the queues for prison and recidivists at the start of the model run. Overall, it is observed that while DES modellers (apart from DES5) did not often think about creating steady state conditions in the model, they ultimately used relevant techniques, such as dealing with the initialisation bias, to achieve a steady state.

> *"…we give ourselves a starting volume…. Start up volume of recidivists, who are going to re-offend…"* (DES4, episode 101)

SD modellers on achieving a model in equilibrium

Achieving a model in equilibrium was important for the SD modellers. This was closely related to the experimentation stage, where scenarios and policies are implemented.

> *"I think it is important to get an equilibrium position, because if you are exploring the impact of different strategies…"* (SD1, episode 33)

It is observed that SD modellers used various terms to refer to a model in equilibrium. Therefore, the number of times the keyword 'equilibrium' and other relevant words appeared in SD protocols is counted to illustrate this fact (Table 7-2). Interestingly, the word 'equilibrium' was used only by two SD modellers (SD1 and SD4). However, this does not include all the instances that the SD modellers thought about the equilibrium in the model. Interestingly, the word ‘steady state’, which is a term usually used by DES modellers, is also mentioned in the context of achieving an equilibrium state by SD2. Other related terms referred to by other SD modellers (SD3, SD4 and SD5) are ‘stable’ or ‘steady’ model (Table 7-2).

As already mentioned in section 7.5.1, the term ‘steady state’ borrowed from DES modelling is viewed slightly differently in SD modelling. A steady state model (equivalent to a model in equilibrium) is achieved by changing the initial value of the stock, so that stable outputs (flat lines) are achieved, whereas in DES modelling, the steady state does not necessarily mean achieving straight lines (section 2.10).

When checking that the model is in equilibrium, most SD modellers (SD1, SD3 and SD4) referred to the graphs and the lines of the main outputs of interest in the model. For example, SD1 and SD3 tested that the model is in equilibrium by checking that the lines of petty and serious in prison became flat.

> *“…in the serious criminals but it’s only a margin of a few hundred above and below 26000.”* (SD1, episode 99)
>
> *“For serious, its 390 here, that’s better now, serious is 390, create a steady state. And, so that’s now 390, let me finish this. So I need to balance that out…”* (SD2, episode 79)
>
> *“We ended up with pretty much straight lines. Serious one is going up a little bit, but not alarmingly, so the prison population is relatively stable.”* (SD4, episode 178)

Table 7-2: Count of the times the keywords "equilibrium", "steady state", "stable" or "steady" appear in the 5 SD protocols.

Modeller	Equilibrium	Steady-state	Stable	Steady
SD1	18	0	0	0
SD2	0	8	0	0
SD3	0	0	3	0
SD4	1	0	1	0
SD5	0	0	0	1

Summary of the comparison of DES and SD modellers' views on steady state or equilibrium state

With regards to hypothesis 2.16, it was observed that DES and SD modellers referred to similar concepts of equilibrium or steady state, but these had different underlying meanings. The following observations were made with regards to the activities followed in creating models in steady state or in an equilibrium condition by DES and SD modellers:

- The DES modellers did not always think about creating a model in a steady state condition. However, this seemed to be an underlying aim in the efforts to reduce the initialisation bias (warm up or initial condition).
- The SD modellers thought about creating a model in equilibrium, but there was not a specific term used. They used the word 'equilibrium', 'steady state' or they referred to achieving flat lines for the model outputs.
- Overall, the 5 SD modellers were inclined to think about creating a model in equilibrium, albeit the extent varied between modellers, all the SD modellers referred to it in one way or another. As already mentioned (section 7.6.2),

creating a model in equilibrium was important for SD modellers from the point of view of creating a valid model, which can then be used to run various scenarios.

7.7.2 Model results in DES and SD models

With respect to hypothesis 2.17, the DES and SD protocols are reviewed in relation to references made to model results. The references to the results are classified into quantitative and qualitative. The results are considered quantitative when modellers referred to numerical values of variables of interest or when predicting or estimating these numbers in the future. On the contrary, when modellers considered or looked at the graphical displays of variables of interest, it is believed that the general dynamic tendencies of the model are looked for, and therefore, these are considered as qualitative results as opposed to numerical results.

DES modellers on the results of interest

3 out of 5 DES modellers used both graphical and numerical displays when discussing the results of the model, while all DES modellers referred to the numerical display of the results (Table 7-3).

Table 7-3: Type of results referred to by DES modellers

	Result of interest	Numerical display	Graphical display
DES1	Prison ‘occupancy/utilisation	Yes	Yes
DES2	Petty in prison	Yes	Yes
	Serious in prison	Yes	Yes
	Total entered prison	Yes	-
	Rehabilitated prisoners	Yes	-
	Re-offenders	Yes	-
DES3	Petty in prison	Yes	-
	Serious in prison	Yes	-
DES4	Total prison population	Yes	Yes
DES5	Petty in prison	Yes	-
	Serious in prison	Yes	-

All DES modellers created variables that counted the number of people in prison or that of prisoners entering jail, which were presented on the screen. The number of prisoners or prison occupancy level was considered a central output by DES modellers. Obviously, due to the limited time available and the simplicity of the model, there was no scope for an in-depth analysis of the results. The references found were mainly comments on the face value of the results obtained:

> *“I need to look at occupancy level.”* (DES1, episode 78)
>
> *“We could look actually at the graph that shows how it is increasing over time.”* (DES2, episode 128)
>
> *“We can see how many people are in prison by looking at how many are in a queue that would work for me.”* (DES3, episode 55)
>
> *“So now we can see that we started with 76,000 but some of these start coming out, you see that dropping.”* (DES5, episode 251)

Obviously, DES models provide the potential to collect a range of detailed statistical outcomes. Only 2 DES modellers (DES2 and DES5) referred to these types of results making connotations of detailed model outputs:

> *"…can work out exactly how long they have left in and what their sentence was"* (DES5, episode 93)
>
> *"If I had more time we could then calculate how many times re-offenders re-offended and what type of person re-offended and how many people that were originally in jail for petty crime re-offended and so you know you get the statistics out.…"* (DES5, episode 181)

SD modellers on the results of interest

In the SD protocols, it is observed that numerical and graphical displays of the model outputs have been variously consulted (Table 7-4). The main output of interest for most SD modellers was the total prison population, or that of petty and serious, which most SD modellers represented in numerical and graphical form. In addition, some other outputs were considered, such as the annual incidence of crime and the recidivism factor. These mainly consisted of numerical displays. On the other hand, total cost was considered as the main model output, by one SD modeller (SD5) and mainly in the form of a graphical display. SD5 specifically suggested that the behaviour of the costs incurred by various policies was the measure of interest. Obviously the cost included a range of other costs, such as: costs incurred in maintaining the functioning of prisons, costs incurred to society, the money spent to implement the policies, etc. However, SD5 was more interested in the shape of the line, which was expected to be different based on the scenarios chosen.

Table 7-4: Type of results referred to by SD modellers

	Result of interest	Numerical display	Graphical display
SD1	Total in prison	Yes	Yes
	Petty in prison	-	Yes
	Serious in prison	-	Yes
	Incidence of crime	Yes	-
	Ex-petty criminals	Yes	Yes
SD2	Total prison population	Yes	Yes
	Recidivism factor	Yes	-
	Prison capacity/ utilisation	Yes	-
SD3	Total prison population	Yes	Yes
SD4	Petty inside	-	Yes
	Serious inside	-	Yes
	Total prison population	Yes	Yes
SD5	Total cost	-	Yes

Some examples extracted from SD protocols follow:

> *"…what we should now have in the graph is how many ex-petty criminals there are in the community…"* (SD1, episode 71)
>
> *"I've got an initial growth in the number of petty criminals and then a stabilisation and a gradual reduction."* (SD1, episode 73)
>
> *"…number of people in prison … producing an output which helps people decide how many prison places they need,"* (SD1, episode 112)
>
> *"So that's the total prison population, and we are going to graph that…"* (SD2, episode 69)
>
> *"…a rough estimate of prison capacity* [utilisation]*."* (SD2, episode 87)
>
> *"…we have a time graph for the prisoners."* (SD3, episode 71)
>
> *"…the variable of the total …80,000 …"* (SD4, episode 162)

"…we ended up with pretty much straight lines." (SD4, episode 178)

Comparative view of DES and SD modellers on model results

From the analysis performed based on hypothesis 2.17, it is clear that DES and SD modellers handled model outputs not that differently. A list of the main findings follows:

- The DES and SD modellers similarly showed an interest in obtaining numerical estimates of variables in the model.
- The behaviour (patterns) of key variables in the model were also observed by some DES and all SD modellers, who looked at the lines of the graphs with a view to pointing out the general tendencies of the model behaviour.
- Optimisation of model results was mentioned only once by DES4, but due to the simplicity of the model, this was not explored further.
- A tendency of some DES modellers to look into the model results in more detail was further identified.

7.7.3 Experimentation

Even though it was not explicitly required in the exercise, a number of modellers thought about the scenarios that could be employed using the models created. Due to this fact, the ideas expressed are mainly spontaneous thoughts raised by the modellers and, therefore, in some protocols no or limited references were found. The thinking behind this is that the stage of experimentation can be considered more an aspect of model use rather than part of model building. Therefore, the following analysis attempts to answer hypothesis 2.18 (section 3.4.2) only to some extent.

Comparing the two groups of modellers, all 5 SD modellers referred to scenario building, while only 3 DES modellers did so. An overview of the scenarios considered, by modeller, follows in table 7-5.

Table 7-5: Scenarios considered by DES and SD modellers

	Variation of people entering prison	Variation of sentence length	Variation of re-offending	Variation of time to re-offend
DES1	X	X		
DES2				
DES3*	X	X		
DES4				
DES5	X	X		
SD1	X	X		
SD2	X	X		
SD3	X	X	X	
SD4	X	X		X
SD5		X	X	

DES modellers on development of scenarios

It is observed that not all the DES modellers thought about creating scenarios (Table 7-5). It should be taken into consideration that the DES modellers took longer to build the code on the computer and due to the limited time this resulted in less time left at the end of the exercise to think about scenarios. In addition, DES3 thought about the variables that could be varied as part of experimentation only at the start of the exercise. His initial plans were not materialised because a complete model was not created to enable experimentation. DES1 spent the longest time of all DES modellers with experimentation. Some examples follow:

"*...how much are we going to reduce the population of people coming into the system due to a stiffer rule?"* (DES1, episode 100)

"…what I'd probably do is, basically look at, if we increase sentences by 10, 20, 30% and maybe do a range of, do a 10, 20, 30% reduction of people coming into the system and look at all the different values and see how our occupancy level goes." (DES1, episode 112)

"So in my actual prison …I will increase the number of people in prison…" (DES5, episode 249)

SD modellers on development of scenarios

More variation was observed in the scenarios considered by SD modellers (Table 7-5). The basic scenarios included variation of sentence length and number of offenders entering prison. The latter was represented in different forms by each modeller: introduction of community sentences, alternatives to custody (SD1 and SD4), changes in crime rates (SD3) and imprisonment rates (SD4). Variations in the values of re-offending and average time to re-offend were some additional factors considered for scenario building, but to a smaller extent.

"… [if] we change the length of sentencing, so let's say that we have 6 years instead of 5 and we see an increase in the population and we see the recidivist effect…" (SD2, episode 107)

"...if we had a policy in place of rehabilitation let's say, which reduced the re-offending rate, then what impact would that have on the prison population..." (SD3, episode 82)

"*...what would happen if the crime rate rose for example?"* (SD3, episode 117)

"...spend some of the resources in trying to ensure that they do not offend again..." (SD5, episode 38)

"...play around with [the] ... policies and get different expressions of total cost." (SD5, episode 59)

Comparative view of DES and SD modellers on development of scenarios

With respect to hypothesis 2.18 (section 3.4.2), a summary of the findings on DES and SD modellers' thoughts about model scenarios follows:

- Contrary to expectations, the development of scenarios came more naturally to most SD modellers as part of the model building process, compared to the DES modellers. Not all DES modellers thought about creating scenarios.
- Variation of sentence length and number of offenders entering prison was the most commonly referred to scenario by DES and SD modellers. However, some SD modellers came up with other ideas such as: variation of the re-offend percentage and time to re-offend variables.

7.8 Implementation

As was the case with experimentation (section 7.7.3), the modelling exercise did not include implementation either, however, some modellers referred to it and therefore, these episodes are analysed in this section. Implementation normally involves interaction with the client. Due to the lack of a real life client in the modelling exercise implemented there is limited scope for verbalisations on implementation. Implementation can include three different aspects: implementing the findings from the simulation study, implementing the model and implementation as learning (Robinson, 2004). In the context of the modelling exercise, implementation of the findings would involve interpretation of the findings and discussions about their contribution to the wider debate of prison population, rather than implementing the findings *per se*. Implementation as learning was most common in the context of improved understanding of the problem situation as an outcome of the simulation study (Robinson, 2004, pp. 203). The analysis implemented here emerged throughout the modelling process and it is not related to any specific hypothesis.

DES modellers on implementation

Overall, most DES modellers did not think about implementation. Only 1 DES modeller (DES3) made references regarding implementation. These were mainly thoughts made relatively early during the modelling exercise regarding how the model would be used by the client.

"...I would normally ask more questions about how will a decision be arrived at..." (DES3, episode 32)

SD modellers on implementation

More SD modellers (3 out of 5) thought about implementation. This involved discussion regarding the learning gained about the problem (SD1), but also thoughts about how the model or its findings can be used. Some examples follow:

"Because what the model is suggesting to me, but I want to get a scale for it. So we are actually growing criminals!!" (SD1, episode 105)

"...that would be up to whoever the policy maker was to decide ... would the policy be appropriate in order to make that [reducing the re-offending rate] happen." (SD3, episode 116)

"So what that model would give you ...it would be a tool to have a conversation about the impact of, can we persuade the judges to shorten the sentence length or can we persuade them about the annual new [new offenders entering prison per year], for example?" (SD4, episode 172)

Comparative view of DES and SD modellers on implementation

The few references made to implementation, especially in the DES protocols do not allow for a great depth of analysis. However, some conclusions can be drawn.

- The DES modellers did not consider implementation to a great degree.

- The SD modellers were more aware of the purpose and use of the model after its development and made considerations mainly about the learning gained from the model and the implementation of the findings by policy makers. This again provides evidence that SD modellers viewed the problem from a wider scope.

7.9 Other modelling issues considered

The episodes coded as 'Other' are now discussed. These include the episodes identified in the protocols that are not covered by the 7 modelling topics already analysed in this chapter.

Other issues considered by the DES modellers

Some DES modellers were concerned about the simulation run time (DES1 and DES2). This is an issue with DES models, where in the next-event list there are a large number of events to process. This is largely a function of the model size. Indeed the other 3 modellers, who developed a different configuration of the large population, did not express particular issues.

Another issue mentioned by one DES modeller (DES3) is the need to be organised at the onset of the modelling project. Some of the aspects mentioned were: systematic naming (names given to people, activities/machines in the model), keeping notes of random number streams used, following a clear stepwise modelling process, etc.

2 DES (DES3 and DES5) modellers felt under time pressure. The main reason for this is due to the longer time taken to build the model. As has already been mentioned (section 5.5.2), their modelling sessions were the longest.

Other issues involved comments regarding the software used. DES4 implied that Simul8 is more flexible compared to Witness. However, DES5 commented on the opposite. Therefore, it seems that software preference depends on the individual modeller.

Other issues considered by SD modellers

The issues considered by SD modellers, did not refer to specific modelling problems. SD1 commented on the clarity of the case study description. SD2 made comments regarding the software used. Due to the fact that an older version of the software, Strategy Dynamics was used, errors were found in measuring units and the equations.

SD3 made references to the models he normally develops as part of his job. He mentioned that in his models he always incorporates policy options, which help the clients to decide on which options to choose. Interestingly, he also mentioned that the SD models he normally develops rarely incorporate softer variables, such as the 'psychological impact'.

> *"we really try quite hard to make sure that all the key sort of policy options are captured so that you can see what you could do and what the outcomes would be."* (SD3, episode 86)
>
> *"The models that we build tend not to have that psychological impact and these sort of effects and they tend to be a bit more straight forward."* (SD3, episode 105)

In addition, a discussion was also held between SD5 and the researcher due to his reluctance to build a model on the computer. SD5 highlighted the importance of conceptual modelling and creating the structure of the model, rather than setting up the model on the computer.

> *"I would have to go to the computer and feed this in to the computer with the numbers. That's a different aspect of modelling. That doesn't have many conceptual issues; it's much more of a mechanical task."* (SD5, episode 80)

7.10 Summary of the findings

In this chapter the 10 verbal protocols are analysed referring to the modelling topics and the aspects of modelling involved in research hypotheses 2.1-2.18 (section 3.4.2). The analysis refers to problem structuring, conceptual modelling, model coding, data inputs, verification & validation, results & experimentation, implementation and other topics. Some of the main similarities and differences between DES and SD model building revealed from the analysis are:

- The SD modellers took a wider view of the modelling objectives.
- The DES and SD modellers did not use specific conceptual modelling diagrams.
- The DES modellers took an analytic perspective when modelling, whereas the SD modellers a more holistic view.
- The DES modellers did not use existing modelling structures but the SD modellers did.
- The DES modellers used mostly quantitative data, whereas SD modellers used quantitative and qualitative data.
- The DES and SD modellers' reactions to missing data were found to be almost the same.
- Randomness was an important aspect for DES modellers, whereas it was not important for SD modellers.
- The DES modellers were familiar with linear relationships whereas the SD modellers showed higher familiarity with both, linear and non-linear relationships.
- Both the DES and SD modellers were concerned with creating accurate models.

- The DES modellers were interested in removing the initialisation bias and thus interested in creating a model in a steady state. SD modellers were interested in creating a model in an equilibrium condition.
- Both the DES and SD modellers showed an equal interest in the quantitative and qualitative aspects of the model results.
- The SD modellers were more keen on running scenarios using the model.

In the next chapter, the evidence for research hypotheses 1.1-1.3 and 2.1-2.18 (section 3.4) found in the model building exercise is discussed.

Chapter 8: Discussion of the findings of the model building study

8.1 Introduction

This chapter provides a discussion of the findings from the analysis of the verbal protocols. The aim is to draw conclusions regarding the comparison of DES and SD model building based on the results obtained from the analysis carried out on the model building process followed (chapter 6) and the approach taken to modelling (chapter 7) by expert simulation modellers.

The chapter outline consists of the two main sections, discussion of the findings from the quantitative analysis and the qualitative analysis. Each research hypothesis is separately discussed, providing a summary of the evidence collected and reaching overall conclusions as to whether they are accepted or not. In addition, issues are raised regarding the validity and the factors that might have affected the results of the study. Next, the main findings are summarised.

8.2 Discussion of the findings from the quantitative analysis

The current section discusses the findings from the quantitative comparison of the model building process followed by expert DES and SD modellers when developing simulation models (chapter 6). This is related to research objective 1 (section 3.3) and hypotheses 1.1-1.3. The discussion considers the implications of the findings regarding the comparison of DES and SD model building.

Each hypothesis is now discussed in detail.

8.2.1 Hypothesis 1.1 - Distribution of attention

Hypothesis 1.1: *DES and SD modellers pay varying attention to the different modelling stages. DES modellers spend most of their time in coding the model on the computer, whereas SD modellers spend most of their time in specifying the structure of the model.*

The statistical analysis (section 6.3), reveals that even though the DES and SD modellers verbalised to similar extents, the attention paid to the 7 modelling topics (identified in this study) differed. Model coding was a central topic in DES modelling. This was the topic that DES modellers paid more attention to as compared to the SD modellers, who paid more attention to conceptual modelling. Based on this finding it is, therefore, suggested that hypothesis 1.1 cannot be rejected. This finding ultimately raises the question: do DES modellers pay more attention to coding or is it inherently harder to code in DES modelling?

One possible explanation could be that DES modellers naturally pay more attention to model coding when building simulation models. This could probably be a result of DES modellers' tendency to creating a lifelike representation of the real system (Baines et al., 1998) and, therefore, engaging in the development of detailed computer models. On the other hand, SD models involve less detail and, therefore, modellers pay less attention to model coding. This, on the other hand, depends on the problem modelled and its complexity.

An alternative explanation to the question raised is related to the modelling software and internal coding logic. Obviously, DES modelling allows more complexity than SD. More flexibility and power is provided by DES software, whose internal logic allows modellers to add more options with the use of code or functions. These are not obvious to the end user and one needs to be knowledgeable about the software to be able to use efficiently the available facilities. On the other hand, most SD software have an easier layout and interface, where different modelling components are available on the screen as well as a choice of equations and relationships between variables. Despite the SD modellers' main focus on conceptual modelling, on the whole it is observed that they spend less time in completing the modelling exercise (Figure 5-1). It is suggested that this is a result of the fact that SD modellers take less time to build the model on the computer[13].

[13] Note that the time taken to complete each modelling stage is only roughly estimated by number of words because the primary focus of the analysis undertaken was the number of words verbalised. The time spent on each individual modelling stage could not be measured with accuracy due to the fact that each modelling stage occurred at various points of time throughout the modelling exercise and the time was not captured for each episode in the verbal protocols.

To conclude, it is believed that both factors considered contribute to the different levels of attention paid to the stages of simulation modelling by the DES and SD modellers. The validity of the data obtained should however, be considered in relation to:

- The higher variation in the attention paid to the seven modelling topics among DES modellers as compared to SD modellers (due to the DES modellers' higher values of the absolute and relative standard deviation - Table 6-3).
- The outliers involved in the DES group of modellers (DES3 and DES5 - section 6.2) could have contributed to the higher values of the standard deviation observed for the DES group of modellers.

However, observations of the ranked verbalisation data for each modelling topic (appendix C.1) show that the outliers identified and the higher variations in attention between the two groups do not affect the differences identified. For example, the verbalisations of DES3 and DES5 on conceptual modelling, model coding and verification & validation do not seem to skew the group values to a great extent and as a result, do not specifically affect the differences in the distributions between the group of DES and SD modellers. The same applies to the differences in the variation of attention paid by the two groups of modellers. Hence, it can be concluded that these factors do not affect the robustness of the findings.

8.2.2 Hypotheses 1.2 and 1.3

Hypothesis 1.2: *DES and SD modellers follow a similar sequence of modelling stages.*

Hypothesis 1.3: *DES and SD modellers follow similarly iterative modelling processes.*

The analysis of the DES and SD timeline plots (section 6.4) reveals that expert modellers followed a different order of steps during the model building process. This finding suggests that DES and SD modellers think about different modelling topics at different points of time during the model building process and hypothesis 1.2 is therefore, rejected.

Furthermore, it is observed that both DES and SD modellers often switched their attention among topics suggesting that an iterative modelling process was followed. Indeed, in the transition matrices almost the same number of transitions among topics is found in the DES and SD protocols. This finding supports hypothesis 1.3 and it is also consistent with the SD and DES literature, which suggests that simulation modelling is a repetitive and an iterative process (Sterman, 2000; Robinson, 2004). In addition, Willemain (1995) spoke about the iterative modelling process followed by expert OR modellers.

Nevertheless, from the comparison of the transitions of attention among topics, different styles of iterations are observed.

- The DES modellers followed a more linear progression through the modelling topics, whereas SD modellers' attention was more scattered among topics throughout the modelling task (section 6.4.1).
- The cyclicality of thinking was more distinctive for SD modellers as compared to DES modellers.

These findings provide useful insights about the process followed during model building, especially when considering the most useful approach in the most appropriate circumstances. For example, in the case when the problem is well defined and the simulation model is the product of interest, it is suggested that the DES model could serve the purposes of the project. On the other hand, if the problem situation has not been well-defined and the conceptualisation of the problem requires more attention, then the SD approach would be preferable. However, another view that can be taken is that the two modelling approaches provide complementary modelling skills. Learning to model, by studying both approaches can inherently enhance modelling skills and can be particularly useful for teaching simulation modelling to novices. It can instil different approaches to modelling and at the same time a different level of attention to various modelling stages. Therefore, the findings of this study may prove useful for educational purposes.

8.2.3 Summary of the discussion on the quantitative analysis of VPA

A summary of the findings by research hypothesis is displayed (Table 8-1). In light of hypotheses 1.1, 1.2 and 1.3, this study suggests that DES and SD modelling are two different modelling approaches from the point of view of the stages and the processes followed. It is also believed that learning both simulation approaches can improve the modelling capabilities of simulation modellers, in terms of the stages followed and the attention paid during modelling. Especially in education, it is suggested that teaching novices both simulation approaches can have a beneficial impact on their modelling skills because of the differing nature of the modelling process.

Table 8-1: Summary of findings of the quantitative analysis of the protocols, by hypothesis

Hypotheses on model building process	Research findings
Hypothesis 1.1: DES and SD modellers pay varying attention to the different modelling stages. DES modellers spend most of their time in coding the model on the computer, whereas SD modellers spend most of their time in specifying the structure of the model.	Cannot be rejected
Hypothesis 1.2: DES and SD modellers follow a similar sequence of modelling stages.	Rejected
Hypothesis 1.3: DES and SD are both iterative modelling processes.	Cannot be rejected, **but** different patterns of iteration are followed

8.3 Discussion of the findings from the qualitative analysis

In this section the findings from the qualitative analysis of expert DES and SD modellers' protocols (chapter 7) are discussed with a view to understanding their model building approaches. The results obtained serve as evidence in support or rejection of the research hypotheses related to the second research objective (section 3.3).

Each hypothesis and the relevant findings are now discussed.

8.3.1 Hypothesis 2.1

Hypothesis 2.1: *Similar modelling objectives are pursued by DES and SD modellers, which are related to understanding how the system behaves over time under different scenarios.*

From the content analysis performed in section 7.2, it was revealed that the DES and SD modellers considered to some extent the same modelling objectives. Most DES and SD modellers mentioned that the primary objective of the modelling exercise was to create a tool that projects the output of interest into the future. However, some SD modellers showed an inherent tendency to consider broader aspects of the problem. In addition, it was revealed that policy testing is an important aspect in SD modelling, while it was not prevalent in DES modelling. Hence hypothesis 2.1 cannot be rejected.

However, the mediating factors to these findings should be considered. An overall limited amount of verbalisations on problem structuring was produced by all modellers and hence most modellers did not make considerable efforts over problem structuring. The main reasons for the limited verbalisations were:

- the use of a structured modelling exercise and

- an overall clearly stated objective of the exercise.

Therefore, these findings provide some indicative evidence in partial support to hypothesis 2.1.

8.3.2 Hypothesis 2.3

Hypothesis 2.3: *There is not a set diagramming method used in DES modelling versus standard diagramming methods (causal loop and stock & flow diagrams) used in SD.*

With respect to creating conceptual diagrams, most DES and SD modellers involved in the model building study did not consider it a priority (section 7.3.1). Most DES and SD modellers conceptualised at the same time as coding the model on the computer. Even though most (3 out of 5) DES and (4 out of 5) SD modellers thought about creating diagrams, these did not resemble a standard diagramming method. It is therefore, concluded that hypothesis 2.3 cannot be accepted.

However, the observations made are only indicative and a final conclusion cannot be reached for a number of reasons:

- In two instances only one modeller from each group (who could be considered zealous in creating a conceptual model) referred to a process flow (DES) and a stock and flow (SD) diagram, respectively.
- In this study, a basic common conceptual diagram was already provided to all modellers, which could have limited to some extent their attention to creating a conceptual model.

- The findings are based on the accounts of 10 modellers only, who were selected to take part in the modelling sessions.

In order to reach more general conclusions, this aspect needs to be explored in the form of a survey with a larger sample of modellers. Hence, further work is required to certify this finding.

8.3.3 Hypothesis 2.4

Hypothesis 2.4: *Causal relationships and feedback structures are an essential part in SD models. In DES, while causal relationships and feedback can be included, they are not explicit.*

From the analysis on consideration of feedback (section 7.3.2), it is revealed that the SD modellers naturally referred to causal relationships and less often to feedback effects in the form of the effect of one variable on another, whereas, the DES modellers did not consider any equivalent structures. It is therefore, suggested that hypothesis 2.4 cannot be rejected. This implies that SD modellers naturally think about causal relationships and feedback, while DES modellers do not think about or refer to them. However, this result cannot be definitive for a number of reasons:

- Not all the SD modellers identified the main feedback effects, (section 9.2.1) in the model, which is a puzzling finding.
- It could be argued that the feedback effect in the prison population case study was not obvious to all the modellers. The closed path of the UK prison population stock and flow network with prisoners and recidivists returning

back to jail coupled with the 2 balancing smaller feedback loops did not represent a conventional reinforcing feedback loop (section 9.2.1). This may explain why most SD modellers downplayed the feedback effects in the prison system.

This provides useful considerations for future work on causal relationships and feedback effects. If a different case study was used, it would be interesting to find out whether similar results would be achieved.

8.3.4 Hypothesis 2.5

Hypothesis 2.5: *DES takes an analytic view of systems, whereas SD a holistic view.*

In the analysis carried out with respect to the systems' view taken by the DES and SD modellers (section 7.3.3) a clear difference was revealed. It was identified that when conceptualising the model, the DES modellers took the approach of breaking the system down into components (Lane, 2000), or the reductionist approach commented on by Han et al. (2005). In contrast, SD modellers took a more holistic perspective, thinking about the elements and their interconnections in the model, the wider factors involved in the prison system and referring to system archetypes (Senge, 1990). Similarly, it has already been mentioned that SD modellers tended to consider the wider issues involved when considering the objectives of the modelling exercise (section 8.3.1).

The views revealed in the study are consistent with the opinions expressed in the comparison literature, which are expressed in hypothesis 2.5. This implies that hypothesis 2.5 cannot be rejected. It seems therefore, that DES modellers take an analytic approach to model building, whereas SD modellers take a more holistic approach. However, one should consider whether this clear difference in the stance taken by the two groups of modellers is representative of the differences between the two modelling approaches or whether this is a result of the specific sample of modellers who took part in this study. Even though care was taken in choosing two comparable samples, comparing the length of experience among the individual modellers, the SD group included more subjects with greater experience. Hence, it should be considered whether the shorter length of experience for most DES modellers affected to some extent the lack of holistic thinking rather than the simulation approach used. However, it has been suggested that the two groups are reasonably comparable (section 6.2). A replication of the current study involving different subjects could verify the findings.

8.3.5 Hypothesis 2.7

Hypothesis 2.7: *DES modelling involves great complexity and detail whereas in SD, models tend to be general and abstract representations of the system, taking a wider focus as opposed to a narrow focus.*

The analysis performed in section 7.4.1 reveals that the models created by the DES modellers involved more detail as compared to the SD models. Some evidence of detailed thinking was identified in the DES protocols, due to the modellers' tendency to add further detail in the model, to think about the individual characteristics of the people and to use relatively small time units. Aspects of detailed complexity were also found in DES modelling arising mainly from the large population size and the inclusion of coding routines and conditional coding. In the SD protocols, the modellers considered aggregate aspects of the system. Furthermore, while detail complexity was not referred to, some limited evidence of dynamic complexity was identified. Based on the evidence found, hypothesis 2.7 cannot be rejected. It is therefore, suggested that DES involves more detailed complexity compared to SD, which leads to the development of more abstract models with a wider focus. However, these findings could have been affected by:

- the case study chosen or
- the sample of participating modellers.

If a different case study situation was utilised, different levels of detail and complexity could have been revealed.

8.3.6 Hypothesis 2.8

Hyp*othesis 2.8:* *In DES modelling there are no prior modelling structures for modelling decision-making processes, whereas in SD standard modelling structures exist, such as asset-stock adjustment. However, information from observations and discussions with decision makers (practical structures) can be added in DES models.*

Based on the analysis in section 7.4.2, the DES modellers did not refer to already existing modelling structures, whereas the SD protocols included some references to already existing modelling structures (either structures used in generic SD models or in previously built models) and to structures borrowed from other disciplines.

Even though some DES modellers mentioned that they would ask the client for information regarding missing data, they did not mention how this information would be embedded in their models.

It can be suggested that the use of already existing structures in SD modelling enhanced modellers' holistic thinking and their considerations of the wider aspects of the problem as identified in hypothesis 2.5 (section 8.3.4), as compared to the DES modellers, who did not make any references to such structures. Furthermore, this practice can help in the improvement of SD modelling skills, where experience in modelling enhances the ability to work in an additive pattern. This suggests that hypothesis 2.8 cannot be rejected. Therefore, it seems that SD modellers use already existing modelling structures, while DES modellers do not refer to them. Nevertheless, this is only an indicative finding determined to some extent by the participating modellers involved and the case study chosen.

8.3.7 Hypothesis 2.10

Hypothesis 2.10: *In DES modelling delays are represented in the form of delayed entities in the system, whereas in SD modelling delays are represented in the form of delayed influence or delayed processing of materials.*

The analysis in section 7.4.3 identified that the DES modellers were aware of the delays in the model, mostly represented in the form of entities (prisoners) kept/delayed in buffers or queues. On the other hand, most SD modellers did not specifically think about the delays in the model. Simple first order delays were most often used. Only one SD modeller used the pipeline delay function available in SD software. This implies that the use of delays is not a common practice in SD modelling. Modelling delays adds a discrete element in SD models (Sterman, 2000). This could explain the reason SD modellers, who tend to think in aggregate terms, avoid using higher order (2^{nd}, 3^{rd}, etc.) delays. These observations suggest that the use of delays is not a mainstream practice in SD modelling. However, this is an inference that needs to be further tested in future studies.

The evidence obtained from the analysis implemented, supports hypothesis 2.10, which therefore, cannot be rejected. Hence it is suggested that delays can be represented by both approaches, depending on the model features and the objectives of the model, whether the representation of discrete or aggregate aspect of delays is of interest. Yet, this finding is based on the result obtained using the case study chosen and the participant samples selected.

8.3.8 Hypothesis 2.11

Hypothesis 2.11: *In DES quantitative data are used which are obtained from concrete/observable processes, whereas in SD modelling quantitative and qualitative data are used, where in the absence of data, opinions and best guesses can be incorporated.*

As expected, the analysis in section 7.5.1 reveals that DES modellers think mostly in terms of concrete/observable data, whereas SD modellers think about softer and also abstract data. It can be concluded that SD modelling provides the modellers with the capacity to consider qualitative aspects of a problem situation. This also supports views expressed by Greasley (2005), that DES can provide a hard technical analysis of the problem, whereas SD can also consider softer issues involved. Regarding missing data, however, contrary to expectations, the analysis (section 7.5.1) reveals that modellers' reactions to missing data depend on their individual characteristics rather than the modelling approach used. Hence, hypothesis 2.11 cannot be accepted as it is only partially supported by the results obtained.

However, this finding could be a result of the specific study design:

- This study involved an experimental modelling exercise. In a real life modelling project, one would assume that modellers' reactions would be different.
- The prison population case study did not necessarily require the involvement of qualitative data. If a more qualitative case study was involved (including more qualitative aspects such as quality of life, incentives, psychological factors, etc.), one would expect to have observed different reactions.

Therefore, this needs to be further verified with the development of different case studies.

8.3.9 Hypothesis 2.12

Hypothesis 2.12: *Randomness is an essential part of DES models, whereas in SD modelling it is not as important.*

The analysis regarding randomness in the model (section 7.5.2) reveals that all DES modellers referred to randomness, whereas none of the SD modellers did.

Randomness was added in the DES models with the use of statistical distributions or conditional coding. In the absence of information on randomness, DES modellers' reactions varied, as already identified with hypothesis 2.11 (section 8.3.8). Therefore, in support of hypothesis 2.12 it can be concluded that indeed randomness is an important aspect in DES modelling, while it is not considered by SD modellers. Hence it is suggested that hypothesis 2.12 cannot be rejected. It should, however, be born in mind that this finding is based on the protocols derived from the experimental modelling exercise with the 5 DES modellers.

8.3.10 Hypothesis 2.13

Hypothesis 2.13: *In DES modelling linear relationships are more common. Non-linear relationships are more commonly used in SD modelling.*

In the analysis regarding the relationships between variables (section 7.5.3), it was observed that DES and SD modellers referred to and modelled linear relationships without particular issues. Non-linear relationships were only modelled by some SD modellers. The fact that some SD modellers thought about non-linear relationships, nevertheless, suggests that these are more naturally considered in SD modelling. Based on this evidence it can be suggested that hypothesis 2.13 cannot be rejected. However, it could be considered that this finding could have been a result of the case study used. Therefore, a test of this finding by utilising different case studies is suggested.

8.3.11 Hypotheses 2.14 & 2.15

Hypothesis 2.14: *DES modellers are concerned with developing an accurate model, close to reality, whereas SD modellers are concerned more with model usefulness.*

Hypothesis 2.15: *DES modellers take a 'black box' approach while validating the model, checking mainly the model outputs, while SD modellers take a 'white box' approach, checking the internal structure of the model.*

In the analysis implemented in section 7.6, it is revealed that the DES modellers checked the code to make sure it worked correctly. Due to the complexity involved in DES software (hypothesis 2.7) mistakes are not always obvious and thus more attention was paid to verification and white box validation as compared to the SD modellers. The SD modellers tended to check the model less frequently. Both the DES and SD modellers undertook black box validation, ensuring that the final results of the model were realistic. Contrary to expectations, the SD modellers did not speak about model usefulness as a way of validating the model. Hence it can be concluded that hypotheses 2.14 and 2.15 are not supported by the current study and it is suggested that they are rejected.

Nevertheless, the verification and validation strategy followed could have been a result of the nature of the study, where:

- An experimental model building exercise was involved.
- Modellers were asked to build a model working on their own, without contact with the client (model users), which is not normal in real life projects.

Therefore, these findings need to be triangulated with studies involving real life modelling projects.

8.3.12 Hypothesis 2.16

Hypothesis 2.16: *DES modellers are interested in steady state conditions whereas SD modellers are interested in achieving a model in equilibrium.*

The analysis regarding the practice of achieving a DES model in a steady state or an SD model in equilibrium (section 7.7.1) reveals that DES modellers subconsciously thought about creating a model in a steady state as part of dealing with the initialisation bias, while SD modellers thought about creating a model in equilibrium. Removing the initialisation bias is a standard practice in DES modelling, which ultimately serves to set the model to representative initial conditions (Pidd, 2004; Robinson, 2004). SD modellers created a model in equilibrium by changing the levels of the stocks or other system parameters (variables). It is, however, concluded that the concept of equilibrium in SD differs to some extent from that of 'steady state' in a DES model. While in SD a system in equilibrium means that the effects between variables stabilise each other and thus resulting in straight lines of outputs, in DES, steady state does not mean that the outputs are not varying (i.e. achieving straight lines), but that the output is varying according to some fixed distribution (Robinson, 2004, pp. 138-139). This suggests that while at first sight it can be considered that DES and SD models reaching an equilibrium or a steady state condition are similar, in fact, these involve slightly different concepts. Hypothesis 2.16 cannot be rejected.

8.3.13 Hypothesis 2.17

Hypothesis 2.17: *The results of interest from a simulation model are quantitative point predictions and the optimisation of performance criteria for DES modellers, while SD modellers are interested in quantitative and qualitative results and point predictions are rarely made.*

The analysis in section 7.7.2 reveals that the DES and SD modellers were similarly interested in obtaining numerical data from the model, but the DES modellers tended to take a more detailed view. As for their interest in qualitative results, it is found that the DES and SD modellers thought alike regarding the qualitative aspect of the model outputs. The only difference was that the SD modellers considered them more often. According to hypothesis 2.17, DES modellers were also expected to show an interest in the optimisation of model performance, but this aspect is not revealed in most of the DES protocols.

This suggests that hypothesis 2.17 can be considered as partially not true and it is, therefore, rejected. However, some mediating factors to this finding should be considered:

- In this study, the analysis is only limited to the considerations made by DES and SD modellers about the qualitative and quantitative results of the model based on the prison population case study.
- The study focused on model building and not on the generation of solutions and understanding.

This outcome could have been different, if there was more scope in the modelling exercise for modellers to make a more elaborate consideration of the results.

8.3.14 Hypothesis 2.18

Hypothesis 2.18: *DES and SD modellers use scenarios to the same extent to compare the performance of alternative system configurations.*

From the analysis (section 7.7.3) it is disclosed that the development of scenarios came more naturally to the SD modellers. Hence this suggests that hypothesis 2.18 is not supported by the findings and it can therefore, be rejected. This suggests that SD modellers are more geared towards experimenting with the model.

These finding should however, be considered with caution for a number of reasons:

- The nature of the modelling exercise involved in the study. The analysis and development of scenarios were not the focus of the task.
- Due to the limited time, the modellers were not expected to spend much time with this stage of model building. The DES modellers specifically, who spent longer to build the model, had less time left to experiment with the model.
- Experimentation and development of scenarios can be considered part of model use rather than model building.

A future study that concentrates more on the model results and experimentation of scenarios is recommended.

8.3.15 Further observations

In addition, some further observations are made as part of the model building study, which were not related to any of the research hypotheses. These, however, are considered to be worth mentioning.

Choice of simulation approach

From the modelling sessions carried out, it is observed that most modellers did not consider alternative modelling approaches before modelling. This conveys that modellers use the simulation approach they are more familiar with rather than the approach that is considered most appropriate.

Model interface

Another issue that emerged during model building sessions is the attention paid by some DES and SD modellers in developing a suitable model interface for the client (section 7.4.4). Most DES modellers use in practice an Excel interface, as the front-end environment on which model inputs are entered and model outputs are displayed. In a similar fashion, SD modellers create a user interface for the client to interact with the model and its results. However, it is fair to mention that more developed interface facilities are available in SD software. Some DES software have facilities for the development of user menus, which can be included inside the model (e.g. Witness), but these were not used or referred to by any of the DES modellers.

Even though DES modellers did not attempt to create a user interface in the models developed, they all showed an interest in it and mentioned that it is an important aspect of the model in real life projects. Some SD modellers considered the user interface more extensively. A possible reason for this could be the fact that the SD model was easier to create compared to the DES models and therefore, they had more time available. The DES modellers considered the user interface as a tool that helps the client's interaction with the model, where the client interacts with the model by changing its inputs and outputs. The SD modellers created sliders and switches with variables that the policy makers would be interested to experiment with, but they did not refer at all to how the users would interact with the model and its structure.

8.3.16 Summary of the findings from the qualitative analysis of VPA

In this section, the findings of the qualitative analysis of the VPA are discussed, based on the relevant research hypotheses. A summary of the hypotheses considered in the qualitative analysis and the outcomes obtained is provided in Table 8-2. It can be concluded that DES and SD modelling are different modelling approaches with respect to the model building approach taken. Some similarities are found regarding the modelling objectives considered, conceptual modelling diagrams and their attitude towards missing data. The main differences are concerned with feedback, systems' view, model complexity and detail, prior modelling structures, representation of delays, randomness, verification and validation ('black-box' and 'white-box' validation) and views about reaching a steady state or equilibrium condition.

Table 8-2: Summary of the findings of the VPA on the prison population case, by hypothesis

Hypotheses on model building process	Research findings
Hypothesis 2.1: Similar modelling objectives are pursued by DES and SD modellers, which are related to understanding how the system behaves over time under different scenarios.	Not rejected
Hypothesis 2.2: DES models problems at tactical/operational level, while SD at a strategic level.	Not tested (depends on nature of the problem situation)
Hypothesis 2.3: i) There is not a set diagramming method used in DES modelling versus, ii) standard diagramming methods (causal loop and stock & flow diagrams) used in SD.	Rejected (overall hypothesis) Partial hypotheses: (i) not rejected (ii) rejected
Hypothesis 2.4: Causal relationships and feedback structures are an essential part in SD models. In DES, while causal relationships and feedback can be included, they are not explicit.	Not rejected
Hypothesis 2.5: DES takes an analytic view, whereas SD a holistic view of systems.	Not rejected
Hypothesis 2.6: In DES modelling objects/people are represented as distinct individuals while in SD aggregate system behaviour is examined.	Not tested (fact)
Hypothesis 2.7: DES modelling involves great complexity and detail whereas in SD models tend to be general and abstract representations of the system, taking a wider focus as opposed to a narrow focus.	Not rejected
Hypothesis 2.8: In DES modelling there are no prior modelling structures for modelling decision-making processes, whereas in SD standard modelling structures exist, such as asset-stock adjustment. However, information from observations and discussions with decision makers (practical structures) can be added in DES models.	Not rejected
Hypothesis 2.9: DES models represent mainly material flows. Information flows can be incorporated but these are not obvious. In SD modelling both material and information flows are equally represented	Not tested (depends on nature of the problem situation)
Hypothesis 2.10: In DES modelling delays are represented in the form of delayed entities in the system, whereas in SD modelling delays are represented in the form of delayed influence or delayed processing of materials.	Not rejected
Hypothesis 2.11: i) In DES quantitative data are used which are obtained from concrete/observable processes, whereas in SD modelling quantitative and qualitative data are used, ii) where (in SD modelling) in the absence of data, opinions and best guesses can	Rejected (overall hypothesis) Partial hypotheses: (i) Not rejected

be incorporated.	(ii) Rejected
Hypothesis 2.12: Randomness is an essential part of DES models, whereas in SD modelling it is not as important.	Not rejected
Hypothesis 2.13: In DES modelling linear relationships are more common. Non-linear relationships are more commonly used in SD modelling.	Not rejected
Hypothesis 2.14: i) DES modellers are concerned with developing an accurate model, close to reality, ii) whereas SD modellers are concerned more with model usefulness.	Rejected (overall hypothesis) Partial hypotheses: (i) Not rejected (ii) Rejected
Hypothesis 2.15: DES modellers take a 'black box' approach while validating the model, checking mainly the model outputs, while SD modellers take a 'white box' approach, checking the internal structure of the model.	Rejected
Hypothesis 2.16: DES modellers are interested in steady state conditions whereas SD modellers are interested in achieving a model in equilibrium.	Not rejected
Hypothesis 2.17: i) The results of interest from a simulation model are quantitative point predictions and the optimisation of performance criteria for DES modellers, ii) while SD modellers are interested in quantitative and qualitative results and point predictions are rarely made.	Rejected (overall hypothesis) Partial hypotheses: (i) Rejected (ii) Not rejected
Hypothesis 2.18: DES and SD modellers use scenarios to the same extent to compare the performance of alternative system configurations.	Rejected

8.4 Summary of discussion

The current chapter provides a discussion of the results obtained from the quantitative and qualitative analysis of the 10 verbal protocols. The reasons behind the findings are sought and explanations are provided in an attempt to understand the differences and similarities between the DES and SD model building approach. The mediating factors, which could have affected the validity of each specific finding, are also considered. Overall, it is concluded that DES and SD modelling are two simulation approaches that share some common features, there are however, differences regarding the modelling approach taken. In the next two chapters, two specific simulation models are compared based on the users' opinions.

Chapter 9: Research design II – study of model use[14]

9.1 Introduction

In this chapter the design of the study of DES and SD model use is described. The objective of the model use study is to test the already specified hypotheses regarding the use of DES and SD models (section 3.4.3), based on the statements found in the literature and the third research objective. For the purposes of this study, a questionnaire survey is used (section 3.5.2), which deals with the comparison of DES and SD models by assessing the differences and similarities based on users' opinions in terms of their use and the learning derived.

The current chapter provides an account of the activities undertaken for the design of the survey. It describes the two simulation models, one DES and one SD created, based on the case study already presented in chapter 4. Next the questionnaire (as the measurement instrument used) is outlined, followed by the selection of the participant sample, the pilot study undertaken and a description of the model use sessions. The survey design was an iterative process, with interdependencies between its different aspects: case study design, model building and survey design.

[14] This chapter is based on Tako A.A. & Robinson S. (Forthcoming) "Comparing discrete-event simulation and system dynamics: Users' perceptions" Journal of Operational Research Society, 17pp, doi:10.1057/palgrave.jors.2602566. Accepted.

9.2 The DES and SD models

The comparison of model use focuses on users' experiences from using DES and SD models, so two separate simulation models using the two simulation techniques are developed. A brief explanation of the two models follows, including the similarities as well as the differences involved.

9.2.1 Model overview

A DES and a SD simulation model were built. Each simulation model is a simple representation of a prison overcrowding problem showing how the prison population evolves over time. The main objective is to build two simple models which enable experimentation with different scenarios/policies, with the intention of using them as tools for decision making. The DES model was developed using WITNESS (www.lanner.com accessed September 2008), a powerful and versatile DES simulation package. For the SD model, Powersim Studio 2005 (www.powersim.com accessed September 2008) was used. This is a package used widely in the field of SD. Both models incorporate a user interface which enables inputs to be set and altered. Witness and Powersim are typical of the simulation software in their respective fields. Although there is some variation in the facilities in alternative packages, there is no specific reason to believe that the choice of package would have much influence on the representation of a simple model such as the prison population case used in this research (Robinson, 2008).

The basic conceptual model, on which both SD and DES models are based, is displayed in figure 9.1. The most important feature of the model is the dominant closed-path of: People in Prison → Release Rate → Released Prisoners→ Re-offend Rate (Recidivists[15]) → People in Prison. As a matter of fact, an increase in the number of first-time offenders causes an increase in the number of people in prison which leads to an increase in the release rate which affects positively the number of released prisoners after a given time period. The increase of released prisoners affects positively the number of rehabilitated prisoners and also recidivists, which in itself causes an increase in the number of people in prison and the circle goes on. In this structure two smaller in scale negative feedback loops are also present, resulting from the interrelationships between People in Prison and Release Rate, and the negative feedback loop between Released Prisoners and Recidivists. These negative loops balance the effect of the dominant positive loop in the system.

[15] People who re-offend after coming out of prison.

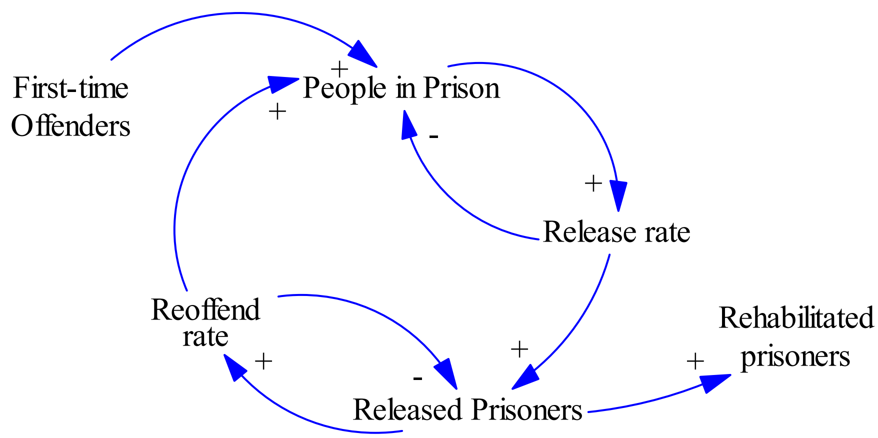


Figure 9-1: Feedback effects in the prison system

9.2.2 Common model features

Building equivalent models was very crucial for the design of the survey. The whole process was iterative revolving around case study design and model building. Initially prototype models were built based on the case study concept and later these were refined as the case study was further developed. At the same time the model building process helped in finalising the case study material.

The preliminary DES and SD models were built simultaneously to enable the comparison of the results and to align model inputs. After finalising the preliminary models, the experimental factors used in the case study were identified. During the modelling process, it was important to remain unbiased between the two simulation approaches given that the worldview is different for a system dynamicist and a DES modeller (Morecroft and Robinson, 2005). For this reason, experts from both areas were consulted, which resulted in a quite fruitful and interesting experience.

While building the two simulation models, the main concern was to develop two equivalent models in terms of inputs and outputs. For this reason the main focus was in aligning inputs and outputs for both models. The underlying time unit was years for the SD model and days for the DES model. However, in the DES model the outputs displayed were in years too, as the model converts internally the time unit to years. The main reason for using different time units for each model was the underlying difference in modelling prisoner admission rates for the two simulation approaches. In SD, prison admissions are represented by admission rates, which represent the number of prisoners entering prison per time unit, e.g. year. In DES prison admissions are represented by parts or entities, which enter the system every n-time units as specified by the inter-arrival time, which is usually represented with the help of a statistical distribution, e.g. negative exponential. So in the DES case, the time between two consecutive arrivals of prisoners needs to be specified. Thus the time unit used was days. In table 9-1, equivalent model parameters used in both simulation models are displayed.

Table 9-1: Model parameters – SD & DES models

	SD model	DES model
Time unit	Year	Day (1 year = 264 days) 1 month = 22 (working) days
Entry point – Petty offenders	3,000/year	Equiv. to 264/30*=8.8 ➔ Petty Mean Arrivals = 8.8 prisoners/day Inter-arrival time: negexp(8.8, 1)

	SD model	DES model
Entry point – Serious offenders	650/year	Equiv. to 264/6.5*=40 prisoners ➔ Serious Mean Arrivals = 40 prisoners/day Inter-arrival time: negexp(40, 2)
Run time	30 years	7,920 days (**30 years** x 264 days)
Initial value for petty offenders	50,000	500*
Initial value for serious offenders	26,000	260*
Initial value for petty recidivists	30,000	200*, +
Initial value for serious recidivists	10,000	26*, +
Time in prison for petty	5 years	1,320 days (**5 years** x 264 days) Length of sentence = Erlang(1320,3,3)
Time in prison for serious	20 years	5,280 days (**20 years** x 264 days) Length of sentence =Erlang(5280,3,7)
Petty Time to re-offend	2 years	528 days (**2 years** x 264 days) Time to re-offend = Erlang(528,3,200)
Serious Time to re-offend	2 years	528 days (**2 years** x 264 days) Time to re-offend = Erlang(528,3,204)

*Note: DES model figures represent 1/100th of SD figures, as 1 entity represents 100 prisoners.

+ Note: The scaling difference of the initial values for petty and serious recidivists is due to the different ways recidivists are handled in the DES and SD models. Different values are required to achieve a 'steady-state' and 'equilibrium' for the DES and SD models.

9.2.3 The DES model

In the first DES models created, the sentencing activity was included after the offenders' entry point. From there, some of the offenders exited the system because they were deemed not guilty and the rest were sent to prison. In the need to keep the model simple, this part was omitted, by only modelling offenders who receive a prison sentence. The process flow diagram for the DES model is provided (Figure 9-2).

As part of the modelling stage, a number of issues had to be resolved. For the queues Prison and Recidivists, it was necessary to set up the initial number of people in prison and the initial number of recidivists at the beginning of the simulation, which is on 1st January 2005. This was finally resolved by adding dummy offenders and dummy recidivists respectively, who enter the system only at the beginning of the simulation run. Due to the large numbers used, it was apparent that the DES model was very slow. At this point, the Batch run feature in Witness, which runs the simulation faster, but without any visual representation, came into use. Due to the nature of DES, dealing with individuals, the large numbers involved affect the simulation speed due to the large number of events scheduled in the next event list. In order to deal with this issue, one individual entity was modelled so that it represents 100 prisoners.

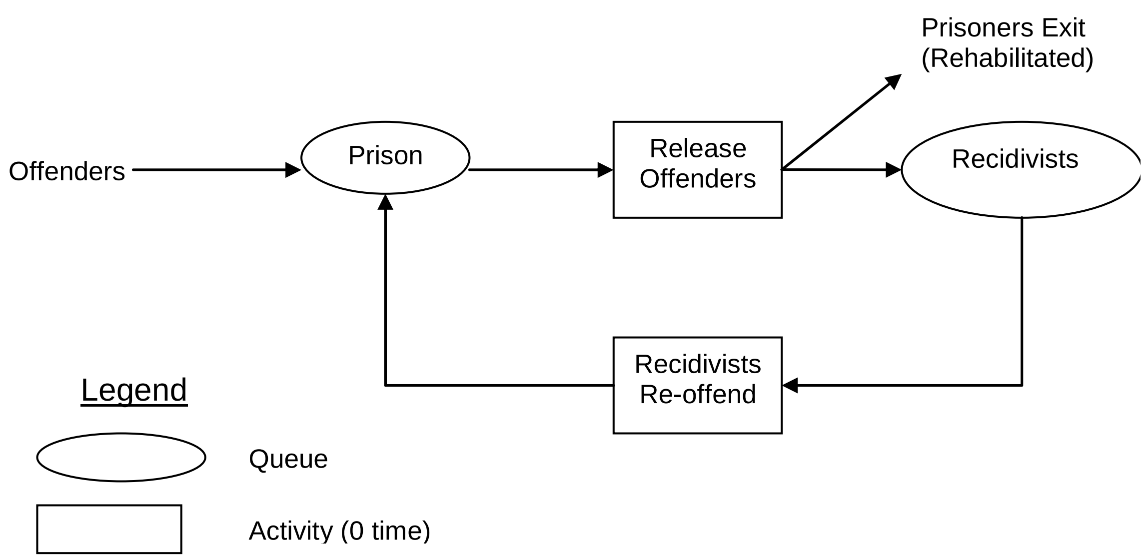


Figure 9-2: Process Flow Diagram - DES prison model

The simulation environment of the final DES model is presented in figure 9-3. The model environment includes a number of different windows which consist of: the model (the box on the left), the input data (on the top, right-hand corner) and the model outputs (at the bottom, right-hand corner). Control buttons (i.e. run, stop, reset etc.) are included in the toolbar (My toolbar) and also a window which reports the time. On the click of the run button, a window appears requesting the users to choose the relevant input data according to the policy or policies chosen. The user can also access a series of graphical results by selecting the graphs button. These include plots of the prison population, plots of rehabilitated prisoners and plots of the recidivists over time, and also bar charts with the distribution of sentence lengths for both petty and serious offenders (Figure 9-4).

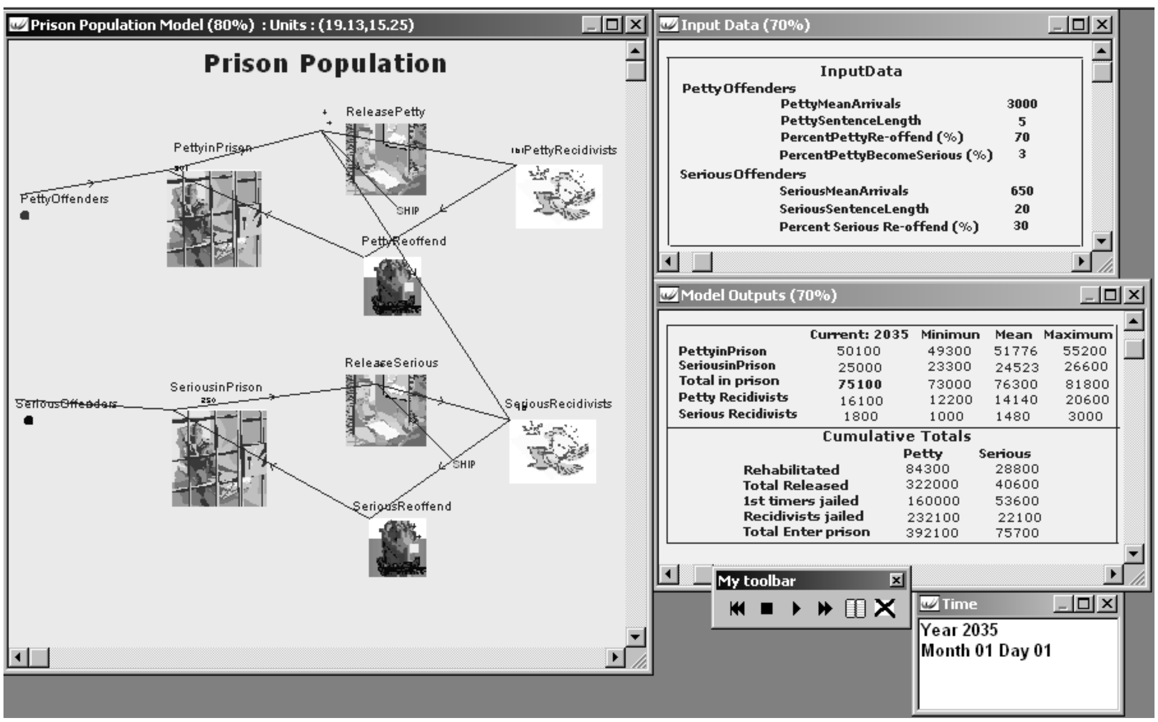


Figure 9-3: DES model representation in Witness, with the model on the left-hand side, input criteria in the top box on the right and in the box below model outputs.

In the DES model, entities enter the system and two attributes are set, 'sentence length' and 'time to re-offend'. The entities then go straight to queues which represent the prison population, either as 'PettyinPrison' or as 'SeriousinPrison'. In these queues, they are delayed according to the attribute 'time in prison'. Prisoners then go into the release activity ('ReleasePetty' or 'ReleaseSerious') from where they are either rehabilitated and exit the system ('Ship') or go to the recidivist queues, ('PettyRecidivists' or 'SeriousRecidivists') according to the seriousness of their offence. In the recidivist queues, the entities stay according to the attribute 'time to re-offend' and then go to the re-offend activities ('PettyReoffend' or 'SeriousReoffend'), where the attribute 'sentence length' is reset. From there the entities re-enter prison.

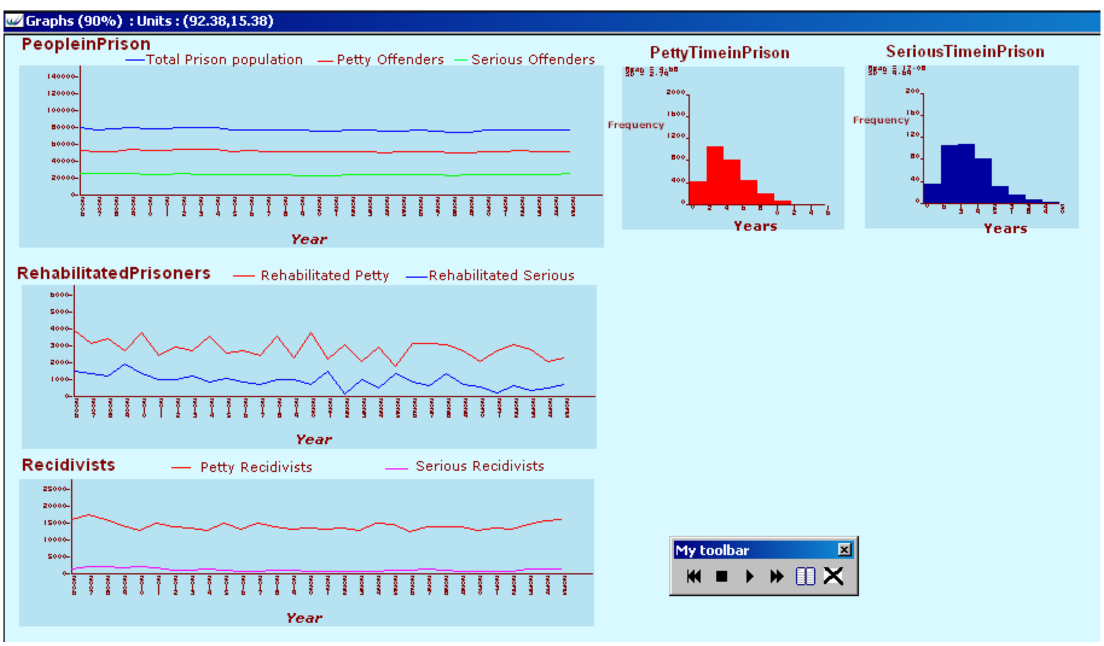


Figure 9-4: DES model outputs

9.2.4 The SD model

In initial SD models people in prison and recidivists waiting to re-offend were modelled by using the delay functions. According to Sterman (2000), “a delay is a process whose output lags behind its input in some fashion”, which seemed to justify the use of delays. Therefore, the function of a first order material delay was included in the equations. However, the outputs derived by running the model were incomprehensible, providing negative numbers for people in prison. It was therefore, decided to use a simple average (exponential) delay equation, dividing the stock by the time to re-offend. (During the model building study, it was later on pointed out by SD3 that in this case a pipeline delay would be most suitable instead.) In system dynamics delay functions are used to represent discrete elements in the model, therefore, in order to maintain a purely aggregate behaviour of the system, it was decided not to include delay functions in the model.

The model was re-formulated by including two separate stocks: ‘People in Prison’ and ‘Released Prisoners’ with their respective outflows. From the stock ‘People in Prison’, the outflow of prisoners depends on the variable ‘Time in prison’. ‘Released Prisoners’ are accumulated and from this stock there are two outflows, ‘Rehabilitated’ and ‘Recidivists’ (re-offending prisoners). The two outflows of prisoners from the stock ‘Released Prisoners’ depend on the variables ‘Proportion Rehabilitated’ and ‘Proportion Re-offend’ respectively, and the variable ‘Time to re-offend’. Recidivists entering prison serves as an inflow for the stock ‘People in

Prison'. The new causal loop diagram, including the respective positive or negative relationships involved, for the SD model is displayed in figure 9.5.

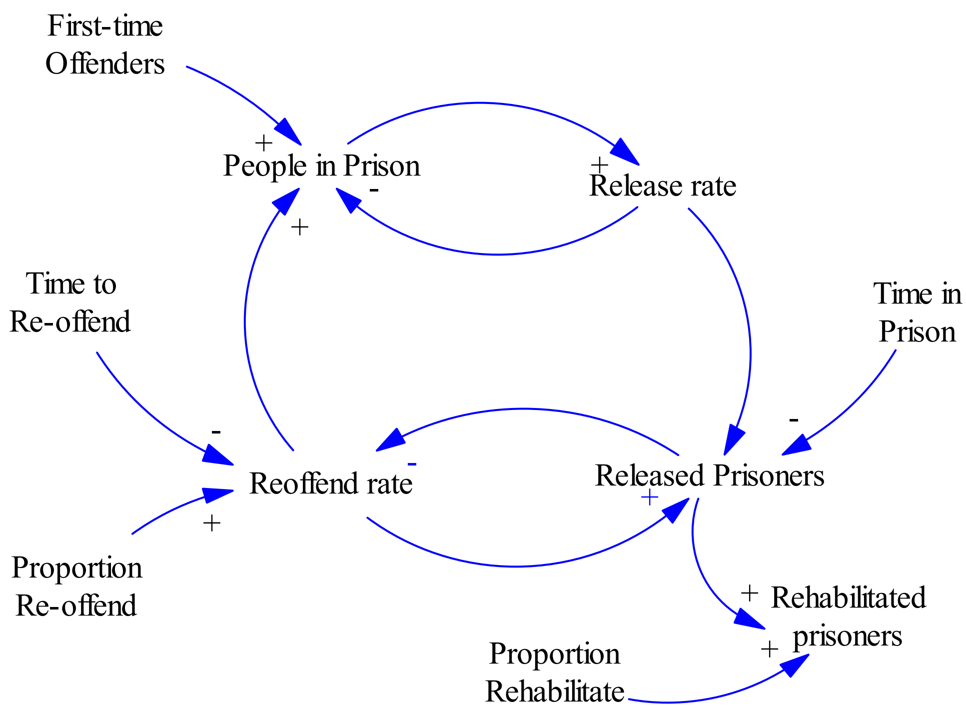


Figure 9-5: Causal loop diagram – SD prison model

The final SD model consists of 4 different pages: introduction, control panel, prison population diagram and the main model. The pages are linked via hyperlinks so that users can easily navigate from one page to the other. The SD model representation is shown in figure 9.6. Two separate flows, petty and serious admissions enter the system and go straight into the prison population stocks ('Petty criminals in prison' and 'Serious criminals in prison'). Prisoners flow out of prison through the outflows, ('Petty Release rate' and 'Serious release rate') to the stocks 'Released petty' and

'Released serious'. Prisoners leave the released prisoner stocks either as rehabilitated prisoners or re-offenders, the latter creating a feedback loop to the prison population stocks. The stock 'Released petty' has an additional outflow, 'Become Serious', which takes a small part of the released prisoners straight to the stock 'Serious criminals in prison'. The structure resulting from the stock and flow network of released and re-offending prisoners is provided in Figure 9-6.

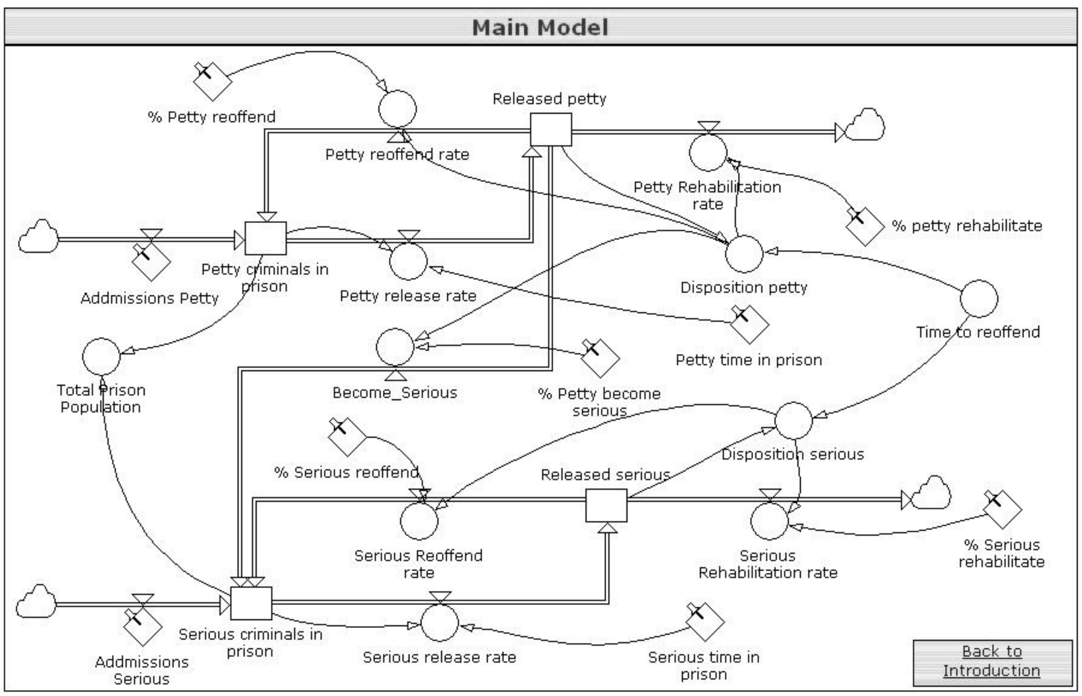


Figure 9-6: SD model representation in Powersim.

The Control Panel (Figure 9-7) is the main working environment where users can interact with the model and enter inputs according to their choice of policy or policies and observe relevant outcomes. The control panel consists of two parts. The user interface includes a set of sliders for the prison admission rates and the sentence time, and three combo boxes which provide choices for the percentage of re-offending. Next to the user interface are the model results, which consist of a set of graphs and tables of key outputs that are simultaneously updated. A larger view of the same graphs is also available on a separate page, which users can access if they click on the link “Enlarge graphs”.

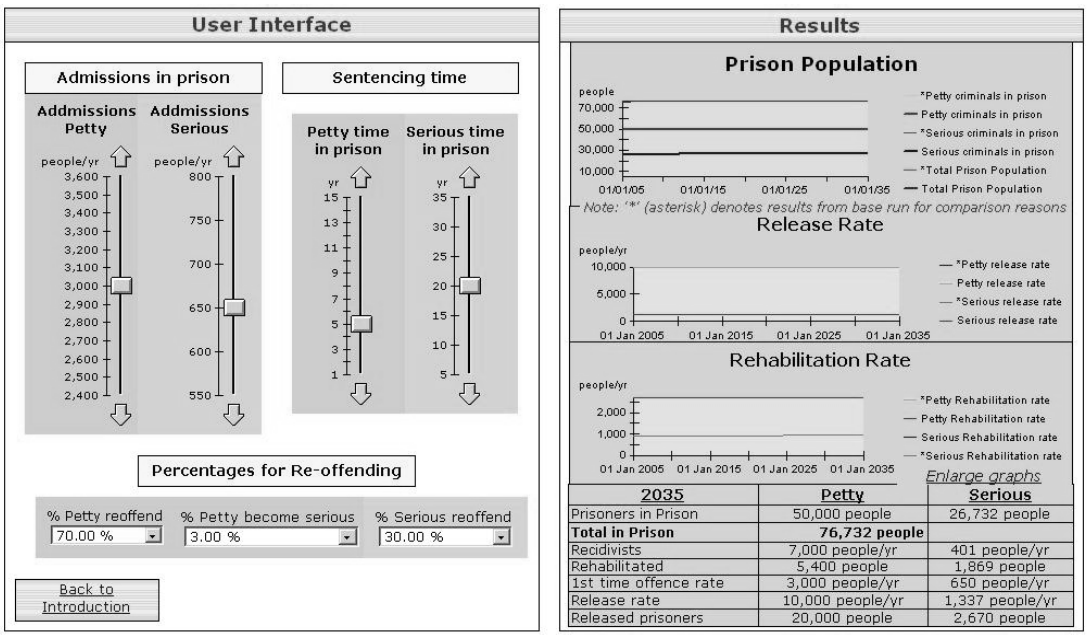


Figure 9-7: SD model Control Panel, which included the user interface and model results page. This was the main working environment.

9.2.5 Differences between the two simulation models created

Some key differences can be observed in the DES and SD models presented above. In the DES model the entities are individually represented and specific attributes assigned to them, i.e. sentence length, offender type, number of times incarcerated etc. Due to the large number of entities the run speed of the model becomes very slow and so the numbers have been scaled down to a fraction of 1/100, where 1 entity represents 100 offenders (section 9.2.2). Grouping entities is a well known practice in DES modelling (Robinson, 2004). Therefore, it can be claimed that there is some level of aggregation involved in the DES model. However, the main feature of DES, which enables the tracking of entities (in this model the group of 100 prisoners) and their attributes is still present. After all, one of the main reasons DES is chosen in practice is its capacity to track individuals in the system.

On the other hand, in the SD model the entities are presented as a continuous quantity, where state changes happen continuously at small segments (Δt) of time[16]. Specific entities cannot be followed through the system. Therefore, it can be claimed that there is a higher level of aggregation in the SD model than in the DES model. Modelling the large number of people in the system does not require any specific handling in SD, which is naturally suited to dealing with large populations.

Key variables in the DES model are sampled using the exponential or Erlang distributions, e.g. admissions to prison, time to re-offend and sentence length. In this way randomness is incorporated into the model. On the contrary, in the SD model these same variables are represented as deterministic average values.

[16] In the SD model, the time-step (Δt) used is 1 year (see table 9-1)

Another difference in the two model representations is related to how the initial number of people already in the system is set up. Powersim and all SD software packages have a facility for setting up the initial stocks at the beginning of the simulation run. In Witness, however, there is no such facility for queues. Two options were available. The model could be run for a warm-up period to allow the system to fill-up to the desired level. The other option was to create dummy entities which enter the model at the start of the simulation run and are assigned to the various queues. The latter option was considered more appropriate, as a warm-up period would have added significantly to the run time of the model and it would have been less intuitive for the users. Because the DES model collects results on the individual entities, each dummy entity had to be given a history of when it had entered the model, otherwise the results would have been skewed. This was achieved by sampling negative times of entry.

A conceptual difference between the two models is the way that released prisoners are dealt with. In the DES model released prisoners who do not re-offend leave the system straight away after being released. In contrast, in the SD model all released prisoners are kept for 2 years in the released stocks and after that a proportion of the stock flows out of the model. The difference arose because in the SD model it is necessary to accumulate all released prisoners into a stock before determining what happens to them next.

Regarding the data requirements, both models required almost the same data inputs. The DES representation more closely resembles the real life system, with variables set (approximately) to the values as described in the case by Grove et al.(1998). However, in the SD model, variables which do not exist in real life were created in order to represent intended behaviours. For instance, some variables were created such as disposition (‘Disposition Petty’ and ‘Disposition Serious’), in order to obtain the correct proportions of re-offending and rehabilitation. Disposition calculates the release rate for all freed prisoners who remain for two years in the stocks of released prisoners, before calculating the rates of re-offending and rehabilitation. In this respect, it seems that SD has more flexible structures.

Despite the differences discussed here, both models depict almost similar behaviour and the key outputs are quite similar (Table 9-2). There are probably some differences in variable definitions from one model to the other, and thus some differences in the results. For example, in the DES model the cumulative number of released prisoners (petty and serious) is displayed in the outputs, while in the SD model the number of released prisoners (petty and serious) at liberty at a specific point of time is displayed. In addition, in the DES model the number of recidivists represents the number of released prisoners at liberty in the community who will re-offend at some point in the future, whereas in the SD model, this number represents the rate of re-offending, that is, the number of prisoners who re-offend annually. Despite these differences the two models are fundamentally equivalent.

Table 9-2: Comparison of DES and SD models outputs.

	DES outputs	SD outputs
Petty in Prison	50,100 people	50,000 people
Serious in Prison	25,000 people	26,732 people
Total in Prison	75,100 people	76,732 people
Petty Recidivists	16,100 people	14,000 people
Serious Recidivists	1,800 people	801 people
Released Petty	-	10,000 people/year
Released Serious	-	1,337 people/year
Total Released Petty	322,000 people (equiv. 10,730 people/year)	-
Total Released Serious	40,600 people (equiv. 1,353 people/year)	-

9.3 Construction of questionnaire survey

The construction of the questionnaire is an important aspect of the survey design (Hutchinson, 2004). A well developed questionnaire needs to directly address the goals of the survey. The overall quality of the research depends directly on the quality of the questions asked (Buckingham and Saunders, 2004) and answers received (Hutchinson, 2004). The rule: "keep it simple and short" (Buckingham and Saunders, 2004) was taken into consideration when designing the questionnaire. Therefore, it was aimed to create a short questionnaire survey, about two pages long, including concise and carefully worded questions to ensure a good understanding and response rate by the survey participants.

The four main elements that require attention during the construction of the questionnaire survey are: determining the questions to be asked by the study, selecting question type for each question, design of the question sequence and overall questionnaire layout. The questions included in the questionnaire are related to the third research objective, which aims to find out how different DES and SD models are from the user's point of view. Hence the specific research questions are based on the research hypotheses, derived from the statements found in the literature, which the current study aims to confirm/refute (section 3.4.3).

The questionnaire consists of two main parts. The first part deals with participants' personal details. In order to ensure confidentiality, the questionnaire does not require participant or company names. The personal details include questions about the industry the respondents work in, the functional area within the organisation and their level of management in the organisational hierarchy. In addition, participants are asked about their prior experience in using a simulation model and if the answer is "yes" they are required to identify the simulation package used. From the name of the simulation package, the researcher can identify the type of simulation technique (DES, SD or other) the participant has used in the past. These questions can be used as control variables (filtering criteria) during data analysis to identify whether any of the personal details affects respondents' answers to the second part of the questionnaire (Buckingham and Saunders, 2004).

The second part of the survey deals with the participants' opinions about the two equivalent DES and SD simulation models used during the session as part of the exercise. The second part of the survey is divided into five main sub-sections regarding participant opinions about the following criteria:

a. Model understanding and complexity

b. Model validity

c. Model usefulness

d. Simulation results

e. Overall model

The model use study is an innovative study in the simulation area looking into managers' perceptions of DES and SD simulation models. Pre-conceived measures on simulation model use were not found in the simulation literature. Therefore, the measures used in this study are created from scratch, based on the researcher's experience as a modeller, on consultations with her supervisor (who is an expert in the area) and on the statements made in the literature concerning DES and SD. More specifically the questions included in the second part of the survey are displayed in table 9-3, by research question/ hypothesis and questions included in the final version of the questionnaire.

Table 9-3: Determining questions to be asked

Research question	**How different DES and SD models are in respect to the criteria:**	
Criteria	**Hypothesis**	**Question focus**
Understanding derived from using the models	**Hypothesis 3.1:** Despite the use of animation, DES models are not transparent to the client, and so difficult to understand compared to SD models, which are transparent to the client.	- Overall understanding of the models (Q. a.1) and their parts: relationship between variables, model structure, model results and how to use the model (Q. a.2.a-d). - The factors that affect model understanding (animation, the description or visual display) (Q. a.3)
Perceived	**Hypothesis 3.2:** Different	- What level of detail is involved in the

model complexity	levels of complexity are involved in DES and SD modelling, resulting from the different levels of detail and perceptions of feedback.	models (Q. a.4) - Is the feedback obvious to the users as a source of complexity? (Q. a.5)
Model validity	**Hypothesis 3.3:** Users find DES and SD models equally credible for giving answers to a problem situation.	- Model representativeness: Is the model representative of the problem described in the case study? (Q. b.1) - Realistic outputs: Does the model generate realistic outputs? (Q. b.2) - Confidence in the model: Do the users feel confident in using the models for decision making? (Q. b.3)
Perceived model usefulness	**Hypothesis 3.4:** DES and SD models are equally helpful as learning and communication tools. **Hypothesis 3.5:** SD models can aid strategic thinking to a higher extent.	- Learning: To what extent the use of the model enhances learning about the various policies (Q. c.1) - Communication of ideas: To what extent users feel that the use of the model facilitates the communication of ideas? (Q. c.4) - Strategic thinking: to what extent using the model helps the user think strategically about the problem. (Q. c. 2) Does knowledge transfer (Morecroft and Sterman, 1994) as a form of learning take place? (Q. c.3)
Result interpretation	**Hypothesis 3.6:** The distinct nature of results derived from DES and SD models result in differences in learning & interpretation of results. - DES aids instrumental learning, while SD conceptual learning. - The interpretation of DES model results is more difficult compared to the SD model. - Different aspects of the models are picked up by the users: Randomness is explicit in DES model results compared to the deterministic nature of SD models.	- Are users driven to use numbers or pictures (graphs) when examining model results? (Q. d.1) - How difficult do users find the interpretation of DES and SD models results? (Q. d.2) - How useful do users find the graphs in the model and what do they learn from them? (Q. d.3) - While examining the results, do users look for the factors that change the results in the model? (Q. d.4)

The main question format used to collect users' opinions on the models was a 5-point Likert scale, ranking from 1 to 5, giving an ordinal, non-metric measurement. The 1 to 5 response scale is commonly used in social science research (Buckingham and Saunders, 2004). Other types of questions included are rank order/multiple choice questions, single select (yes/no) questions and open ended questions. The latter were included to avoid leading participants' answers, despite the fact that open questions have a lower response rate. This is considered to be a trade-off for better quality information. A multiple-choice question is used in order to collect information on ranking among factors that aided model understanding (Q. a.3). The ranking scale is deemed as suitable in this case because the aim is to find out the factors that aid model understanding. A single select question is used in Q. d.1 asking for the type of results (numerical or graphical) participants mostly used. The aim of this question is to find out which type of learning is associated with each simulation approach, instrumental versus conceptual learning, i.e. learning from graphs or tables with numbers. So asking users to select one of the two output options they mostly used is considered sufficient.

The initial version of the questionnaire survey was revised a few times, after consultations with colleagues and especially after running the pilot study (section 9.5). The pilot study provided useful insights in improving the survey and in making sure that the questions would be clear to the participants.

9.4 Survey participants

As part of the current survey the simulation models are evaluated from the users' point of view. In any organisation it is the managers who are the ultimate users of a simulation model, whether it be directly experimenting with the model or as recipients of the results. In the latter case the manager would normally interact with the model to, at least, gain some confidence in the results. Managers, therefore, were considered the most relevant participants for the purposes of this study. Since ready access to executive MBA students at Warwick Business School was available, these were chosen as the subjects of the study.

The executive MBA students at Warwick are highly representative of managers working in the public and private sector. They have on average 12 years of work experience (www.wbs.ac.uk/students/mba/learn/class-profile-mod.cfm accessed September 2008) and at the same time as studying are holding managerial positions in their organisations. During the first year of their studies they take a core module, Modelling and Analysis for Management (MAM)(Robinson et al., 2003), on which the researcher's supervisor teaches. The executive MBA students work in a wide range of organisations, with 56% working in International organisations, 20% national, 13% small to medium sized enterprises and 11% work in public or not for profit organisations[17].

[17] Statistics taken from WBS website, Profile of a typical modular study mode class on the Warwick MBA, www.wbs.ac.uk/students/mba/learn/class-profile-mod.cfm (accessed on 2nd September 2008).

The study was implemented with two different groups of MBA students who took the MAM module at two different times, the first in June 2006 and the second in February 2007. The first group consisted of 52 participants, this group used the DES model. The second group was made up of 38 participants and evaluated the SD model.

9.5 Pilot study

Pilot testing is an important element of a research project in order to reveal potential problems and to assess the suitability of the survey. It thus enables the refinement of specific issues or problems related to the instruments chosen to be used.

Pilot testing involved the assessment of all the materials created for the sessions of the model use study, including the simulation models, the questionnaire survey and the paper-based materials (description of the case study, the model description and a user guide explaining how to use the model). Five pilot sessions with 5 PhD students from Warwick Business School were implemented to check the clarity of the questions and the layout of the survey. The subjects involved in these sessions were volunteers who offered to help at this stage. They were mainly PhD students at Warwick Business School, studying various subjects, such as Industrial Relations and Organisational Behaviour, Marketing and Strategic Management and also some PhD students from the Operational Research and Management Science group, who had experience of simulation.

During the pilot tests, mini sessions were run with each volunteer separately, imitating as closely as possible the actual session. The subjects were first introduced to the case study and model description and then were asked to run the computer model according to the task given. Next, discussions took place between the volunteer and the researcher, where the former was intentionally left to take a primary role in using the models. These discussions were run in a similar format to the group discussions intended to run as part of the real model use sessions (section 9.6). After that the subjects were asked to complete the questionnaire. In the meantime, comments about problematic areas related to any of the materials used during the sessions were made. Due to the discussions involved, the pilot sessions took approximately 2 hours, significantly longer than the real sessions.

The main aims of the pilot study were:

- To check that the simulation models were comprehensible and clear to the subjects and to identify any problems.
- To test the accompanying paper-based material
- To test the questionnaire

The main aspects raised and the changes made are now discussed regarding the simulation models, the paper-based material and the questionnaire.

1. *The simulation models*

Pilot testing helped in calibrating the models and gave a sense of reassurance that they were working as intended. During pilot testing, the models were generally found correct and some comments about the visual appearance were made. These ranged from modelling errors to visual appearance problems. The comments and the changes made follow:

- The graphical representation of simulation results (improvement of the appearance of tables, addition of the total number of people in prison, etc.)
- The addition of more possible values/options available for the user to choose from when entering input variables at the beginning of the simulation. For this reason more in-between values were added for ‘Petty Offenders’ and ‘Serious Offenders’ admission rates without changing the upper and lower limits which were accordingly 2500 up to 3500 and 500 to 900 prisoners admitted per year. The same applies for the input variables length of sentence for petty and serious offenders. These comments were also taken into account for the SD model.
- Two subjects were concerned with the fluctuations in the number of rehabilitated prisoners in the DES model. This was in fact evidence that randomness is obvious in the system, even though the reasons why this is happening were not understood.
- One of the volunteers suggested changing the percentage of the re-offending rate for people who have already been in prison once. Even though this was

considered a useful comment, this was not relevant to the assumptions of the model and so no action was taken.

2. *The paper-based material*

Some of the comments made about the models were utilised in order to clarify the description provided in the paper-based material. Some subjects suggested that a growth factor should be embedded in the number of annual admissions in prison. This was perceived as a misconception of the problem because the increase in the number of people in prison does not necessarily mean that there are more first-time offenders entering jail. This resulted in changes to the case study description where:

- The nature of a steady prison population was reinstated.
- A graph with the annual prison population totals and forecasted numbers was also added to visualise the stabilising effect.
- The effect of recidivism, as an important factor that affects the totals of prison population was emphasised.
- Simplifying assumptions were added.

Further comments regarding the paper based material dealt with:

- The clarification of the task objectives and the policies in the case study description.
- The necessity of introducing simplifying assumptions for the case study, which would line it up with the models created. The assumptions made were

added in the prison model description (*Critical model assumptions* in 'A model of the prison population' - Appendix A.4).

3. The questionnaire survey

During pilot testing the questionnaire survey was also revised. Subjects' comments included issues like: vague questions, problematic wording and incomprehensible questions.

- The demographic questions were further refined and additional options included.
- A few questions were found not to be serving their purpose by not receiving the intended answers based on the underlying hypothesis. These comments were taken into consideration and the questions were partly or fully amended. For example, question Q. a.5 (table 10-3) about the existence of feedback effects in the models, was converted into an open-ended format to avoid leading the respondents' answers (Bynner and Stribley, 1979, p.144-145). A closed question was considered to be leading the subjects to think about the feedback effects, which they might have not initially thought about.

The pilot tests assisted immensely in making sure that all the material was appropriately prepared and ready to be used in the real model use session.

9.6 The survey sessions

The survey was administered in two stages, using separately the DES and SD models. The survey on the DES model took place on June 26th 2006, while the SD one on February 5th 2007, involving two different groups of executive MBA students from Warwick Business School.

The sessions were arranged as part of the Modelling & Analysis for Management (MAM) core module on the Monday evening lecture, the first day of the modular week[18]. This particular time slot was purposely chosen in order to make sure that the subjects were not biased towards any of the two simulation approaches. Only some introductory lectures on modelling had been delivered earlier on that day and simulation is not taught until later in the week. The exercise was presented to the subjects as an example of a simple simulation model, which they would use themselves as a decision making tool.

Before the sessions, the subjects were given the case study description to read in advance. The sessions started with a brief presentation introducing the case study, the basics of the simulation models and how they work. Two further sets of hand-outs were given. The one consisted of the description of the model, including the statistics and assumptions made as well as an explanation of the task (Appendix A.4). The other handout consisted of guidance as to how to use each model (one for

[18] Modules for the part-time MBA course are delivered in one week blocks, from Monday to Friday. Information available on the Warwick Business School website, http://www.wbs.ac.uk/students/mba/learn/index.cfm (accessed 2nd September 2008)

the DES model and one for the SD model – see Appendices A.5 and A.6). The participants were then divided into syndicate groups and were asked to work on the task for 30-40 minutes. During this time they were asked to take the role of a government consulting service and to identify solutions to the problem. The groups consisted of 4-6 participants. All group members were involved in group discussions for half an hour. During the group sessions, three facilitators (the researcher, a colleague PhD student and the supervisor who teaches the course) were roaming from group to group providing support for technical problems and answering the questions raised. Each group worked separately in their assigned syndicate room where a computer with a running model was available. The models had been previously set up on each individual computer. While working in their syndicate groups, all members were actively involved in the task. One of the members was making notes on the board, one was mainly working with the computer model, while the rest of the group was making suggestions and taking part in group discussions. As a general observation, no major problems or difficulties with running the models were encountered. The MBA students were asked to prepare a presentation on their findings.

A feedback session followed, where two random syndicate groups for each session presented their findings and further discussions and comments were made by all participants. The participants raised issues about the simplicity of the models. This was dealt with by the professor teaching the course who highlighted the scope of the model and spoke about the advantages of simple models. At the end of the session, questionnaire surveys were handed out, which the participants were asked to complete, after having briefly explained that some research is being carried out at Warwick Business School on the use of models.

For the DES model survey, 34 out of 52 participants present on the day completed and returned questionnaires, resulting in a 65% response rate. For the SD model survey, 30 usable questionnaires from a total of 38 participants present in the February session completed and returned the questionnaires, resulting in an approximate 79% response rate. The 65% and 79% response rates are considered satisfactory.

9.7 Summary of the model use study design

This chapter describes the design of the model use study. The main activities followed are reported, including the development of the DES and SD simulation models, the construction of the survey questionnaire, the choice of the participant sample, the pilot study. It also describes the sessions implemented. The data collected are analysed in the next chapter.

Chapter 10: Results of the model use study[19]

10.1 Introduction

The current chapter presents the results of the statistical analysis on the data obtained from the study of model use, which was described in chapter 9. Non-parametric statistical tests are carried out to compare respondents' opinions. Conclusions are drawn based on the differences and similarities of DES and SD from the users' point of view. Overall, the empirical work does not identify significant differences for most of the comparison criteria for DES and SD model use.

The chapter is outlined as follows. First the respondents' profile is described. Then the statistical analysis performed is presented by survey question and the related research hypothesis (Table 9-3). At the end, the findings are discussed.

10.2 Respondents' profile

From the questionnaire survey with 2 different groups of executive MBA students, 34 usable questionnaires were derived from the DES group (implemented in June

[19] This chapter is based on Tako A.A. & Robinson S. (Forthcoming) "Comparing discrete-event simulation and system dynamics: Users' perceptions" Journal of Operational Research Society, 17pp, doi:10.1057/palgrave.jors.2602566. Accepted.

2006) and 30 from the SD group (implemented in February 2007). This gave response rates of 65% and of 79% respectively. Two mixed groups of executive MBA students in terms of background and management level participated in this survey.

In relation to the industry sector participation in the survey sample (Table 10-1), the majority of the DES group came from the public services sector (32% - 11 respondents) and from manufacturing (21% - 7 respondents), whereas the SD group, had no representation from the public services. It can be argued that participants from the public services sector are more familiar with problems in the prison population case study and so the DES group could be considered more predisposed to the exercise and the models. The majority of the respondents in the SD group came from the manufacturing sector (40% - 12 respondents). There was a smaller representation of the other sectors in both groups.

Table 10-1: Sample representation by industry sector

Industry	*DES group*	*SD group*
Public Services	32%	-
Manufacturing	21%	40%
Business Services	18%	13%
Financial Services	9%	3%
Transport & Communic.	9%	13%
Energy & Mining	6%	13%
Trade	3%	3%
Construction	3%	3%
Other	-	10%

The respondents were also asked to indicate their functional areas (Table 10-2) and their position in the management hierarchy. Participants in the DES group consisted of 34% working in the production/operations area, 20% in sales & marketing and 9% in computing/IT services, with a lower representation from finance, procurement, R&D and customer services. A somewhat similar picture was observed in the SD group, with 23% of respondents working in the production/operations area, 27% in sales & marketing, and 13% in computing/IT services and a lower representation of the other areas.

Table 10-2: Sample representation by functional area

Functional area	*DES group*	*SD group*
Production / Operations	34%	23%
Sales & Marketing	20%	27%
Other, i.e. Corporate Affairs, Analysis, etc	17%	13%
Computing/IT	9%	13%
Finance / Accounts	9%	3%
Customer Services	6%	7%
R & D	3%	10%
Procurement	3%	-
Human Resources	-	3%

Regarding the participants managerial level (Table 10-3), the majority of the DES group (61%) came from the lower (line) manager level with higher and middle management having a lower representation. Meanwhile, the SD group had a somewhat different representation, with the proportions being 40% and 47% for middle and lower level management respectively, while higher management had a lower representation. This suggests that both groups had a somewhat different mix regarding managerial level, which might affect the answers and thus the results. However, middle and line managers counted together represented 88% of the DES group and 87% of the SD group. Having a high representation of line management positions in both samples is considered to be beneficial for the survey. It is believed that managers of a lower level tend to use simulation to a greater extent as a problem solving tool. In fact, considering both groups together (Table 10-4), line managers made up the majority of respondents with prior experience 12 (out of 18) and only 4 (out of 18) middle managers had prior experience. There was only one instance of a higher level manager with prior experience of simulation.

Table 10-3: Managerial level for each group

Management level	*DES group*	*SD group*
Executive	12%	10%
Middle management	27%	40%
Manager	61%	47%
Other	-	3%

Table 10-4: Prior experience by management level (includes both DES and SD samples)

Management level by prior experience	**Count**	
	No	**Yes**
Executive	5	1
Middle management	17	4
Manager	21	12
Other	2	1
Total	***45***	***18***

10.3 The statistical analysis implemented

In order to test for differences in users' opinions, non-parametric statistical tests are carried out due to the nature of the data obtained from the questionnaire (ordinal and nominal data). According to Siegel (1957) meaningful statistics for nominal data are frequency counts and the mode, and for ordinal data, the median. Diagnostic P-P (probability-probability) plots are used to graphically explore differences in the distributions of ordinal data comparing answers received from the two groups of users. Fisher (1983) and Law (2007) suggest the use of P-P plots in order to compare two distributions. When the plot is linear or close to linear, the two distributions of answers fit one another, meaning that the variables have identical distributions (Wilk and Gnanadesikan, 1968; Fisher, 1983; Law, 2007). The chi-square test for the nominal data and the Mann-Whitney-Wilcoxon test (Fisher, 1983; Siegel and Castellan, 1988) for the ordinal data are used to check that the differences are statistically significant.

In the following sections, the statistical analysis for each research hypothesis on the comparison of DES and SD model use is provided.

10.4 Comparison of DES and SD model understanding

In line with hypothesis 3.1, users' opinions are measured on their understanding of the simulation and parts of it and on the factors that help model understanding. Participants' answers from both groups (DES and SD) are compared.

10.4.1 Level of understanding using the DES & SD models

Question a.3, in the main part of the survey questionnaire (Appendix A.7), included a series of statements regarding the respondents' understanding of the models when using either the DES or the SD model. Understanding deals with: overall model understanding, understanding of the relationship between variables, understanding of the model structure, understanding of how to use the model and understanding of the model outputs. The level of understanding for each of these items is measured on a scale of 1 to 5, where 1 means 'understand very little' and 5 means 'understand very well'. The aim is to measure the users' opinions regarding their understanding of the simulation model and parts of it and subsequently to compare the answers from both groups.

The P-P plots reveal differences in DES and SD model users' opinions only for the variables: understanding of the relationship between variables and understanding how to use the model (figures 10-1 and 10-2). The P-P plots consist of 5 data-points, where each dot represents the cumulative probability for each Likert scale measure (1 = understand very little, up to level 5 = understand very well), with 1 on the left and 5 on the right. The DES probabilities are plotted on the x-axis and for SD on the y-axis.

Looking more closely at both graphs, it can be observed that the lines are skewed towards the DES model. This means that the DES model users gave a higher proportion of responses in the mid-range (understand little, moderate and understand well, levels 2, 3 and 4), while the SD model was mostly rated at the higher levels of the scale (levels 3, 4 and 5). This implies that SD model users perceived that they had a better level of understanding regarding the relationship between variables and how to use the model. A two-sided Mann-Whitney-Wilcoxon test, however, does not identify these differences as significant.

In the P-P plots for the other items on understanding (overall understanding of the model, understanding its structure and understanding of the model outputs) there is little difference between the two groups. This is confirmed by a lack of statistical significance in the differences as well.

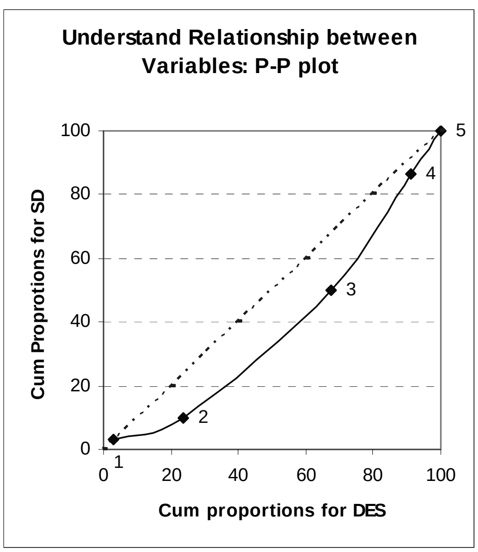


Figure 10-1: P-P plot on understanding of the relationship between variables, SD vs. DES answers, where 1 means understand very little and 5 understand very well.

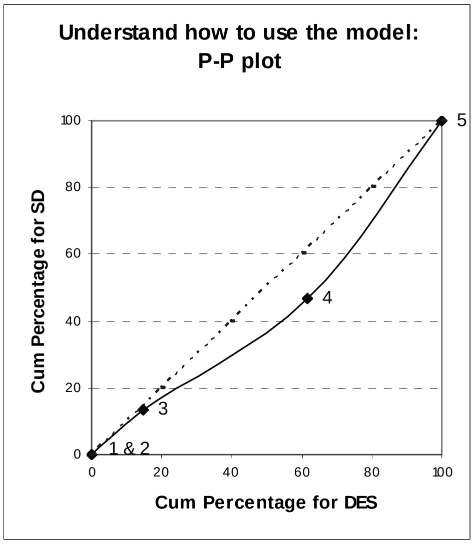


Figure 10-2: P-P plot on understanding of how to use the model, SD vs. DES answers, where 1 means understand very little and 5 understand very well. Points 1 & 2 coincide with the origin of the coordinates (0, 0) because none of the respondents answered with: understand very little, and little, for either model.

10.4.2 Factors that help in model understanding

The question regarding the factors that help model understanding asked the user to rank in order of importance the factors: paper-based material, visual display of the model and animation as the model runs. Looking at the answers received for each factor in table 10-5, there is a clear difference in the rankings of the DES group and the SD group for the factors paper-based description and animation as the model runs. This shows that the DES group identified animation as the most important factor that aided model understanding (58.8%), followed by the paper-based description as very important (55.9%). Meanwhile, the SD group identified the paper-based material as the most important factor (62.1%). However, there is no clear difference in the two groups' rankings regarding the visual display of the model. DES and SD users equally rated it as the least important factor.

Table 10-5: Ranking of factors that helped user understanding of the models (DES & SD)

Factor by model type		Important (%)	Very important (%)	Most important (%)
Paper-based material	DES	17.6	55.9	26.5
	SD	10.3	27.6	62.1
Visual display	DES	73.5	20.6	5.9
	SD	75.9	17.2	6.9
Animation	DES	11.8	29.4	58.8
	SD	24.1	48.3	27.6

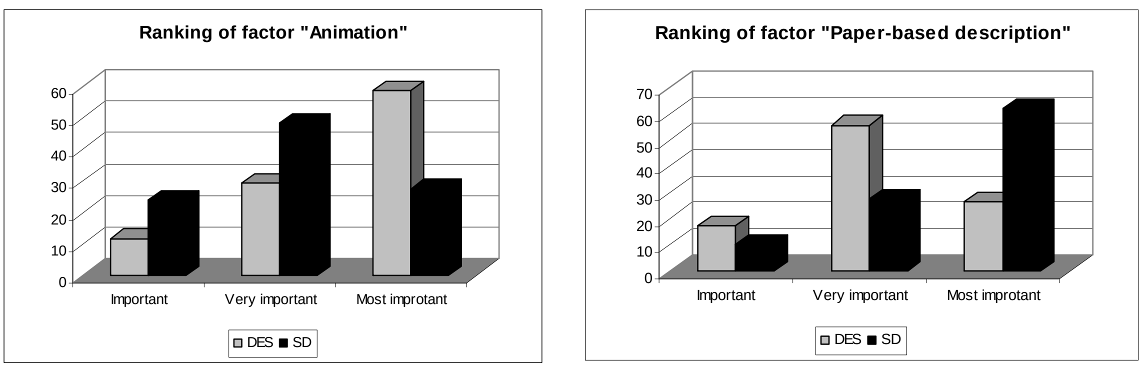


Figure 10-3: Frequency diagram showing importance of animation and paper-based description as factors that helped user understanding of the model (DES & SD)

The differences in the ranking of the factors animation as the model runs and paper based material between the two groups are graphically presented (Figure 10-3). The Mann-Whitney-Wilcoxon test shows that there is indeed a significant difference in users' opinions regarding these factors at a 1.4% and 2.9% significance respectively. This suggests that for the DES model animation has the greatest impact in helping model understanding, while for the SD model the paper-based material has most effect.

10.5 Comparison of perceived model complexity

The questions derived from hypothesis 3.2, deal with the perceived level of detail and the identification of feedback present in the two models. The relevant analysis follows.

Concerning the level of detail, a Likert-type question asked the user to rate the simulation models, where 1 represents very detailed and 5 a very high level perspective. The P-P plot (Figure 10-4) reveals a skew towards the SD model, with the SD model having a higher proportion of answers at the lower level of the scale, corresponding to a greater level of model detail. This is an unexpected finding because based on the relevant comparison literature, it is generally thought that DES models are more detailed.

The users could have perceived the SD model as more detailed due to the fact that all the components of the SD model are explicitly presented on screen (Figure 9-6), whereas for the DES model the structure may not be so explicit (Figure 9-3). The actual relationships between variables in DES models are not so apparent to the users when compared to SD models where the stocks, flows and auxiliary variables are displayed on the screen. Despite some skew towards the SD model results in the

P-P plot (Figure 10-4), the chi-square and Mann-Whitney-Wilcoxon tests do not reveal any significant differences between the two samples.

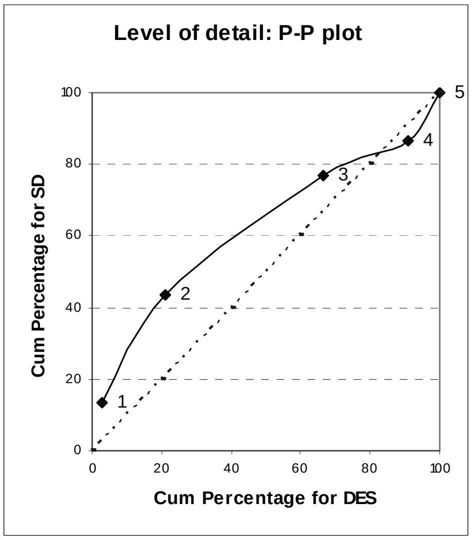


Figure 10-4: P-P plot on level of detail of the model, SD vs. DES answers, where 1 means very detailed and 5 meant very high level.

Next, the questionnaire consisted of an open-ended question asking the users to identify the sources of complexity in the model. The aim was to find out how obvious the feedback in each model (DES and SD) is to both groups of survey participants, without specifically mentioning ‘feedback’ in the question. It should be noted that the students had received no instruction on feedback as part of the MBA module. It was expected that users would identify the feedback in the model by considering the complexity that arises due to prisoners re-entering prison. Only 20% of the DES group and only 3% of the SD group answers were found as correct. Correct answers are considered as those that refer to the relationship and the interdependency between variables or to variables such as release of prisoners and re-offending. A chi-square test reveals a significant difference in the distribution of answers between the two groups, with a chi-square value χ =4.33, significant at 3.7% level. Contrary to expectations based on hypothesis 3.2, this suggests that the feedback effects were more explicit to the DES model users as opposed to the SD model users.

One possible reason for this counterintuitive result might be that the users did not actually explore the models enough in order to pick up on their underlying features. In the case of the SD model, the users would not be able to pick up the feedback effects between variables unless they navigated away from the model interface (Figure 9-7) to the model representation page (Figure 9-6). On the other hand, the closed path of the stock and flow network of offenders and recidivist is not a conventional reinforcing feedback loop in SD. Hence, it can be argued that the feedback effects were not obvious to the users. A low response rate was received for this question (the response rate was 35.3% for the DES group and 13.3% for the SD group), so this finding should be considered with caution.

10.6 Comparison of model validity

The questions in section b of the questionnaire deal with model validity, based on hypothesis 3.3 (section 3.4.3). The users were asked to provide their opinions to the extent they find the models representative of the case study situation and the outputs realistic. They were also asked to rate their level of confidence in the models.

The P-P plots do not show a difference between the two groups, apart from the plot on model representativeness. Observing the P-P plot (Figure 10-5), the line is skewed towards the DES group, revealing that the users of the DES model rated the model as being less representative, mostly levels 2 and 3 (little and moderate respectively), while SD model users rated it higher, mostly levels 3 and 4 (moderate and much). This implies that the SD model was perceived to be more representative of the case study compared to the DES model. Furthermore, a Mann-Whitney-Wilcoxon test identifies a somewhat significant difference at a 6.5% level.

When performing a Mann-Whitney-Wilcoxon test on the answers of users with no prior experience of simulation models, there is a more significant difference between the DES and SD groups; significant at a 1.7% level. This finding suggests that, for users with no prior simulation experience, the SD model was perceived to be more representative of the case study as opposed to the DES model. An obvious explanation for this result could be that, as discussed in section 10.5, the SD model structure is more explicit than the DES model structure. One DES model user commented that they would be interested to see the underlying mathematics.

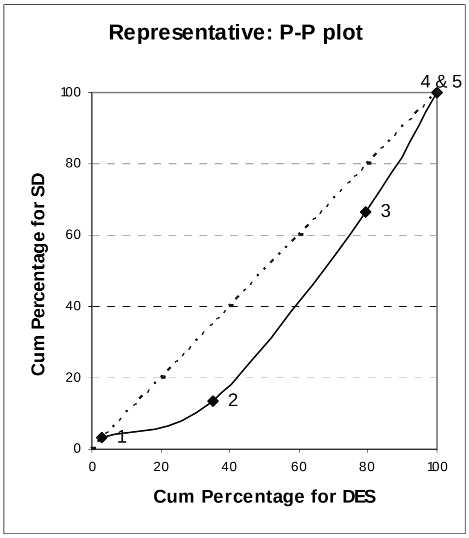


Figure 10-5: P-P plot on model representativeness, SD vs. DES answers, where 1 means very little and 5 very much. Points 4 and 5 coincide because none of the respondents considered the models representative at level 5.

10.7 Comparison of perceived model usefulness

Section c of the questionnaire dealt with users' opinion regarding the usefulness of the two simulation models, in connection with the research hypotheses 3.4 and 3.5 (section 3.4.3). 3 Likert-type questions and 1 open-ended question were used.

The three Likert-type questions asked users to express their opinions as to whether the use of the models enhanced their learning, it helped them think strategically about the problem and it facilitated the communication of ideas. The P-P plots do not identify any differences in the responses to the Likert-type questions apart from the question as to whether the use of the models facilitates the communication of ideas.

In the P-P plot (Figure 10-6) the line is skewed towards the SD axis in the lower end of the scale. This suggests that the users rated mostly high and very high the DES model in facilitating the communication of ideas. The Mann-Whitney-Wilcoxon test, however, does not identify any significant differences for any of the 3 Likert-type questions.

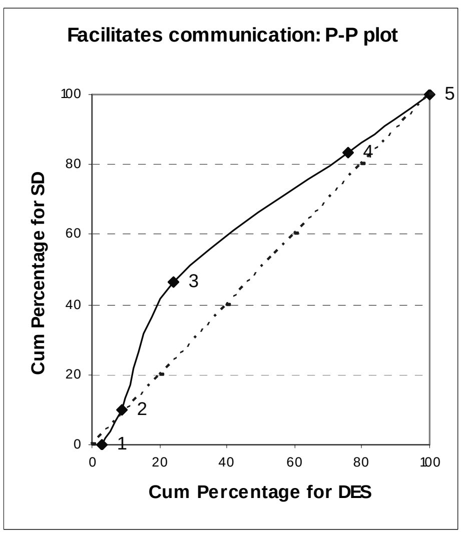


Figure 10-6: P-P plot on the capacity of the model to facilitate the communication of ideas, SD vs. DES answers, where 1 means very little and 5 very well.

To support the answers to the Likert-type question on the models' capacity to enhance learning, the open-ended question asked the participants to identify systems that are similar to the context of the prison population model. This question aimed to identify whether after using the prison population model the participants could transfer the knowledge gained to other similar systems. ‘Knowledge transfer’ can be used as an indicator of the learning achieved (Morecroft and Sterman, 1994).

Considering the answers to this question, only 23% of responses from each group are deemed as appropriate. Examples of correct answers are hospital/bed occupancy and job seeking/unemployment services. This indicates that the same level of learning was achieved by both groups. However, these findings are considered with caution because there was a high level of no-response to this question (the response rate was 44% and 36.6% for the DES and SD group respectively). On the other hand, it is not clear why some participants did not answer it. It might be that no-answer reflected a lack of learning and so a lack of ability to transfer the knowledge gained.

10.8 Comparison of model results

The questions in section d of the survey dealt with users’ opinions about model results and their interpretation. This part is based on research hypothesis 3.6 (section 3.4.3). The three aspects involved in the interpretation of model results in hypothesis 3.6 are: type of learning achieved, level of difficulty and considerations made about the behaviour of model results. These are separately considered in the analysis that follows.

In terms of type of learning, an issue of importance is the type of results users look at when running a simulation model. Based on hypothesis 3.6, DES model users were expected to focus on 'instrumental learning' and so were expected to look more at numerical data. Meanwhile, SD model users were expected to use graphs to a higher extent with more of an interest in 'conceptual learning'.

The questionnaire results show that, almost the same proportion of participants from both groups used the numerical results (instrumental learning). On the other hand, a higher proportion of respondents in the SD group claimed to have used the graphs (conceptual learning) as compared to the DES group. The bar chart (Figure 10-7) reveals the differences in the level of use of graphs between the two groups. Indeed, a relaxed chi-square test reveals a significant difference between the two groups at a 9.2% level of significance (Table 10-6).

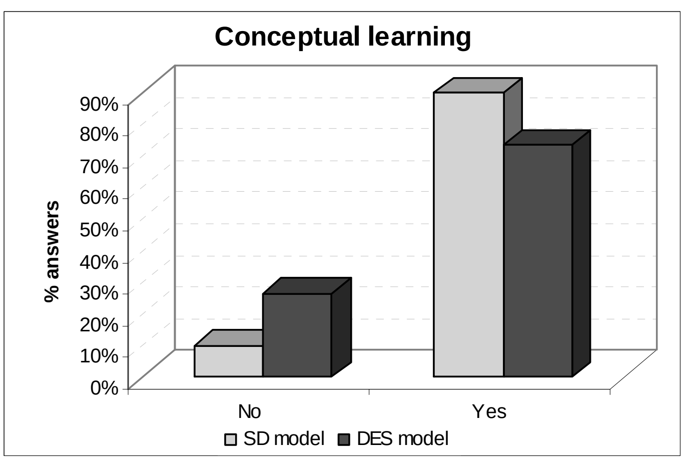


Figure 10-7: Bar chart with frequencies of DES & SD model users who used graphical outputs (conceptual learning) – a higher proportion of SD model users.

Table 10-6: Results from answers on type of outputs DES & SD users refer to when running the model

	DES model			SD model		
		Frequency	**Percent**		**Frequency**	**Percent**
Instrumental learning	No	12	35,3	No	12	40,0
	Yes	22	64,7	Yes	18	60,0
	Total	34	100,0	Total	30	100,0
	Chi-square: χ^2=0.15, 0.69 – Difference is not significant					
Conceptual learning	No	9	26,5	No	3	10,0
	Yes	25	73,5	Yes	27	90,0
	Total	34	100,0	Total	30	100,0
	Chi-square: χ^2=2.8, 0.092 – Significant difference between groups					

A Likert-type question dealt with users' perceived difficulty in the interpretation of results. The data reveal a difference in the two groups' opinions. The P-P plot (Figure 10-8) is significantly skewed towards the SD model, meaning that the SD model users found the results interpretation less difficult. Also a Mann-Whitney-Wilcoxon test reveals a significant difference at a 3.6% level.

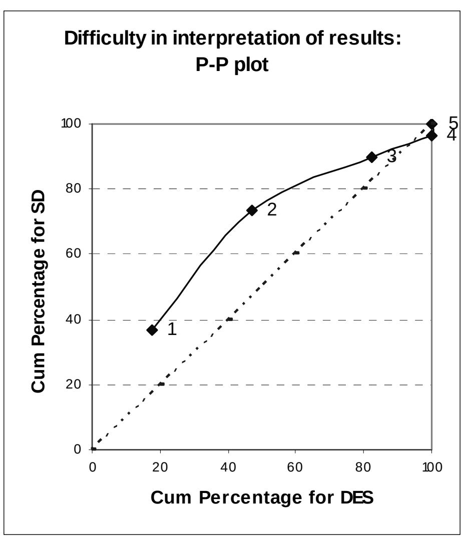


Figure 10-8: P-P plot on perceived difficulty in the interpretation of model results, SD vs. DES answers, where 1 means very straightforward and 5 very difficult.

Regarding the users' attitude when interpreting the model results, two Likert-type questions and one open-ended question, were used. The Likert-type questions were aimed at users' opinions about the usefulness of graphs and the extent to which they examine the factors that cause differences in the results. The open-ended question asked the user to identify the main learning point from the graphs, with the aim to understand whether the users became aware of the randomness in the models (aimed mostly at DES model). These questions were intended at finding whether the two models trigger model users to employ different attitudes towards model results. It was expected that DES model users would take notice of the randomness present in the outputs, and therefore, in response to the question they would mention randomness as their main learning point from the graphs. In the case of the SD model, the users were expected to be looking for the endogenous factors that cause the changes in the variables' behaviour and so give high scores to the Likert-type question. A somewhat higher percentage of the SD group observed the trend in the graphs as compared to the DES users (Table 10-7). The chi-square test, however, does not identify any significant differences in the answers between the two groups of users. Furthermore, no significant differences were found in the two user groups responses to the 2 Likert-type questions regarding their approach to interpreting the results.

Table 10-7: Comparison of answers to open-ended question about the observation of trends in the graphs

	DES model			SD model		
		Frequency	**Percent**		**Frequency**	**Percent**
Observe trends	No	30	88,2	No	23	76.7
	Yes	4	11,8	Yes	7	23.3
	Total	34	100,0	Total	30	100,0
Chi-square: χ^2=1.5 (0.221) – not significant difference						

10.9 Discussion of the findings of the model use study

This section discusses the findings of the model use study, based on the relevant research hypotheses (Table 3-3). The topics considered are the main aspects considered in this study, such as: model understanding, complexity, model validity, model usefulness and model results.

10.9.1 Model understanding

Hypothesis 3.1: *Despite the use of animation, DES models are not transparent to the client, and so difficult to understand compared to SD models, which are transparent to the client.*

Understanding, defined as the users' opinion on the level of understanding gained from using the two simulation models, is not found to be significantly different for the two (DES and SD) groups (section 10.4). Some differences were observed regarding the factors that helped users understand the model and parts of it.

Animation is found as the factor that mostly aided model understanding for the DES group, while for the SD group it was the paper-based description of the model (section 10.4.2). This complies with the views of Sweetser (1999) and Morecroft and Robinson (2005) that animation and on-screen displays can help model understanding. These results suggest that the understanding gained from using a DES model (because of animation and on-screen displays) is not necessarily more than the understanding achieved when using an SD model. Even though, these findings suggest that the level of user understanding is the same, it can be argued that users gain different insights from the two models. However, observation of the DES and SD groups using the models suggests that this was not the case in the current study, because similar issues and policies were considered by both groups during their discussions. The case and accompanying materials were, of course, the same for both sets of users, and so this is not unexpected. It can be concluded that the findings do not support hypothesis 3.1 and that it is, therefore, suggested that there is no difference between understanding of the DES and SD models.

10.9.2 Complexity

Hypothesis 3.2: *Different levels of complexity are involved in DES and SD modelling, resulting from the different levels of detail and perceptions of feedback.*

In this study, complexity is defined in terms of the level of detail and the feedback effects perceived in the model. It is found that users from both groups rated the two simulation models as involving a similar level of detail (section 10.5). A counterintuitive finding of this study is that the feedback effects were found to be more explicit to the users of the DES model. Contrary to hypothesis 3.2 and the general belief that SD is more appropriate in representing feedback structures (Coyle, 1985; Sweetser, 1999; Morecroft and Robinson, 2005), DES can represent feedback effects, which in this case appear to be more explicit to the user. This could suggest that the same level of complexity and feedback can be represented using DES and SD. Therefore, hypothesis 3.2 cannot be accepted and it is concluded that the same level of complexity can be involved in two equivalent DES and SD models.

On the other hand, the model building study revealed that DES modellers tended to create models involving more detailed complexity, as opposed to SD modellers (section 8.3.5). Therefore, the subjectivity in the choice of the two model representations should be taken into consideration. The DES and SD models could have been represented in many different ways. Obviously, for the purpose of this study, only one mode of display could have been chosen for each model. Caution needs to be raised about the finding on explicitness of feedback because of the small number of answers received to the open-ended question on model complexity. Furthermore, it should be considered whether the feedback effects involved in the prison population case study are representative of feedback effects modelled in SD. The use of different case studies is therefore, suggested to verify these findings.

10.9.3 Model validity

Hypothesis 3.3: *Users find DES and SD models equally credible for giving answers to a problem situation.*

Regarding model validity, this study suggests that the extent to which the users perceive the models to be representative of the case study is different between the two groups. From the analysis performed, the SD model was found to be slightly more representative (section 10.6). For both models the outputs were perceived to be equally realistic and both groups of users had the same level of confidence in them. The higher level of perceived representativeness related to the SD model can probably be attributed to the overall picture of the system provided within the SD model representation. On the other hand, the finding that model outputs and the confidence in the model were equally rated by both groups implies that overall the level of users' acceptance of both models was not different. Therefore, hypothesis 3.3, suggesting that DES and SD models are equally credible cannot be rejected. These findings provide useful evidence about the validity of models as perceived by model users, however, in this case the model users are not the real life clients and more to the point are rather detached from the problem modelled.

10.9.4 Model usefulness

Hypothesis 3.4: *DES and SD models are equally helpful as learning and communication tools.*

Hypothesis 3.5: *SD models can aid strategic thinking to a higher extent.*

Model usefulness, defined as a measure of a simulation model's capacity to enhance learning, communication of ideas and strategic thinking, was not identified as different between the two models (section 10.7). Against some opinions expressed in the literature (Sweetser, 1999), the findings suggest that both simulation approaches can be used as learning tools and can both trigger the communication of ideas. Even though in the SD literature a range of examples exists that illustrate the use of models for learning and for the communication of ideas (Vennix, 1996; Sterman, 2000), there are also cases where DES models have been used in facilitating group discussions and problem understanding (Robinson, 2001; Robinson, 2002). It is therefore, suggested that hypothesis 3.4 cannot be rejected. On the other hand, it is generally believed that SD modelling supports strategic thinking (Taylor and Lane, 1998; Sweetser, 1999; Lane, 2000). However, in this study both DES and SD models were found to assist strategic thinking to the same extent. Therefore, hypothesis 3.5 cannot be accepted.

10.9.5 Model results

Hypothesis 3.6: *The distinct nature of results derived from DES and SD models result in differences in learning & interpretation of results.*

Hypothesis 3.6 is tested against the findings of the analysis implemented in section 10.8. For model results, the findings indicate that the users of both, the DES and SD models, used the numbers (numerical displays) to the same extent. Meanwhile, the SD users focused on graphical displays more than the DES users, suggesting that SD models can aid conceptual learning and thus help users look at the bigger picture. Regarding the level of difficulty in the interpretation of results, the findings support the literature (Brailsford and Hilton, 2001) that the DES model results are more difficult to interpret, even though this specific model and the results were fairly simple. No differences are identified in the users' opinions on the use of graphs and in the attitude employed by the users when interpreting the model results. Therefore, it is suggested that hypothesis 3.6 cannot be rejected.

However, a difference in attitudes was observed by the researchers during the group discussions. The SD model users tended to take a 'goal seek' approach, where they blindly changed the inputs in order to get the right output, and then reflected on what policies might be employed to achieve these inputs. The DES group did not employ the same approach and focused on the effect a policy might have on the inputs to the model and then set the input values accordingly.

10.9.6 Summary of the findings of the model use study

The current study adds to the discussion on the comparison between DES and SD. This is the only empirical study that tests the differences in using DES and SD simulation models. It provides empirical evidence as to how users perceive the differences between DES and SD. The comparison criteria used in the survey are based on the generally accepted opinions/statements regarding the differences in using DES and SD found in the literature. Overall, it was not possible to identify many significant differences in the users' opinions regarding the specific DES and SD models used. This may imply that from the user's point of view the type of simulation approach used makes little difference if any. Akkermans (1995) reaches a similar conclusion, identifying that clients are usually indifferent to the simulation language being used. This may not be too surprising, as users are likely to be more interested in what they can learn from a model rather than how the model works; that is, as long as the modelling approach is able to address the problem situation. A summary of the main findings derived from the model use survey is provided (Table 10-8).

Table 10-8: Summary of the findings on the comparison of DES and SD model use, by hypothesis

Hypotheses on model use	Research findings
Hypothesis 3.1: Despite the use of animation, DES models are not transparent to the client, and so difficult to understand compared to SD models, which are transparent to the client.	Rejected
Hypothesis 3.2: Different levels of complexity are involved in DES and SD modelling, resulting from the different levels of detail and	Rejected

perceptions of feedback.	
Hypothesis 3.3: Users find DES and SD models equally credible for giving answers to a problem situation.	Not rejected
Hypothesis 3.4: DES and SD models are equally helpful as learning and communication tools.	Not rejected
Hypothesis 3.5: SD models can aid strategic thinking to a higher extent	Rejected
Hypothesis 3.6: The distinct nature of results derived from DES and SD models result in differences in learning & interpretation of results.	Not rejected

10.10 Summary of the model use study results

In this chapter the results of the model use study are discussed. First, the profile of the two groups of MBA students involved in the study is presented. Next, the results of the statistical analysis implemented on the responses collected in terms of model understanding, complexity, model validity, usefulness and model results, are reported. The findings obtained and the implications of the study are further discussed.

Chapter 11: Conclusions

11.1 Introduction

This final chapter brings to a close the work presented in the thesis. It outlines the objectives driving the research undertaken. Next, it presents the key findings of the study, followed by an assessment of the extent the research objectives are achieved, taking into consideration the limitations of the study. The chapter also identifies the main contribution of the study and areas of further research.

11.2 Summary of research objectives

The thesis sets out to empirically verify or refute the statements made in the literature, with an ultimate view to empirically understand the similarities and differences between the DES and SD modelling approach. The main research question that drives the empirical work is:

“What are the differences and similarities between the DES and SD modelling approaches and the use of DES and SD models?”

The thesis compares DES and SD modelling, focusing only on the model building process and model use. More specifically, the research objectives are:

1. To empirically determine how different the modelling process followed by DES and SD modellers is.

2. To establish the differences and similarities in the modelling approach taken by DES and SD modellers in each stage of simulation modelling.

3. To assess how different DES and SD models of an equivalent problem are from the users' point of view.

The choice of methods utilised in this thesis is made based on each specific research objective and the requirements involved. Two separate studies were implemented: a model building study based on the first and second research objectives, and a model use study dealing with the third research objective. In the former study, Verbal Protocol Analysis was used, where expert DES and SD modellers were asked to 'think aloud' while developing simulation models. Next, the model use study involved a questionnaire survey with Executive MBA students, who were asked to provide opinions about two equivalent DES and SD models.

11.3 Key research findings

In the two studies implemented in the thesis, the comparison of DES and SD modelling approaches focuses respectively on model building (research objectives 1 and 2) and model use (research objective 3). Overall, the model building study found a number of differences between the DES and SD modelling approaches (findings 1-9). Nevertheless, it was not possible to identify many significant differences in the users' opinions regarding the use of two equivalent DES and SD models (findings 10-14). The key findings of the study are provided under the respective research objectives:

Objective 1: To determine how different the modelling process followed by DES and SD modellers is.

1. DES modellers focus significantly more on model coding and verification and validation of the model, whereas SD modellers concentrate on conceptual modelling. (section 6.3)

2. DES modellers progress more linearly among modelling topics compared to SD modellers. (section 6.4)

3. DES and SD modellers follow an iterative modelling process, but their pattern of iteration differs. (section 6.4.1)

- DES and SD modellers switch their attention frequently between topics, and almost to the same extent (505 times for DES modellers and 507 times for SD modellers) during the model building exercise.

- The cyclicality of thinking during the modelling task is more distinctive for SD modellers compared to DES modellers. For example:
    - DES modellers tend to create all the parts of the model and then enter the data in the model, and after having validated the model they go back to change it on the computer. They less often go back to conceptual modelling.
    - SD modellers jump from conceptual modelling to model coding or data inputs and then go back to conceptual modelling to add further parts into the model.

Objective 2: To establish the differences and similarities in the modelling approach taken by DES and SD modellers in each stage of simulation modelling.

4. DES and SD modellers consider similar modelling objectives. The primary objective considered is to create a tool that projects the output of interest into the future. SD modellers show a tendency to consider broader aspects of the problem modelled. Furthermore, policy testing is an important aspect of SD modelling, which is not as prevalent in DES modelling. (section 7.2)

5. Standard conceptual modelling diagrams are not used by most modellers in the study. Causal relationships and feedback are an important aspect in SD modelling, whereas in DES modelling these are not prevalent. DES

modellers take an analytic perspective when modelling, whereas SD modellers think in a more holistic pattern. (section 7.3)

6. Detailed thinking characterises DES modellers' protocols. SD modellers refer to dynamic complexity and tend to use existing modelling structures. (section 7.4)

7. DES models are almost entirely built on quantitative data, whereas SD modellers use both quantitative and qualitative data. DES and SD modellers' attitudes towards missing data vary depending on individual modellers. Some required additional data, whereas others were willing to make assumptions. As expected, randomness is an important aspect in DES modelling, whereas no references are made to it by SD modellers. DES and SD modellers are familiar with modelling the linear relationships between variables in the model, whereas SD modellers pay limited attention to non-linear relationships too. (section 7.5)

8. Both DES and SD modellers are concerned with creating an accurate model. DES modellers pay more attention to verification and white box validation compared to SD modellers. SD modellers engage in 'black-box' validation more than DES modellers, checking the numerical outcomes, but also the patterns of the results. Contrary to expectations, SD modellers do not refer to model usefulness as a way of validating the model. (section 7.6)

9. DES modellers subconsciously think about creating a model in a steady state, while SD modellers think about creating a model in equilibrium. Both DES and SD modellers show an interest in the quantitative and qualitative aspects of the results of the models. DES modellers tend to consider more detailed results, whereas SD modellers consider most often the model outputs in a qualitative way. SD modellers are more keen on developing scenarios for experimentation with the model. (section 7.7)

Objective 3: To assess how different DES and SD models of an equivalent problem are from the users' point of view.

10. The level of understanding gained from using the two (DES and SD) simulation models is not significantly different. Some differences are observed regarding the factors that help users understand the model and parts of it. Animation aids understanding when using the DES model. Furthermore, the study suggests that users gain similar insights from the two models. (section 10.4)

11. A similar level of complexity (detail) can be delivered by two equivalent DES and SD models. Contrary to expectations, feedback effects can appear equally or even more explicitly in the DES model. (section 10.5)

12. Regarding model validity, the SD model is found to be slightly more representative of the problem situation. However, the overall acceptance of both (DES and SD) models (in terms of model outputs and confidence in using the models for decision making) is not different. (section 10.6)

13. Model usefulness is not identified as different between the two models. The findings suggest that both simulation approaches can be used as learning tools and can both trigger the communication of ideas. (section 10.7)

14. For model results, users achieve the same level of instrumental learning from using the DES and SD models. Meanwhile, SD models can aid more conceptual learning and thus help users look at the bigger picture. DES model results are more difficult to interpret, even though the model used in this study and its results were fairly simple. No differences are identified in the users' opinions about the use of graphs and in the attitude employed by the users when interpreting the model results. (section 10.8)

11.4 Achievement of objectives

After summarising the key findings of the study, this section considers the extent to which the analysis performed has met the research objectives stated in chapter 3.

11.4.1 Comparison of the DES & SD model building process (objective 1)

By undertaking a quantitative analysis of the verbal protocols it was possible to compare the modelling process followed by DES and SD modellers. The results supported the views expressed in the literature (section 2.4.1) and were on the whole as expected. As with generic OR modelling, both DES and SD involve iterative modelling processes. A new insight gained from this analysis was that DES modellers' thinking followed a more linear process, whereas for SD modellers it involved more cyclicality. As expected, differences were identified in the attention paid to different modelling stages. Clearly, the results are dependent to some extent on the case study used and the modellers selected. Therefore, considerations are made about the limitations of the study and the consequent validity of the findings (section 11.6).

11.4.2 Comparison of DES & SD modellers' approach to model building (objective 2)

With the qualitative text analysis of the modellers' verbalisations obtained from the model building sessions it was possible to compare the DES and SD modellers' approach taken during the stages of simulation modelling. Overall, the comparison of the model building approach taken by DES and SD modellers has identified some marked differences between the two groups of modellers. It is however, suggested that some of these aspects require further investigation.

Some similarities were found regarding the modelling objectives considered by DES and SD modellers (hypothesis 2.1), use of conceptual modelling diagrams (hypothesis 2.3) and their attitude towards missing data (part of hypothesis 2.11). Differences were found regarding modelling of causal relationships and feedback effects (hypothesis 2.4), the perspective taken towards systems' view (hypothesis 2.5), the complexity and the detail involved (hypothesis 2.7), use of prior modelling structures (hypothesis 2.8), representation of delays (hypothesis 2.10), randomness (hypothesis 2.12), focus of attention during verification and validation ('black-box' and 'white-box' validation) (hypothesis 2.15) and views on reaching a steady state or equilibrium condition (hypothesis 2.16).

The model building study was not able to investigate some of the statements found in the literature and the research hypotheses identified. Modelling philosophy does not naturally occur during the VPA sessions and therefore, it was not possible to investigate in this research. In addition, some research hypotheses were specified as non-testable or factual and the reasons for this have already been provided in section 5.5.3. Partial answers were given to some of the hypotheses regarding the use of diagramming methods (hypothesis 2.3), type of data inputs used (hypothesis 2.11), use of non-linear relationships (hypothesis 2.13), model usefulness as a measure of model validity used in SD (hypothesis 2.14) and optimisation of model results in DES (hypothesis 2.18) (Table 8-2). It is suggested that the findings obtained in the analysis regarding the aforementioned hypotheses are a result of the case study chosen and the sample of modellers selected. Hence, it is concluded that these aspects need to be further considered in future research by either utilising different case studies or larger samples.

11.4.3 The comparison of DES and SD models use (objective 3)

A questionnaire survey was successfully used to compare users' opinions regarding two equivalent DES and SD models. Overall, the results obtained were not expected. The empirical work did not identify any significant differences for most of the comparison criteria for DES and SD model use.

Contradictory views on DES and SD model understanding were verified in this study. The findings suggest that regardless of the nature of the models and their transparency, the users' understanding is not different. Contrary to expectations, the same level of complexity was perceived in the DES and SD models, whereas the feedback effects were more explicit to users of the DES models. As expected, both DES and SD models and their results were found equally credible. In terms of model usefulness, while SD models are established in the literature as tools that assist in learning, this study found the equivalent DES and SD models as equally useful in their capacity to enhance learning, in aiding communication of ideas and strategic thinking. As expected, model results were found more difficult to interpret in the case of DES models. However, this study was not able to explore in detail the learning achieved from the results. This could be further studied employing focus groups, which could reveal the learning achieved from the use of models.

While the comparison study undertaken does not reveal many significant differences from using two equivalent DES and SD representations, considerations should be made as to whether this is a result of the case study and models used, as well as the samples involved. Therefore, the results obtained should be considered with caution and generalisations are to be avoided. These findings need to be further verified with future research (sub-section 11.6.2).

11.5 Contribution of the thesis

The research carried out in the thesis provides an empirical comparison of the DES and SD model building process and model use. The work is the first of its kind and contributes towards the comparison of the two widely used simulation approaches in OR, DES and SD. Work on the comparison of DES and SD is limited, consisting mainly of some conference papers. The few comparisons that can be found in the literature are mostly based on the authors' personal opinions. To date, there has been no empirical study reported that provides an evidence base for the comparison of the two approaches. The key contribution of this study is that it provides empirical evidence on the differences and similarities between DES and SD from the model building and model use point of view.

The findings of this study could be useful to academics and practitioners alike. For the former, this study contributes to the theory of simulation (DES and SD). The findings of the study provide new insights into the comparison of the two simulation approaches. Beyond the actual findings on the comparison of DES and SD, this study contributes to the wider area of simulation. It opens new areas of exploration, sharing concepts between the two fields that could be beneficial for the advancement of each field. For example, this study introduces new aspects in DES modelling such as: the development of systems thinking, use of prior modelling structures, using the models for the communication of ideas, and learning from models, which are underdeveloped areas in this field.

It is however accepted that there are aspects (i.e. feedback effects in SD modelling, randomness in DES modelling) which each approach is better at dealing with. In addition, DES and SD modellers think differently during the different modelling stages, implying that indeed different aspects of a problem situation can be considered under different lenses if modelled with the DES or SD approach. Therefore, the complementary use of both approaches can provide useful insights into the problem studied. Teaching DES and SD modelling simultaneously to novice modellers can also prove beneficial for educational purposes. Learning to model, by studying both approaches can inherently instil different approaches to modelling and at the same time a different level of attention to various modelling stages. In addition, the findings provide academics from the two fields of modelling with a better understanding of the two fields, helping to leave behind generally accepted beliefs about the other field. It also promotes the use of two alternative simulation approaches in a complementary way.

Furthermore, this study introduces new methods that can be used in the field of simulation modelling. It introduces a new case study, that of the UK prison population, which has been used to study the two modelling approaches. Furthermore, the questionnaire survey developed contributes to the area of the evaluation of simulation models. New concepts have been created, which can be used in future studies for the evaluation of simulation models. These are probably geared towards the comparison of simulation models, however, with some adaptations these can be re-used for other purposes, e.g. assessing the learning achieved from the use of models, client satisfaction with simulation models, etc.

Considering the contribution of the work undertaken in this study from a practical point of view, this study can be beneficial for simulation practitioners in each field. First of all, it raises the awareness about the modelling steps followed as part of simulation modelling. It also helps in understanding, but also in learning from the two simulation approaches leading to an improved modelling practice. While this study suggests that the modelling process followed in DES and SD modelling is different, as well as the approach taken to various aspects of modelling, there is potential for the exchange of insights between the two fields. For example, it highlights the importance of conceptual modelling as a means of understanding the structure of the model, before creating one on the computer. On the other hand, it is suggested that the complementary use of both approaches can provide different insights to modellers and subsequently to model users. Therefore, practitioners who use both simulation approaches would be better equipped to tackle various problem situations from different angles. Furthermore, this study raises awareness that in practice the type of simulation approach used makes little difference to the users of the model and its results, that is, as long as the modelling approach is able to address the problem situation, however, it is important to determine the modelling approach bearing the users in mind.

The findings of this study can ultimately help in the selection of the appropriate simulation approach when modelling a particular problem situation, albeit specific answers are not provided. The aspects studied in this research can be further used in setting criteria for the development of a framework that will help in the choice between DES and SD also including other simulation approaches (i.e. agent-based simulation).

11.6 Limitations of the study

After having considered the findings of the study and the way these relate to the achievement of the study objectives, the limitations of the two specific studies undertaken should be considered.

11.6.1 Limitations of the model building study

The model building study undertaken is the only empirical study that compares DES and SD model building based on data gathered from experimental exercises involving expert modellers themselves. Some distinct differences between DES and

SD modelling were found regarding model building. There are however, a few limitations to this study which should be considered.

Obviously it should be noted that the findings of this study are based on the researcher's interpretation of participants' verbalisations. Subjectivity is involved in the analysis of the protocols, as well as in the choice of the coding scheme. This cannot be avoided in qualitative research. A different researcher might have reached different conclusions (using a different coding scheme with different definitions). To deal with the subjectivity involved in the coding process, the transcripts of all protocols were coded twice, involving a gap of 3 months between codings (section 5.5.2). In addition, a third party blind checked the coded protocols, to ensure that a consensus in coding was reached. This ensured that some level of objectivity was maintained in the coding of the verbal protocols.

Additionally, the current findings are based on the verbalisations obtained from a specific sample of modellers who participated in the exercise. The selection of modellers to participate in the survey was made based on convenience sampling (section 5.2), therefore, consideration should be made as to whether this is a representative sample. An attempt was made to involve a representative sample of the DES and SD community. The SD group included experts with a greater experience than the DES group (section 6.2). It could therefore be suggested that the greater experience among SD modellers was the cause of the shorter protocols for the SD group. Two outliers were identified, but these are not considered to have affected the results. The statistical analysis performed on the total amount of verbalisations by the DES and SD modellers did not find any significant statistical differences (section 6.3). Furthermore, due to the limited number of potential participants available for this study, differences (i.e. background, education, etc.) among modellers were difficult to control. A bigger sample size could have also provided more representative results. Due to project timescales and the difficulties encountered in finding expert modellers from each field, using a random and a larger sample was not feasible.

Considering the data (verbal protocols) obtained from the modelling sessions implemented, these are derived from artificial laboratory settings, where the modellers at times felt the pressure of time or the pressure of being observed. In one case one modeller depicted an exam-type attitude where he was trying "*to tick as many boxes as possible*". In all other protocols, this type of attitude was not noticed. But this does not mean that it was not present. Unless observing a real modelling task, it is not possible to distinguish the effect of the artificial settings on the modellers' behaviour. However, the researcher attempted to reduce these effects to a minimum. At the start of the modelling sessions the researcher gave clear instructions about the exercise and the aim. Among others, it was mentioned that the models created as part of the exercise would not be scored and it was highlighted that the aim of the exercise was to understand the modellers' thinking process. It was also attempted to create a most relaxed atmosphere so that modellers felt at ease during the sessions. Verbalisation exercises were also run at the beginning of the sessions for the modellers to practice and aid a more natural verbalisation.

In addition, despite the researcher's attempts to run all modelling sessions under similar conditions, this was not always possible. It should be noted that the modelling sessions were arranged in or close to the modellers' working environment. The sessions had to be arranged taking into consideration the specific circumstances of each modeller, resulting in the modelling sessions being run under different conditions and time during the day. This could have affected the accuracy of the protocols. However, it is believed that the protocols derived are representative of each field. Hence, the protocols obtained were the best that could be obtained for the available research resources and conditions. It is, however, suggested that for future research, observation of real-life modelling projects can provide more representative accounts of DES and SD modellers' behaviour, albeit that they would have the limitation of comparing behaviours in different contexts.

It should be acknowledged that the case study chosen involved a simple and a quite structured task to ensure completion of the exercise for the limited time available with the modellers. These factors have to some extent affected the smaller amount of verbalisations for modelling topics such as: problem structuring and maybe conceptual modelling. Additionally, the task did not specifically require the modellers to consider scenarios and the results of the model and so limited verbalisations and partial findings were obtained on results & experimentation and implementation.

The prison population case study chosen was considered a suitable instrument that can potentially instigate the modelling of both DES and SD models. If a different case study was used, different protocols and thus different results would have been obtained. More than one case study could have been used, but this was not possible given the participants' limited time availability. Moreover, the participants were required to fully engage in thinking about the model, so it was considered that shorter modelling sessions (less than 1 hour) would not be suitable or effective for the purposes of the research. However, for future research, the use of different case studies to the UK prison population is suggested. This would mitigate the effect of the case study used on participant responses. Therefore, a random allocation of case studies to each modeller would be beneficial for the quality of the research. Using two or more different case studies can provide more representative results regarding the differences between the two modelling approaches.

11.6.2 Limitations of the model use study

The participant groups involved in the exercise were two mixed groups of executive-MBA students in terms of background and level of management and thus comparable to each-other. There was a high representation of first line managers, who tend to be more involved in simulation projects, as compared to higher level managers. It should be noted that from the sample used in this study, the proportion of participants with prior experience was lower than those without any. Both groups commented on the simplicity of the models, but at the same time they appreciated their usefulness for the purpose at hand. Participants tended to be looking for more sophisticated models, considering a wider range of factors such as: costs, deaths, and other types of sentences.

In the model use study, the best possible samples to which the researcher had access to were used. Of course, the study could be improved with larger samples. In the DES group, the highest proportion of participants had a background from the public services and manufacturing sectors, while in the SD group, it was the manufacturing sector (section 10.2). In the latter group, there was no representation from the public sector. As a general comment, the DES group expressed a greater interest in the exercise. A possible explanation is that for a considerable proportion (32%) of the group this was a problem related to their jobs. This in itself could have biased their answers. It would be considered a more fair experiment if the participants were randomly allocated into each group. However, random assignment of participants in the two groups was not possible because each MBA group took the same course at different times (May 2006 and February 2007). It was observed that because the users were exposed to only one of the two simulation models, they tended to take for granted the features of each simulation model, and did not pick up the specific features of each approach, which differ from one another. A solution to this would be to get the participants to work with both simulation models. This was not possible due to the limited amount of time available, but it is suggested for future research as an extension of the current study.

There is some level of subjectivity in the choice of the case study and the simulation models. The case study was chosen because it was amenable to both DES and SD modelling. Use of an alternative case study may have provided different findings in terms of the comparison of DES and SD. Meanwhile, specific DES and SD models were built of the prison population problem. These were only one representation in each approach out of many (if not an infinite number of) possible representations. Would different DES and SD models of the problem have led to different findings? To mitigate this effect, the DES and SD models were developed with the help of experts in their respective fields. It is believed that these models are typical DES and SD models, but it cannot be claimed that they are the only possible models. Future work could compare DES and SD using different case studies, and a range of different models and simulation packages could be investigated for each case study.

A natural extension of the model use study, presented in the thesis, would be to further explore the findings of the study in focus groups with managers. This would enable checking the findings of the research, but it would also provide an in-depth understanding of some counter-intuitive results of the study.

11.7 Further work

Finally, in this section further work is suggested, which takes forward the limitations of the model use and model building studies raised in section 11.6. This includes ideas for the potential advancement of the findings from the current study, aspects of the comparison of DES and SD modelling that have not been considered in the current study as well as new ideas that emerged from the current study.

Ideas for the improvement and advancement of the current findings are:

- Comparing DES and SD model building using different and more than one case study (varying aspects such as: hard and soft data inputs, linear and non-linear relationships, etc.). The use of larger samples can improve the validity of results.
- Comparing the use of simulation models using different case studies and subsequently different models. It is also useful to change the study to provide model users with both simulation models and subsequently their opinions on both models collected.
- The comparison of DES and SD model validation in real life projects.
- Comparison of diagramming methods used in DES and SD modelling through a survey of modellers.
- Study the model building with modellers who use both simulation approaches.

Studies that deal with aspects of the comparison of DES and SD modelling that have not been covered in the current thesis:

- Comparison of the DES and SD modelling philosophy involving focus groups with expert modellers.
- Comparison of the simulation literature regarding the use of DES and SD models for strategic and tactical aspects.

New areas of research that emerge from the current study are:

- In-depth understanding of the learning achieved from using simulation models.
- Study of the effect of the number of iterations in the modelling process on model quality.

The collection of the work suggested and the work undertaken in this thesis can eventually support the development of a framework with criteria for the selection between the two simulation approaches.

11.8 Summary and author's comments

In this chapter, an overall summary of the thesis is provided. The objectives of the study are presented, followed by a report of the key findings and the contribution of the study. The achievement of the objectives is discussed followed by suggestions for improvement and extension of the current study. It is believed that this comparison study will be useful to modellers (including novices in DES and SD field) and to practitioners in simulation and in OR in general.

The process of completing this study was fascinating, at times difficult and challenging, but overall rewarding. The author enjoyed dealing with a range of people, including modellers, administrative staff, students and colleagues. Furthermore, the author gained useful experience in designing a research study, in project managing, approaching participants and convincing others about the benefits of this research.

References

Akkermans, H. A. (1995). Modelling with Managers. Participative Business Modelling for Effective Strategic Decision-Making. Technology Management. Eindhoven, Technische Universiteit Eindhoven. PhD thesis.

Artamonov, A. (2002). Discrete-event Simulation vs. System Dynamics: Comparison of Modelling Methods. Warwick Business School. Coventry, University of Warwick. MSc Dissertation.

Baines, T. S. and D. K. Harrison (1999). "An opportunity for system dynamics in manufacturing system modelling." Production Planning & Control **10**(6): 542-552.

Baines, T. S., D. K. Harrison, J. M. Kay and D. J. Hamblin (1998). "A consideration of modelling techniques that can be used to evaluate manufacturing strategies " The International Journal of Advanced Manufacturing Technology **14**(5): 369-375.

Banks, J., J. S. CarsonII, B. L. Nelson and D. M. Micol (2001). Discrete-Event System Simulation, 3rd ed. in Industrial and Systems Engineering series. Upper Saddle River, New Jersey: Prentice Hall.

Bard, J. F. (1978). "The Use of Simulation in Criminal Justice Policy Evaluation." Journal of Criminal Justice **6**(2): 99-116.

Bellinger, G. (2004). "Systems." Retrieved 28 October 2008, from http://www.systems-thinking.org/systems/systems.htm.

Berg, B. L. (2004). Qualitative research methods, 5th ed. ed. Boston: Pearson Education.

Borshchev, A. and A. Filippov (2004). From System Dynamics and Discrete Event to Practical Agent Based Modelling: Reasons, Techniques, Tools. Proceedings of the 22nd International System Dynamics Conference, Oxford, England, UK.

Brailsford, S. and N. Hilton (2001). A Comparison of Discrete Event Simulation and System Dynamics for Modelling Healthcare Systems. Planning for the Future: Health Service Quality and Emergency Accessibility: proceedings from the 26th meeting of the ORAHS 2000. J. Riley. Glasgow Caledonian University**:** 18-39.

Brailsford, S. C., V. A. Lattimer, P. Tarnaras and J. C. Turnbull (2004). "Emergency and on-demand health care: modelling a large complex system." Journal Of The Operational Research Society **55**(1): 34-42.

Brennan, A., S. E. Chick and R. Davies (2006). "A taxonomy of model structures for economic evaluation of health technologies." Health economics **15**(12): 1295-1310.

Bryman, A. (2008). Social research methods, 3rd ed. ed. Oxford: Oxford University Press.

Buckingham, A. and P. Saunders (2004). The survey methods workbook : from design to analysis. Cambridge ; Malden, Mass.: Polity.

Bynner, J. and K. M. Stribley, Eds. (1979). Social Research: Principles and Procedures. London, Longman; Open University Press.

Cassel, R. A. and M. Pidd (2001). "Distributed discrete event simulation using the three-phase approach and Java." Simulation Practice and Theory **8**(6-7): 491-507.

Checkland, P. (1999). Systems thinking, systems practice Chichester John Wiley.

Cox, G. B., P. Harrison and e. al. (1978). "Computer Simulation of Adult Sentencing Proposals." Evaluation and Program Planning **1**: 297-308.

Coyle, R. G. (1985). "Representing Discrete Events in System Dynamics Models: A theoretical Application to Modelling Coal Production." The Journal of the Operational Research Society **36**(4): 307-318.

Coyle, R. G. (1996). System Dynamics Modelling: a practical approach, 1st. ed. London: Chapman & Hall.

Curram, S., J. Coyle and D. Exelby (2000). Planes, Trains And Automobiles... A System Dynamics Model With Discrete Elements. In the Proceedings of the 18th International Conference of the System Dynamics Society, 6-10 August 2000, Bergen, Norway.

Daniel, W. W. and J. C. Terrell (1995). Business Statistics for Management and Economics, 7th ed. ed: Houghton Mifflin Company.

Donzelli, P. and G. Iazeolla (2001). "A hybrid software process simulation model " Software Process: Improvement and Practice **6**(2): 97 - 109.

Ericsson, K. A. and H. A. Simon (1984). Protocol Analysis: Verbal Reports as Data, The MIT Press.

Fisher, N. I. (1983). "Graphical methods in nonparametric statistics: a review and annotated bibliography." International Statistical Review **51**: 25-58.

Forrester, J. W. (1958). "Industrial Dynamics: A Major Breakthrough for Decision Makers." Harvard Business Review **36**(4): 37-66.

Forrester, J. W. (1961). Industrial Dynamics. Cambridge: MIT Press.

Forrester, J. W. (2007). System dynamics - a personal view of the first fifty years. 23**:** 345-358.

Forrester, J. W. and P. Senge (1980). Tests for building confidence in system dynamics models. Studies in the Management Sciences (System Dynamics). A. A. Legasto, J. W. Forrester and J. M. Lyneis. Amsterdam: North-Holland**:** 209-228.

Greasley, A. (2005). "Using system dynamics in a discrete-event simulation study of a manufacturing plant." International Journal of Operations & Production Management **25**(6): 534-548.

Grove, P., J. Macleod and D. Godfrey (1998). "Forecasting the prison population." OR Insight **11**(1).

Hair, J., A. Money, M. Page and P. Samouel (2007). Research methods for business. Chichester: John Wiley & Sons.

Han, S., M. Park and F. Pena-Mora (2005). Comparative study of discrete-event simulation and system dynamics for construction process planning. ASCE Proceedings of construction research congress: Broadening Perspectives, San Diego, California, USA.

Helal, M., L. Rabelo, J. Sepulveda and A. Jones (2007). A methodology for Integrating and Synchronizing the System Dynamics and Discrete Event Simulation Paradigms. Proceedings of the 25th International Conference and 50th Anniversary Celebration, July 29 – August 2, 2007, Boston, Massachusetts, USA, The System Dynamics Society.

Hieber, R. and I. Hartel (2003). "Impacts of SCM order strategies evaluated by simulation-based 'Beer Game' approach: the model, concept, and initial experiences." Production Planning & Control **14**(2): 122-134.

Hillier, F. S. and G. J. Lieberman (1990). Introduction to Operations Research 5th ed. ed. New York: McGraw-Hill.

Hurrion, R. D. and R. J. R. Secker (1978). "Visual Interactive Simulation: An Aid to Decision Making." OMEGA - The International Journal of Management Science **6**(5): 419-426.

Hutchinson, S. R. (2004). Survey research. Foundations for Research: Methods of inquiry in education and the social sciences. K. deMarrais and S. D. Lapan. Mahwah, N.J., Lawrence Erlbaum Associates.

Korporaal, R., A. Ridder, P. Kloprogge and R. Dekker (2000). "An analytic model for capacity planning of prisons in the Netherlands." Journal of the Operational Research Society **51**(11): 1228-1237.

Kwak, N. K., P. J. Kuzdrall and M. J. Schniederjans (1984). "Felony Case Scheduling Policies and Continuances - a Simulation Study." Socio-Economic Planning Sciences **18**(1): 37-43.

Lakey, P. B. (2003). A Hybrid Software Process Simulation Model for Project Management. In Proceedings of the International Workshop on Software Process Modeling and Simulation (ProSim '03), Portland, OR.

Landry, M., J.-L. Malouin and M. Oral (1983). "Model validation in operations research." European Journal of Operational Research **14**(3): 207-220.

Lane, D. C. (1995). The folding star: A comparative reframing and extension of validity concepts in system dynamics. The 13th International Conference of the System Dynamics Society, Tokyo, Japan, 111-130.

Lane, D. C. (2000). "Should system dynamics be described as a 'hard' or 'deterministic' systems approach?" Systems Research and Behavioural Science **17**(1): 3-22.

Lane, D. C. (2000). "You Just Don't Understand Me: Modes of failure and success in the discourse between system dynamics and discrete event simulation." LSE OR Working Paper 00.34: 26.

Lane, D. C. and R. Oliva (1998). "The greater whole: Towards a synthesis of system dynamics and soft systems methodology." European Journal of Operational Research **107**(1): 214-235.

Law, A. M. (2007). Simulation modeling and analysis, 4th ed. McGraw-Hill series in industrial engineering and management science. Boston; London: McGraw-Hill.

Lee, Y. H., M. K. Cho, S. J. Kim and Y. B. Kim (2002a). "Supply chain simulation with discrete-continuous combined modeling." Computers & Industrial Engineering **43**(1-2): 375-392.

Lee, Y. H., M. K. Cho and Y. B. Kim (2002b). "A discrete-continuous combined modeling approach for supply chain simulation." Simulation-Transactions Of The Society For Modeling And Simulation International **78**(5): 321-329.

Mak, H.-Y. (1993). System dynamics and discrete event simulation modelling. Department of Statistical and Mathematical Sciences. London, London School of Economics and Political Science**:** 228.

Martin, R. H. and D. Raffo (2000). "A Model of the Software Development Process using both Continuous and Discrete Models." Software Process: Improvement and Practice **5**(2-3): 147 - 157.

McKelvie, D., S. Hadjipavlou, D. Monk, S. Foster, E. Wolstenholme and D. Todd (2007). The use of SD methodology to develop services for the assessment and treatment of high risk serious offenders in England and Wales Proceedings of the 25th International conference of the System Dynamics Society, Boston, USA.

Meadows, D. H. (1980). The Unavoidable A Priori. Elements of the System Dynamics Methods. J. Randers. Cambridge, Productivity Press.

Morecroft, J. (2007). Strategic Modelling and Business Dynamics: a feedback systems approach. Chichester: John Wiley.

Morecroft, J. D. W. and S. Robinson (2005). Explaining Puzzling Dynamics: Comparing the Use of System Dynamics and Discrete-Event Simulation. In Proceedings of the 23rd International Conference of the System Dynamics Society, Boston.

Morecroft, J. D. W. and J. Sterman, Eds. (1994). Modeling for learning organizations. System dynamics series. Portland, Or., Productivity Press.

Music, G. and D. Matko (1999). "Combined simulation for process control: extension of a general purpose simulation tool." Computers in Industry **38**(2): 79-92.

Onggo, B. S. S. (2007). Towards a unified conceptual model representation. Lancaster, The Department of Management Science, Lancaster University Management School**:** 20pp.

Onggo, B. S. S., M. M. Gunal and W. Maden (2008). The conceptual model of a distribution warehouse simulation. In proceedings of 2008 Simulation Conference (SW08), 1-2 April 2008, Worcestershire, England, 129-134.

Oppenheim, A. N. (1968). Questionnaire Design and Attitude Measurement. Heinemann books on Sociology. London: Heinemann.

Oral, M. and O. Kettani (1993). "The facets of the modeling and validation process in operations research." European Journal of Operational Research **66**(2): 216-234.

Oyarbide, A., T. S. Baines, J. M. Kay and J. Ladbrook (2003). "Manufacturing systems modelling using system dynamics: forming a dedicated modelling tool." Journal of advances manufacturing systems **2**(1): 71-87.

Pala, O., J. A. M. Vennix and J. P. C. Kleijnen (1999). Validation in soft OR, hard OR and system dynamics: A critical comparison and contribution to the debate. 17th International Conference of the System Dynamics Society and 5th Australian & New Zealand Systems Conference, Wellington, New Zealand, The System Dynamics Society.

Patrick, J. and N. James (2004). "Process tracing of complex cognitive work tasks." Journal of occupational and organizational psychology **77**(2): 259-280.

Petropoulakis, L. and L. Giacomini (1998). "Development of a hybrid simulator for manufacturing processes." Computers in Industry **36**(1-2): 117-124.

Pidd, M. (2003). Tools for Thinking: Modelling in Management Science, 2nd edition ed. Chichester: John Wiley & Sons.

Pidd, M. (2004). Computer Simulation in Management Science, 5th edition ed. Chichester: John Wiley & Sons.

Powell, S. G. and T. R. Willemain (2007). "How novices formulate models. Part I: qualitative insights and implications for teaching." The Journal of the Operational Research Society **58**(8): 983-995.

Rabelo, L., M. Helal, A. Jones and H.-S. Min (2005). "Enterprise simulation: a hybrid system approach." International Journal of Computer Integrated Manufacturing **18**: 498-508.

Randers, J. (1980). Elements of the System Dynamics Methods. Cambridge: Productivity Press.

Randers, J. (1980). Guidelines for model conceptualisation. Elements of the System Dynamics Methods. J. Randers. Cambridge, Productivity Press.

Richardson, G. P. (1991). Feedback thought in social science and systems theory. Philadelphia. PA: University of Pennsylvania Press.

Ritchie, J. and J. Lewis, Eds. (2003). Qualitative research practice: a guide for social science students and researchers. London, SAGE Publications.

Ritchie, J., J. Lewis and G. Elam (2003). Designing and selecting samples. In Qualitative research practice: a guide for social science students and researchers. J. Ritchie and J. Lewis. London, SAGE Publications**:** 77-108.

Robinson, S. (2001). "Soft with a hard centre: discrete-event simulation in facilitation." Journal Of The Operational Research Society **52**(8): 905-915.

Robinson, S. (2002). "Modes of simulation practice: approaches to business and military simulation." Simulation Modelling Practice and Theory **10**: 513-523.

Robinson, S. (2004). Simulation: The Practice of Model Development and Use. Chichester: John Wiley & Sons.

Robinson, S. (2005). "Discrete-event simulation: from the pioneers to the present, what next?" Journal of the Operational Research Society **56**(6): 619-629.

Robinson, S. (2008). "Conceptual modelling for simulation Part I: definition and requirements." J Oper Res Soc **59**(3): 278-290.

Robinson, S., M. Meadows, J. Mingers, F. A. O'Brien, E. A. Shale and S. Stray (2003). "Teaching OR/MS to MBAs at Warwick Business School: A turnaround story." Interfaces **33**(2): 67-76.

Rouwette, E. A. J. A., J. A. M. Vennix and T. v. Mullekom (2002). "Group model building effectiveness: a review of assessment studies." System Dynamics Review **18**(1): 5-45.

Senge, P. M. (1990). The Fifth Discipline: The art and practice of the learning organisation. London: Random House.

Sheskin, D. J. (2007). Handbook of parametric and nonparametric statistical procedures, 4th ed. ed. Boca Raton Chapman & Hall/CRC.

Siegel, S. (1957). "Nonparametric Statistics." The American Statistician **13**(3): 13-19.

Siegel, S. d. and N. J. Castellan (1988). Nonparametric statistics for the behavioral sciences, 2nd ed. ed. New York; London: McGraw-Hill.

Silverman, D. (2000). Doing qualitative research: a practical handbook. London: SAGE Publications.

Stchedroff, N. and R. C. H. Cheng (2003). Modelling a Continuous Process with Discrete Simulation Techniques and its Application to LNG Supply Chains. In proceedings of the 2003 Winter Simulation Conference, Institute of Electrical and Electronics Engineers, Piscataway, New Jersey**:** pp. 1607.

Sterman, J. D. (2000). Business Dynamics: Systems Thinking and Modeling for a Complex World. London: McGraw-Hill.

Sweetser, A. (1999). A Comparison of System Dynamics and Discrete Event Simulation. In Proceedings of 17th International Conference of the System Dynamics Society and 5th Australian & New Zealand Systems Conference, Wellington, New Zealand.

Swenseth, S. R., J. R. Olson and P. B. Southard (2002). "Extending product profiling through simulation." International Journal of Operations & Production Management **22**(9): 956 - 971.

Taylor, K. and D. Lane (1998). "Simulation applied to health services: opportunities for applying the system dynamics approach." Journal of health services research and policy **3**(4): 226-232.

Tocher, K. D. (1963). The art of Simulation. Electrical Engineering Series. London: The English Universities Press.

Todd, P. and I. Benbasat (1987). "Process tracing methods in decision support systems research: Exploring the black box." MIS Quarterly **11**(4): 493-512.

Usano, R. R., J. M. F. Torres, A. C. Marquez and R. Z. d. Castro (1996). System Dynamics and discrete simulation in a constant work in process system. A comparative study. In the proceedings of the 14th International Conference of the System Dynamics Society, 1996, Cambridge, Massachusetts, USA, 457-460.

Venkateswaran, J. and Y. J. Son (2004). "Design and development of a prototype distributed simulation for evaluation of supply chains." International Journal Of Industrial Engineering-Theory Applications And Practice **11**(2): 151-160.

Vennix, J. A. M. (1996). Group Model-building: Facilitationg Team Learning Using System Dynamics. Chichester: Wiley.

Wilk, M. B. and R. Gnanadesikan (1968). "Probability plotting methods for the analysis of data." Biometrika **55**(1): 1-17.

Willemain, T. R. (1994). "Insights on modeling from a dozen experts." Operations Research **42**(2): 213-222.

Willemain, T. R. (1995). "Model formulation: What experts think about and when." Operations Research **43**(6): 916-932.

Willemain, T. R. and S. G. Powell (2006). "How novices formulate models. Part II: a quantitative description of behaviour." J Oper Res Soc **58**(10): 1271-1283.

Willemain, T. R. and S. G. Powell (2007). "How novices formulate models. Part II: a quantitative description of behaviour." The Journal of the Operational Research Society **58**(10): 1271-1283.

Wolstenholme, E. F. (1990). System Enquiry: a system dynamic approach. Chichester: John Wiley.

Appendix A: The case studies

A1. Fresh Dairies case study

Fresh Dairies deliver six days a week. Doorstep delivery is convenient for the family and for working people. It means no worries about running out of milk. All one needs to do is put their orders and then next morning milk is on their doorstep. Customers are sure that milk is always fresh and also they know that they are doing their bit for the environment. Bottles are reusable, which saves energy and resources. When they can no longer be used they are recycled and made into new bottles.

The Fresh Dairies is a small doorstep delivery company based in Wales, Flintshire. The company is dedicated to fulfil customer demand and requirements. They provide a flexible service, responding to their customer needs, i.e. special orders for the weekend. Three types of milk are provided, whole, semi-skimmed and skimmed milk. Customers have been with Fresh Dairies for many years. New customers are generated by delivering leaflets, special offers, advertising, word of mouth, etc. The company has around 200 customers, who place regular and special orders. Customers purchase 1.5 pints[20] a day. Milk products have a short life and if any products are not delivered in 2 days they are thrown away. Thus, Fresh Dairies cannot carry too much spare stock. They also aim to keep their waste down to 5%. Problems are faced with maintaining service during holidays or when the supplier, Mavel has difficulties with deliveries.

The role of Fresh Dairies lies in the last phase of the whole process, distributing finished products to the customer. The products come from Mavel, a wholesaler for milkmen in Wales and the Southeast of England situated in Newtown in the region of Montgomeryshire. Fresh Dairies receive orders from their customers, add a safety margin (forecasting system) and place orders to their supplier Mavel. Mavel delivers orders for Fresh Dairies to a cold store in Mold. The next morning the Fresh Dairies' drivers pick up the delivery to distribute to the customers. Fresh Dairies employs 4 deliverymen who are responsible to serve a particular area every morning.

* * *

A milk delivery driver for Fresh Dairies, Roy's working day starts well before his customers are out of bed. At 1 a.m. Roy takes his truck to the cold store in Mold and loads it for the morning's deliveries. It takes him about 15 minutes to load the crates from the cooler to the truck. Before going, he needs to know his itinerary for this day's route. Roy enters a small office next to the loading area, where Mike, the customer services manager enters the customer orders into a computer and prints out the updated route sheet, showing how much milk will be delivered to each customer this morning.

[20] 1 pint = approx. 0.5 litres

After having loaded the truck and with route sheets in hand, Roy sets off his route. He pulls up to the first house, 33 Holy Road, and this is his first delivery for this morning. He glances at the route sheets once more, then grabs the order from the back of the truck and dashes to the porch where a small box resides. He quickly places the fresh order inside the box, taking away any empty bottles the customer may have left in return. His truck grinds as the engine starts. Off to the next house. Roy has to deliver to about 50 customers on the route this morning before 7 a.m.

* * *

The supplier, Mavel serves 10 doorstep milk delivery businesses similar in size and service to Fresh Dairies. It has no contact with the end customer. It only takes orders from milkmen and calculates the overall demand, which will consist of their order to the milk processor, Sunrise. Mavel have a long-term relationship with Sunrise, which has been supplying them with milk for over a decade. Sunrise has a big milk processing plant and is situated in Swansea. In the actual plant the unprocessed milk moves around through numerous tubes and machines and gets pasteurized and homogenized. It is then separated into skimmed, semi-skimmed and whole milk and ready to be bottled. Skimmed milk gets a red seal, semi-skimmed green and whole a blue seal. The differently coloured lids allow deliverymen to identify at a glance what kind of milk is in the bottle.

* * *

Fresh Dairies does not keep any contacts with the milk processor. Back in the past the manager had considered the option to contact Sunrise and get milk directly from the production plant. This would mean that instead of placing orders to Mavel 2 days in advance, Fresh Dairies would have to place orders one week in advance. However, as Sunrise is situated further down South Wales in Swansea, a very long drive from Mold, travelling everyday to Sunrise premises to load on average 1,800 litres of milk would prove to be expensive for Fresh Dairies.

The company is wary of the new situation in the doorstep delivery market and is concerned to maintain sales over the long term. The number of doorstep delivery companies is declining as demand for doorstep deliveries is falling. The application of advanced technologies has significantly increased milk shelf-life and people now more and more buy milk at supermarkets. Fresh Dairies have called you to help them understand their supply chain. They have asked you to give them advice on how to deal with falling demand and suggest ways of making their supply chain more effective. Would the option of getting a step forward and getting orders from Sunrise help Fresh Dairies keep their business running for the next 20 years?

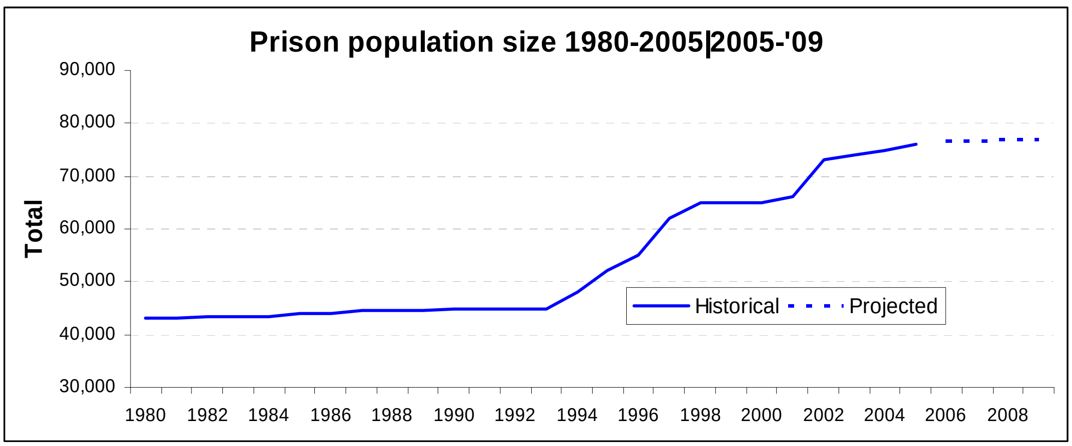
Based on: Ace Dairies case study from Waters, D. (2003) Logistics: An Introduction to Supply Chain Management. Palgrave Macmillan, New York.

A.2 The prison population case study for the model building study

Case Study: Prison Population

I. Introduction

Prisons are correction institutions where offenders are confined for a period of time according to the crime committed. The UK has the highest imprisonment rate in the European Union. The prison population reached 76,000 people in 2005. An inherent stability has been generally observed in the prison population system. However, in 1993 the government introduced a new line of tougher sentences and as a result in the period between 1993 and 2003, a dramatic increase was observed. The number of offenders sentenced to custody increased by 83% and the length of sentences given increased by 31%. According to Home Office projections the prison population is expected to stabilise. On the other hand, 82 out of the 139 prisons in England and Wales are overcrowded. The occupancy level has reached 120%. A part of released prisoners, re-offend and get back to jail known as recidivists. Recidivism is considered to be the main reason for the existence of overcrowded prisons.



II. The problem: Prison overcrowding

Prison overcrowding causes disruption and undermines the rehabilitation effect of the Prison Service. The problem of overcrowded prisons has come to the attention of the authorities and action needs to be taken. The overall objective of the criminal justice system is to prevent crime and/or deter its repetition. When considering solutions to the problem one needs to consider ways of improving the prison system in line with its objectives. Looking into the problem of prison overcrowding, the authorities are currently considering two possible scenarios, either to increase the current prison capacity and so facilitate the introduction of stiffer rules, or the alternative of reducing the size of the prison population by introducing alternatives to jail and/or enhancing the social support provided to prisoners.

Your Task

Using your preferred simulation tool build a model of the prison population to be used as a decision-making tool by government authorities. For a starting point, assume that it is the 1st January 2005.

Consider two types of offenders, petty and serious offenders. Both types of offenders receive a sentence and enter the system as first-time offenders, where they serve time in prison according to the sentence length received. After serving time in prison offenders are released. Some released prisoners re-offend and go back to jail, known as recidivists. The majority of prison sentences passed by courts are short. Petty offenders enter the system at a higher rate, due to a higher rate of offending, but receive a shorter sentence length. Such prisoners have the highest reconviction rates amongst released prisoners. They are more likely to commit new offences due to the fact that prison may serve as a school for crime. Furthermore, after being released a few petty prisoners commit even more serious crimes and then are re-convicted as serious offenders. Serious offenders represent a small proportion of the total offender population. They receive a longer sentence, but have a lower offending rate compared to petty offenders. Serious offenders have the lowest rates of recidivism and are the least likely to be rearrested for another violent crime.

In figure 1, a graphical representation of the prison system is provided.

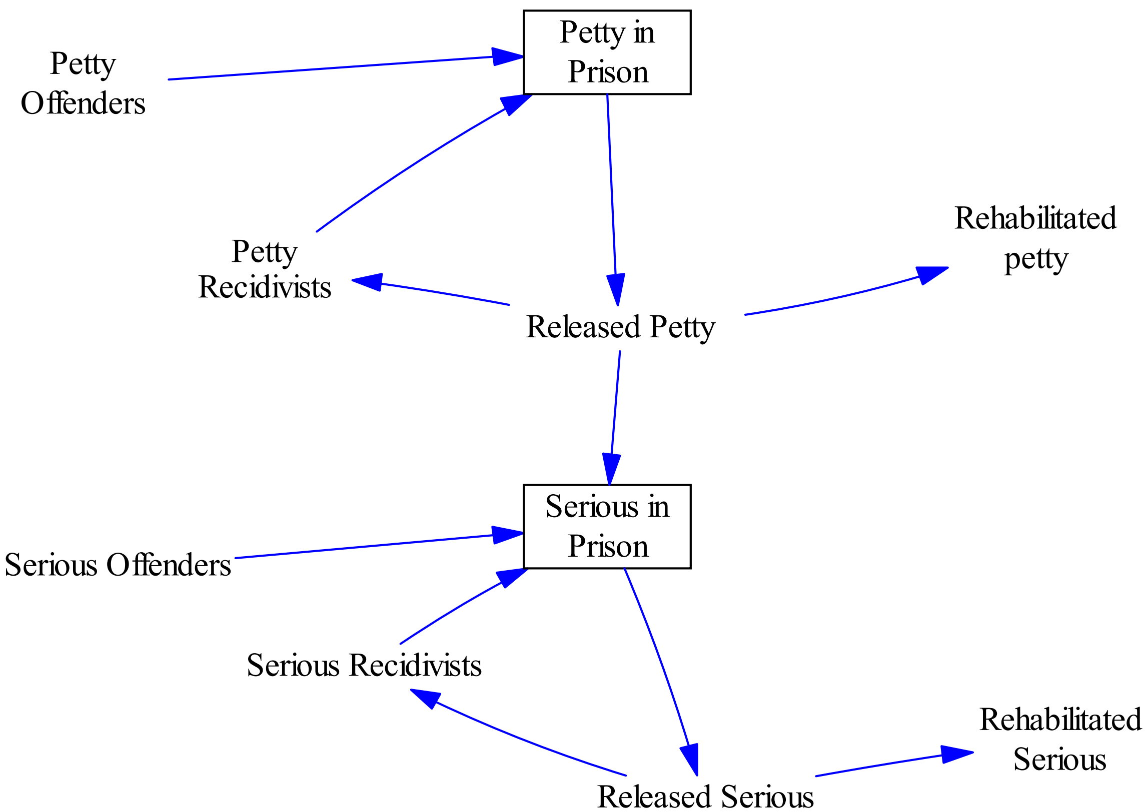


Figure 1: Prison population model diagram

Some possible assumptions to consider:
When building the simulation model, feel free to make your own assumptions about any model parameters. Some possible initial assumptions you could consider are:

- Current prison population totals 76,000 prisoners consisting of 50,000 petty and 26,000 serious offenders.
- Prison admissions: on average 3,000 petty and 650 serious offenders enter prison per year.
- The average sentence length is 5 years for petty offenders and 20 years for serious offenders.
- Recidivists return back to jail on average after 2 years.

If you have any queries during the model building exercise I'd be happy to answer them.

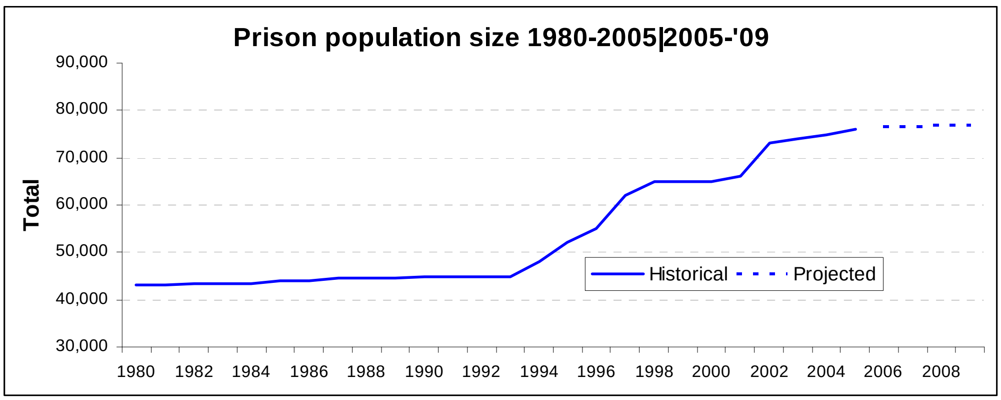
THANK YOU FOR YOUR TIME!

A.3 The prison population case study for the model use study

Case Study: Prison Population

I. Introduction

Prisons are correction institutions where offenders are confined for a period of time according to the crime committed. The UK has the highest imprisonment rate in the European Union. The prison population reached 76,000 people in 2005. An inherent stability has been generally observed in the prison population system. However, in 1993 the government introduced a new line of tougher sentences and as a result in the period between 1993 and 2003, a dramatic increase was observed. The number of offenders sentenced to custody increased by 83% and the length of sentences given increased by 31%. According to Home Office projections the prison population is expected to stabilise.
On the other hand, 82 out of the 139 prisons in England and Wales are overcrowded. The occupancy level in March 2005 was at 106%. 60% of all prisoners are reconvicted (become 'recidivists') within two years of release. The reconviction rates are higher for less serious offences. According to Home Office statistics, 4 out of 5 shoplifters and car thieves appear before the courts within two years of leaving.



II. The problem: Prison overcrowding

Prison overcrowding causes disruption and undermines the rehabilitation effect of the Prison Service. The problem of overcrowded prisons has come to the attention of the authorities and action needs to be taken. When considering solutions to this problem there are two main issues to be considered: "Is it worth building new prisons?" or "Would the use of alternatives to prison reduce the burden?" Two options are available: either to follow the route of stiffer rules (longer sentences), resulting in an increase in the number of inmates or introducing alternatives to jail, e.g. electronic tagging, which would result in decreased numbers of prisoners in the system.

The overall objective of the criminal justice system is to prevent crime and/or deter its repetition. So in making a decision one needs to consider ways of improving the prison system in line with its objectives. Thus the main criteria that need to be considered are: prison capacity and the efficiency of prisons in rehabilitating wrongdoers. When considering a solution to the problem, one needs to bear in mind that the community requires protection, but at what expense? A balance should be found between the costs incurred and the level of protection achieved.

III. Potential options for policy makers

The two available lines of action for policy makers are to either increase or reduce prison capacity. We will assume here that the policies described below are the only options available.

1. **Policies resulting in increase of prison capacity**

If the objective of introducing stiffer rules is considered, supported by a decision to increase the current prison capacity, the available policies/options are as follows:

- **Tougher sentences** for serious offences, e.g.: murder, public violence, sex offences, etc.
  Way of implementing**: introduction of longer sentences for serious offenders.**
  Potential benefits**:** public protection from dangerous elements.

- **Sentence enhancement**. In line with the "prison works" theory, it is believed that increased use of prison reduces crime. Penal laws are amended by introducing longer sentences for various offences. Justification of this line of action is mainly based on the so called incapacitation effect i.e. that while in jail offenders cannot break the law. Another factor considered is the deterrence effect i.e. longer sentences serve as a deterrent for offenders to engage in further criminal behaviour.
  Way of implementing**:** Introduction of **longer sentences**.
  Potential benefits: **reduced recidivism**, but overcrowded prison population.

- **Increased imprisonment sentences**. In line with this, it is believed that the increase in the number of imprisonment sentences given from courts, will act as a deterrent to crime, and will thus reduce the level of crime in the future.
  Way of implementing: **Increased prison admissions**.
  Potential benefits: reduced level of crime in the future, public protection, but overcrowded prisons.

2. **Policies resulting in reduced prison capacity**

On the other hand, if considering reducing the current prison capacity the available options/policies are:

- **Introduction of electronic tagging (ET)**. ET is an alternative punishment to imprisonment that restricts the freedom of the criminal by attaching an electronic device on the offenders' body. It is usually given to inmates with short term sentences who are eligible for early release and tagging. An issue related to ET is that shorter sentences might have an adverse effect through reduced deterrence.
  Ways of implementing ET:
  ✓ Use of electronic monitoring where short custodial sentences would otherwise be imposed and thus would **reduce prison admissions**. About 1,400 remand receptions per annum could be avoided through the use of electronic tagging.
  Potential benefits**:** reduced prison admissions for petty offences
  ✓ **Shorter sentences.** Early release of offenders with a condition of electronic monitoring. Short term prisoners may be released early with a condition of ET.
  Potential benefits: reduced number of prisoners (petty and serious).

The introduction of ET is directly related to **reduced levels of recidivism.** In both cases ET can help in tackling offending behaviour and prevent re-offending by providing close supervision and/or treatment to released prisoners.

There are also **cost benefits** related to ET. Sending someone to prison for a year costs a minimum of £24,000. Tagging an offender costs just £2,000 and eases the strain on the prison system.

- **Reforming prison institutions**. Reform in prison can potentially improve the quality of the prison service. Reform focused on rehabilitation, with an integration of the social work practice, the introduction of training & rehabilitation programmes in prison, has much to offer to inmates, who after serving time in prison will potentially integrate more smoothly into society and will thus not return to crime.
  Potential benefits: Higher quality of prison services, resulting in lower re-offending rates - **reduced recidivism.**
- **Commutation of sentence** – reduction of the legal penalty due to a prisoner's good behaviour.
  Potential benefits: **Reduction of time in prison**, resulting in reduced number of inmates.
- **Intensive supervision**. Community services and closer supervision, provided to offenders after serving their time in prison.
  Potential benefits: Higher quality of community support to ex-prisoners and thus **reduced levels of recidivism.**

A.4 Prison population model handout

A Model of the Prison Population

A simple model of prison population has been created, which predicts the size of the overall prison population for the next 30 years. In this model, two types of offenders are included, petty and serious offenders. Both types of offenders receive a sentence and enter the system through Prison Admissions as first-time offenders, where they serve time in prison according to the sentence length received. After serving time in prison offenders are released. Some released prisoners re-offend and go back to jail, known as recidivists. The majority of prison sentences passed by courts are short. Petty offenders enter the system at a higher rate, due to a higher rate of offending, but receive a shorter sentence length. Such prisoners have the highest reconviction rates amongst released prisoners. They are more likely to commit new offences due to the fact that prison may serve as a school for crime. Furthermore, after being released a few petty prisoners commit even more serious crimes and then are re-convicted as serious offenders.

Serious offenders represent a small percentage of the total offender population. They receive a longer sentence, but have a lower offending rate compared to petty offenders. Serious offenders have the lowest rates of recidivism and are the least likely to be rearrested for another violent crime.

The prison population model with initial assumptions made and initial values, is presented in figure 1 overleaf. The model starts with an initial overall prison population of 76,000 in 2005. By introducing various policies (refer to **Potential options for policy makers**, in the case study), one can observe their impact on the size of prison population up to 2035.

Looking into the problem of prison overcrowding, the authorities are currently considering two possible scenarios, either to increase the current prison capacity and so facilitate the introduction of stiffer rules, or the alternative of reducing the size of the prison population by introducing alternatives to jail and/or enhancing the social support provided to prisoners. In line with the overall objectives of the criminal justice system and the two separate scenarios, budget limitations are as follows:

- Scenario 1: While increasing prison capacity, a maximum of 100,000 prison places can be supported.
- Scenario 2: While reducing prison capacity, a maximum of 50,000 places in prison can be supported.

Your Task

Taking the role of the government consulting services and based on the policies mentioned in the "Prison Population" case study, consider separately the two possible scenarios and use the model provided to suggest how the problem of prison overcrowding can be solved. Consider the wider implications of your findings in relation to the policies that might be implemented. Bearing in mind the specific budget constraints, what policy changes should be implemented in order to achieve an optimal prison service?

When implementing various scenarios, you will be able to change the model input variables according to the policy chosen. Make your own assumptions about the relationship between the policy and its resulting effect on the input data (i.e. the implementation of severe/longer sentences may act as a deterrent to crime and result in reduced re-offending rates and thus recidivism, but may also result in reduced offending rates and thus prison admissions). The introduction of one policy may mean more than one change in the input data. Also a mixture of policies can be introduced at the same time.

Prepare a short presentation (2 slides) of your findings in PowerPoint and place this on the network drive.

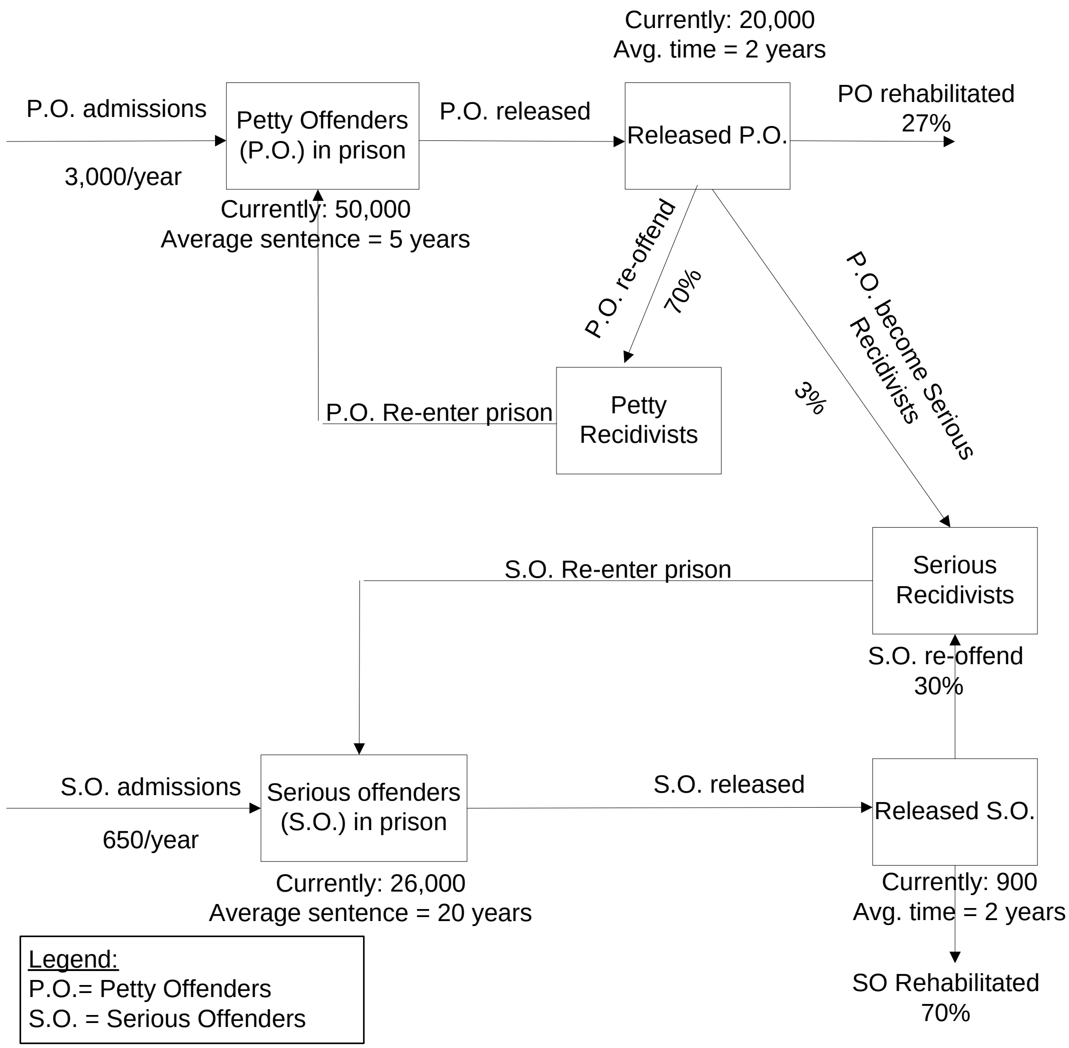


Figure 1: Prison population model diagram

Critical model assumptions:

- Annual admission rates and sentence length represent average numbers.
- This is a simple model of the prison population, assuming that other social factors remain unchanged for the next 30 years (i.e. there are no political shifts, no changes in crime patterns or appearance of new types of crime, etc.)
- The available current prison capacity is 72,000 beds.
- The model does not account for the effect that different policies have on crime levels, but only for the effect on the prison population size.

A.5 Explanations about the DES model

Running the Model

How to run the model

When opening the Prison Population model, the screen as in Figure 1 will appear. In the left hand window, the Prison model is created based on the description provided in the handout: A model of the prison population. On the top right hand side window the input variables are presented and in the window below, model outputs are displayed. In the bottom right hand corner, the clock displays the current year and date.

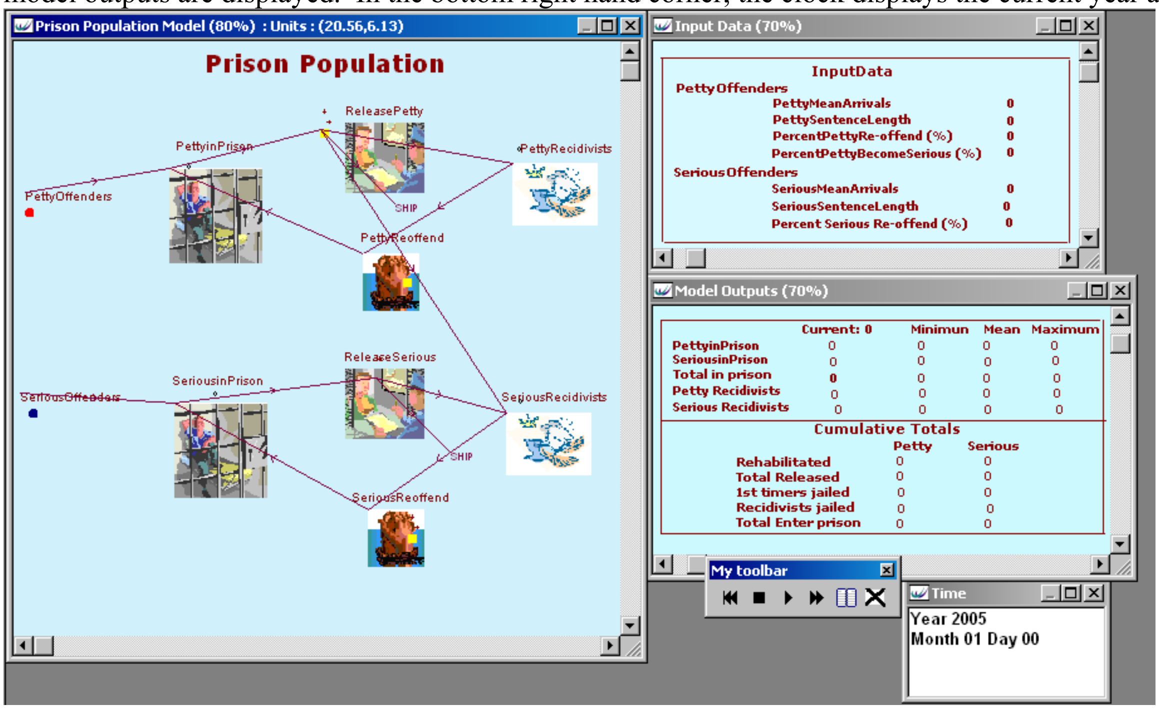


Figure 1: Model screenshot

The toolbar called "My toolbar" (See Figure 2) can be used in order to control the simulation model run. The following buttons have been included in it:

- Run button – by clicking on it the model runs at normal speed
- Stop button – stops the simulation
- Batch button – runs the simulation in high speed with no display
- Begin button – resets the simulation to year 2005
- Open graphs button – opens the graphs window. See Figure 3.
- Close graphs button – closes the graphs window.

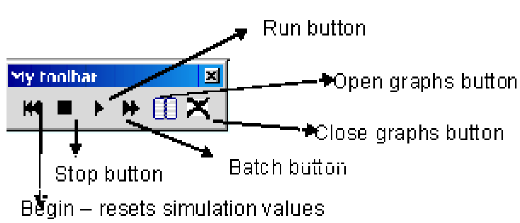


Figure 2: My toolbar

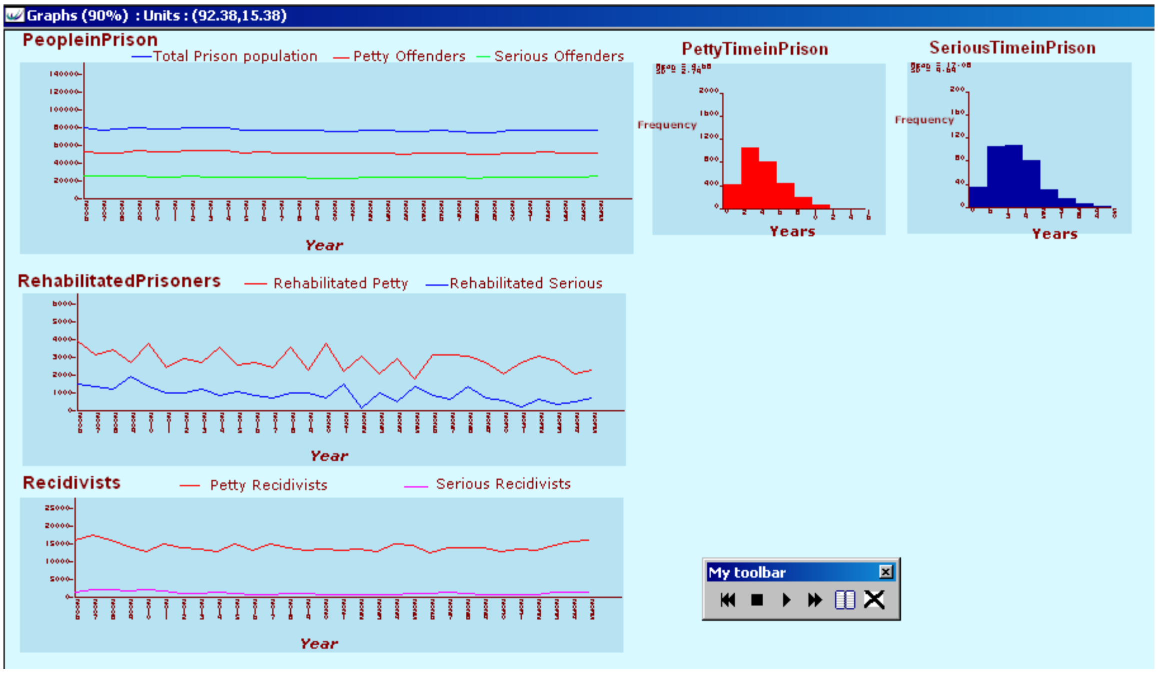


Figure 3: Screenshot of Graphs window displayed when clicking Open graphs button

When clicking the "Run" button, a pop-up window as in Figure 4, will appear warning that after clicking the "Start Simulation" button the simulation model will start.

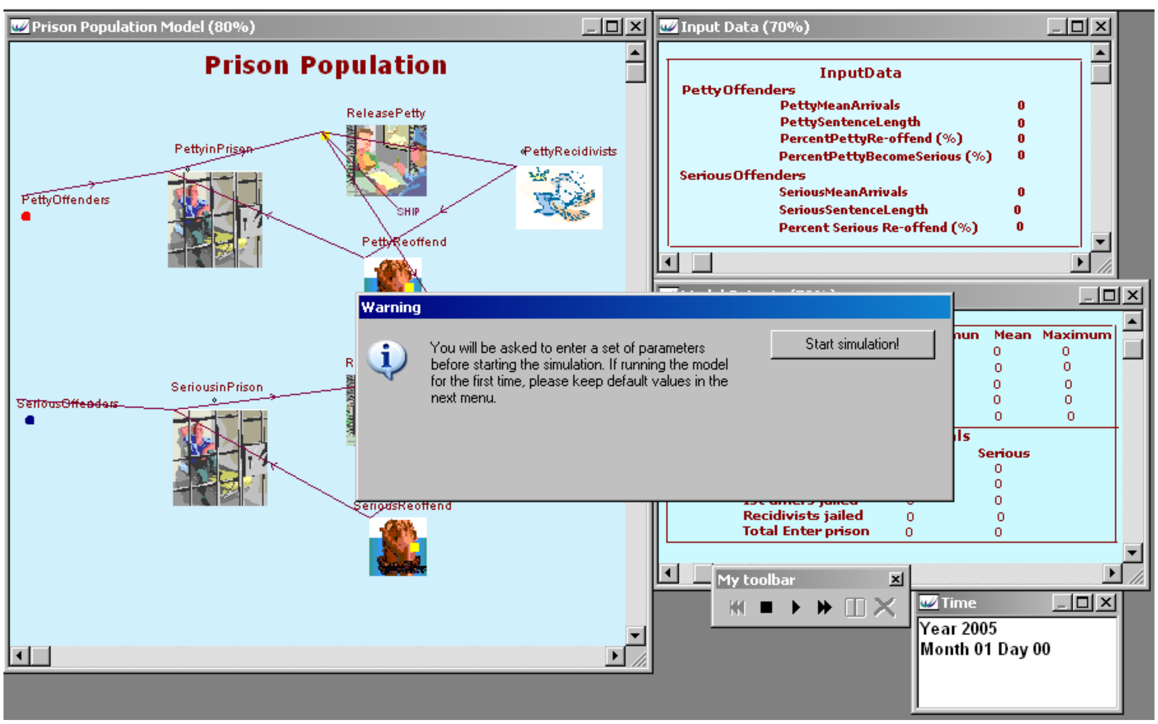


Figure 4: First Pop-up window

By clicking on the "Start simulation!" button, the "Model Input data" window appears (Figure 5) asking the user to enter the input data for the simulation. If running the simulation for the first time keep the default parameters (please see diagram in handout "A model of the prison population"), and then click OK.

When re-running the model the user is asked to choose different parameters in line with possible scenarios (as explained in the Prison Population case study in the sub-section: Potential options for policy makers). After entering the values for each parameter in the window "Model Input data" click the "OK" button and the simulation runs until year 2035. In order to speed up the simulation run, the user can click on "Batch" button on "My toolbar".

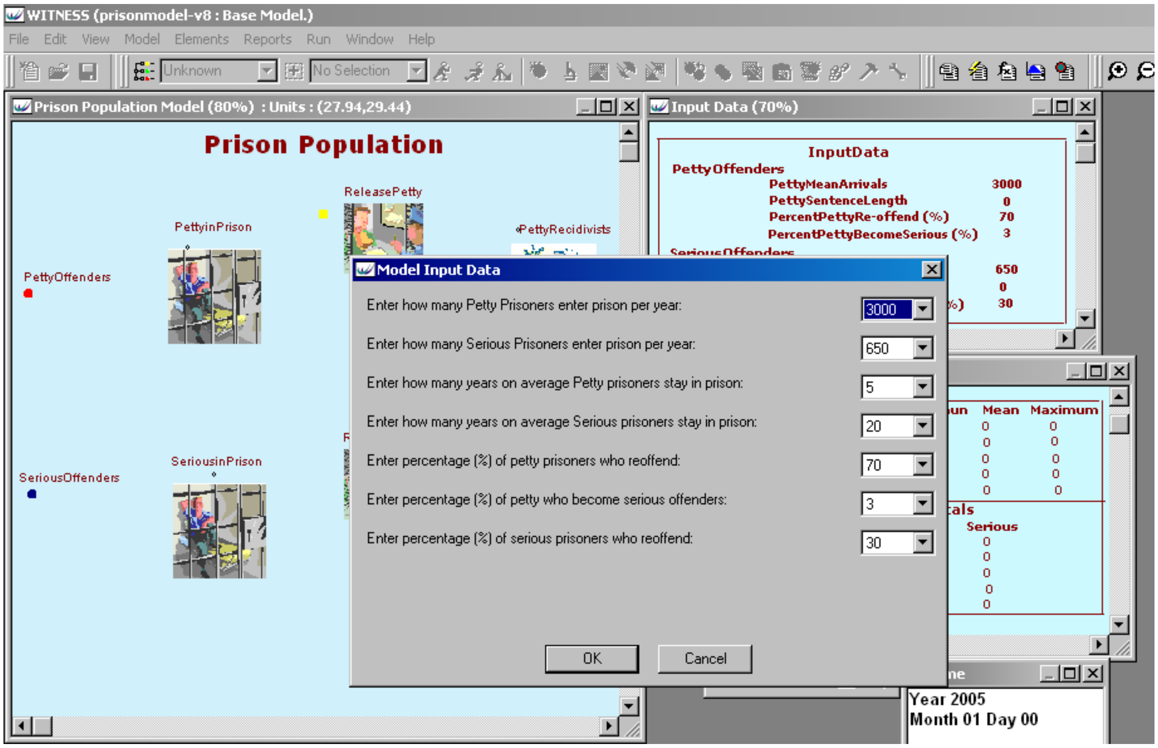


Figure 5: Second pop-up window, Model input data window

After running the model, the results will be displayed in the Model Outputs box. In order to observe the overall system behaviour, the user can open the Graphs window by clicking on the "Open graphs" button (Figure 2). By clicking on the "Close graphs" button the graphs window will close and the initial model window as in Figure 1 will appear.

To re-run the model under new settings, click on the button "Begin", which will reset all the numbers and will turn the simulation clock to year 2005.

Remember, after each run, the user needs to reset the simulation, otherwise it will run for ever, and the Input data menu will not appear.

A.6 Explanations about the SD model

Running the Model

How to run the model

When opening the Prison model, the screen as in Figure 1 will appear. This is the home page, which gives a brief introduction to the simulation model and to the pages available. The Prison model is created based on the description provided in the handout: **A model of the prison population**. By clicking on the picture or the Control Panel link, the user can get to the **Control Panel** page (see Figure 2), which is the working environment where users can enter values and check the results.

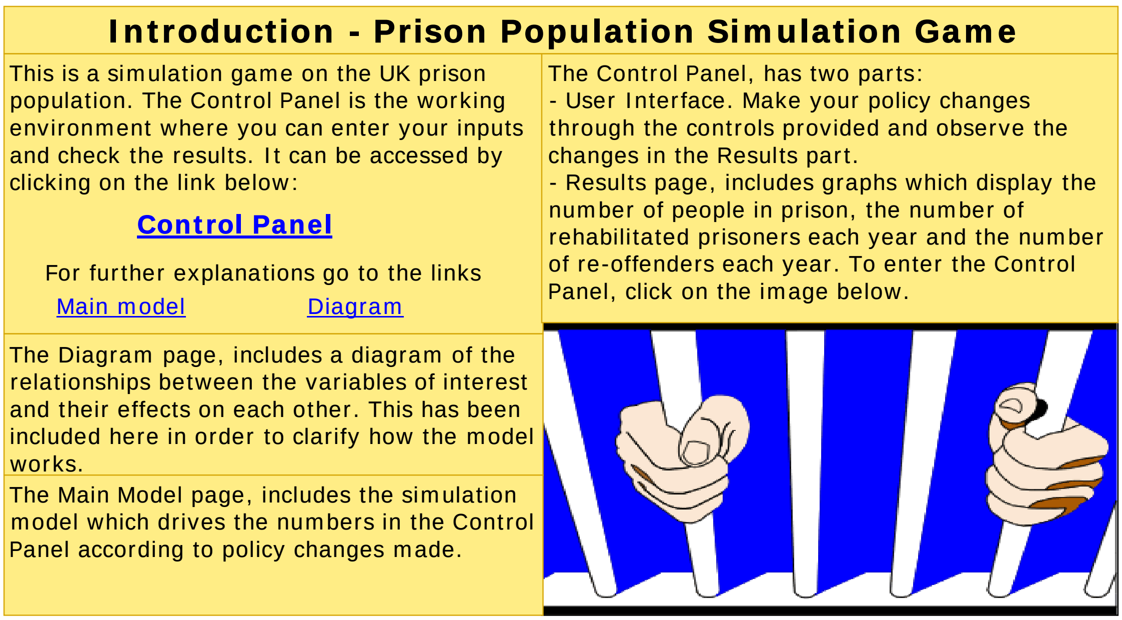


Figure 1: Simulation game Homepage

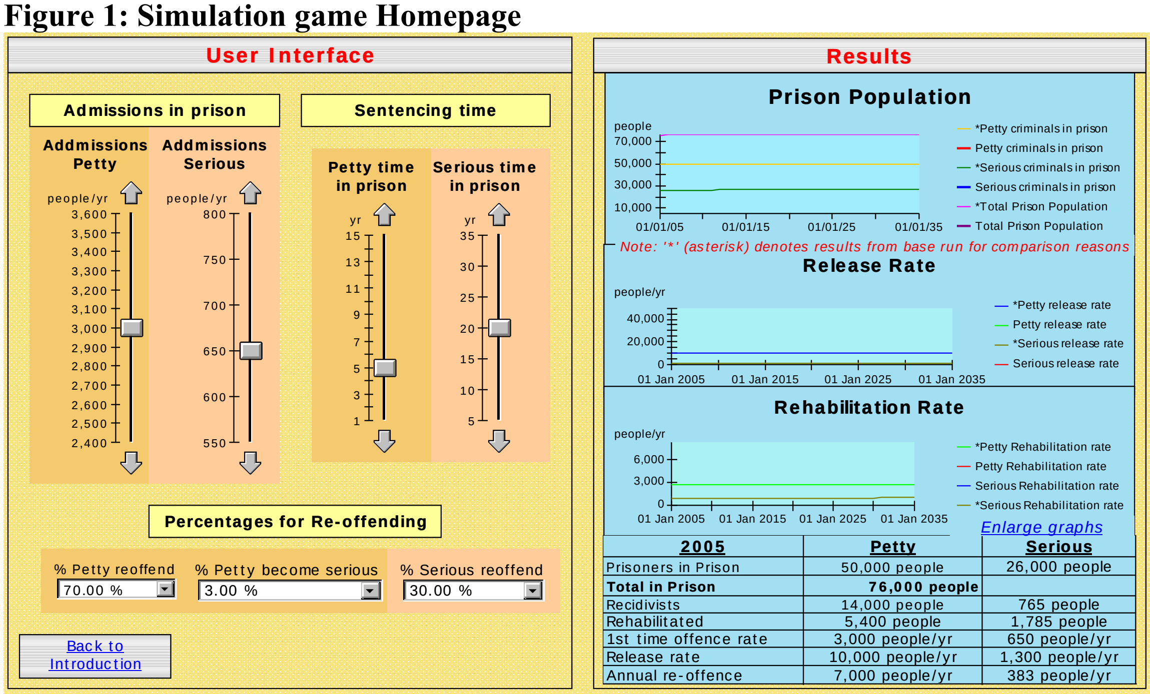


Figure 2: Control Panel

The Control Panel page has two parts, the Input Data part on the left hand-side and on the right is the Results page, where the outputs are displayed through graphs and the table at the bottom.

In order to run the simulation, the user should use the Simulator toolbar (see Figure 3), which is found on the top of any page. It is advisable to use the Simulator Toolbar, once the user is on the Control Panel page. A quick description of the buttons and their functionality follows below:

- Back & Forward buttons – By clicking on this buttons, you can go back and forward between pages.
- Homepage – click on this button, to get to the Homepage.
- Run button – Click on it to run the simulation. It is advisable to click on this button once you are in the Control Panel page.
- Reset simulation button – resets the simulation to 1st January 2005, but keeps the changes in the input variables made in the previous run. **You'll need to reset the simulation before changing the values in the User Interface.**
- Step run button – runs the simulation in steps, by year. It is advisable to use this button, when you want to observe what changes happen every year.
- Copy area/picture – When clicking this button, a dotted line rectangle appears on the right of the cursor and keeping the right mouse clicked you can select the area. After selecting the area of interest it gives you the option to copy it and then paste it in any document.
- Advanced commands – Clicking on advanced commands button, a drop-down menu appears where the most important command is "Restore permanent variables".
- Restore permanent variables – by clicking on the command Restore Permanent Variables, the initial inputs are restored into the simulation (initial values as displayed in the Prison population diagram in "A model of the prison Population")

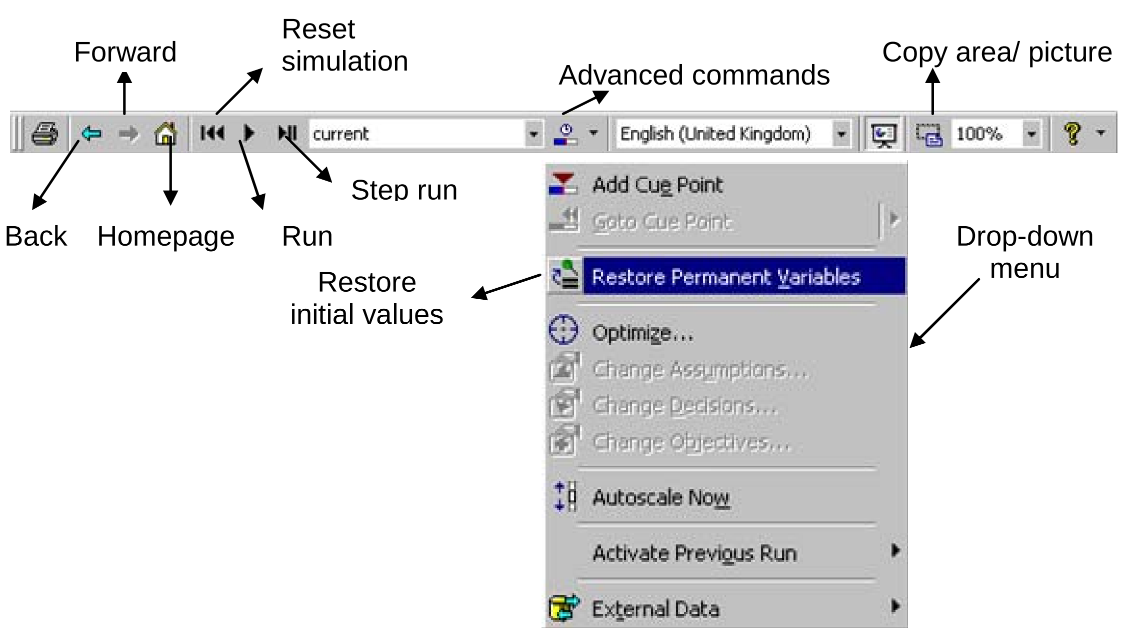


Figure 3: The simulator toolbar

When considering specific scenarios, the new numbers should be entered in the User Interface part (see Figure 4) in the Control Panel page. The initial values for Admissions in prison, Sentencing and Percentages for re-offending, are set based on the initial values described in the Prison population diagram in "A model of the prison Population". In order to change the values for Admissions in prison and Sentencing time the user can click on the top and bottom arrows to change the position of the bar along the slider. To change the values of Percentages for Re-offending, the user can choose any value from the drop down menu. **Please remember,**

in order to re-run the model, it is necessary to click the "Reset simulation" button on the simulator toolbar (Figure 3) and then enter the new changes. After clicking the Reset simulation button, the date will go back to 1st of January 2005, but the values for Admissions, Sentencing and Percentage of re-offending will be still the same as set in the previous run.

.

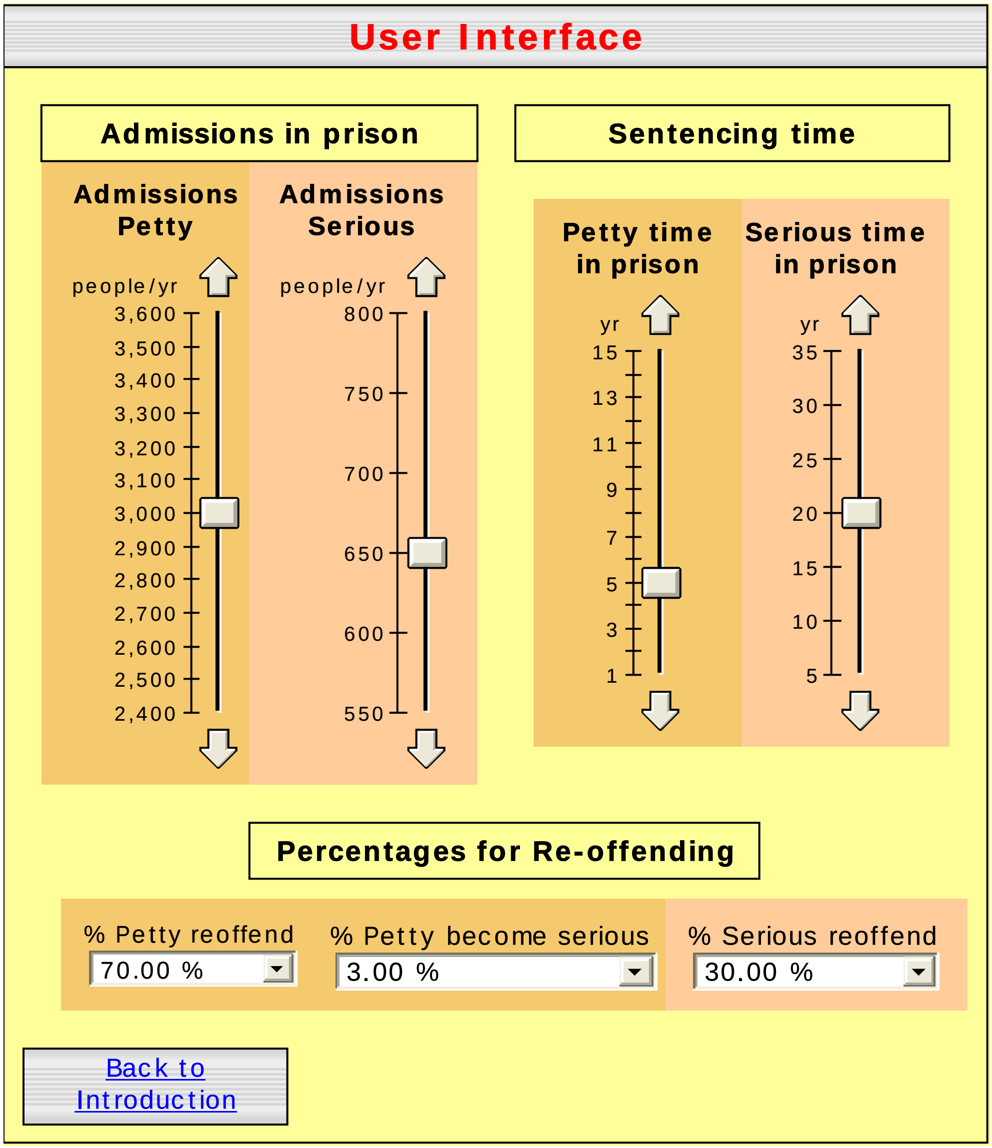
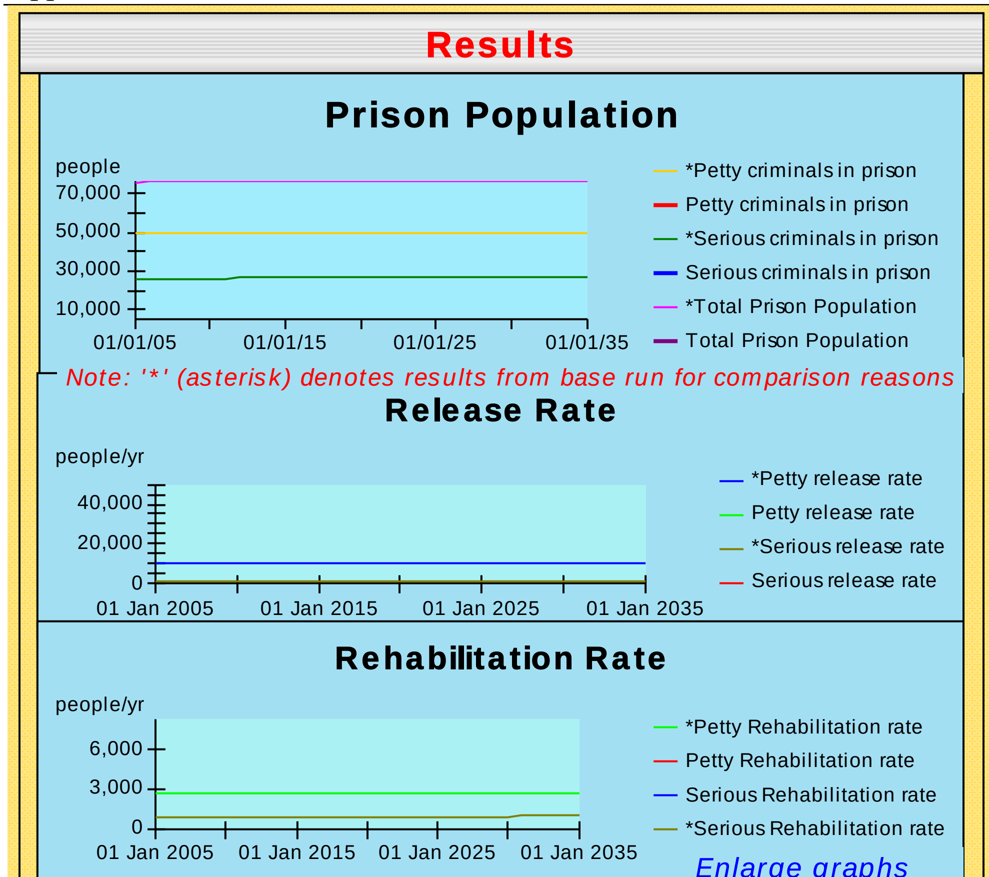


Figure 4: User interface

After entering the desired values in the User Interface part, the user can observe the outcomes in the "Results" page (Figure 5). In the results page there are 3 graphs displayed, and at the bottom there is a table with numbers. If clicking on the "Enlarge graphs" link, the user can zoom out the graphs (see Figure 6) and make necessary observations. By clicking on the button "Back" it goes back to the Control Panel.



2005	Petty	Serious
Prisoners in Prison	50,000 people	26,000 people
Total in Prison	**76,000 people**	
Recidivists	14,000 people	765 people
Rehabilitated	5,400 people	1,785 people
1st time offence rate	3,000 people/yr	650 people/yr
Release rate	10,000 people/yr	1,300 people/yr
Annual re-offence	7,000 people/yr	383 people/yr

Figure 5: Results

In the graphs, the line with an “*” (asterisk) represent the results for the base run, i.e. the model running with the initial numbers as explained in the handout “A model of the prison Population”.

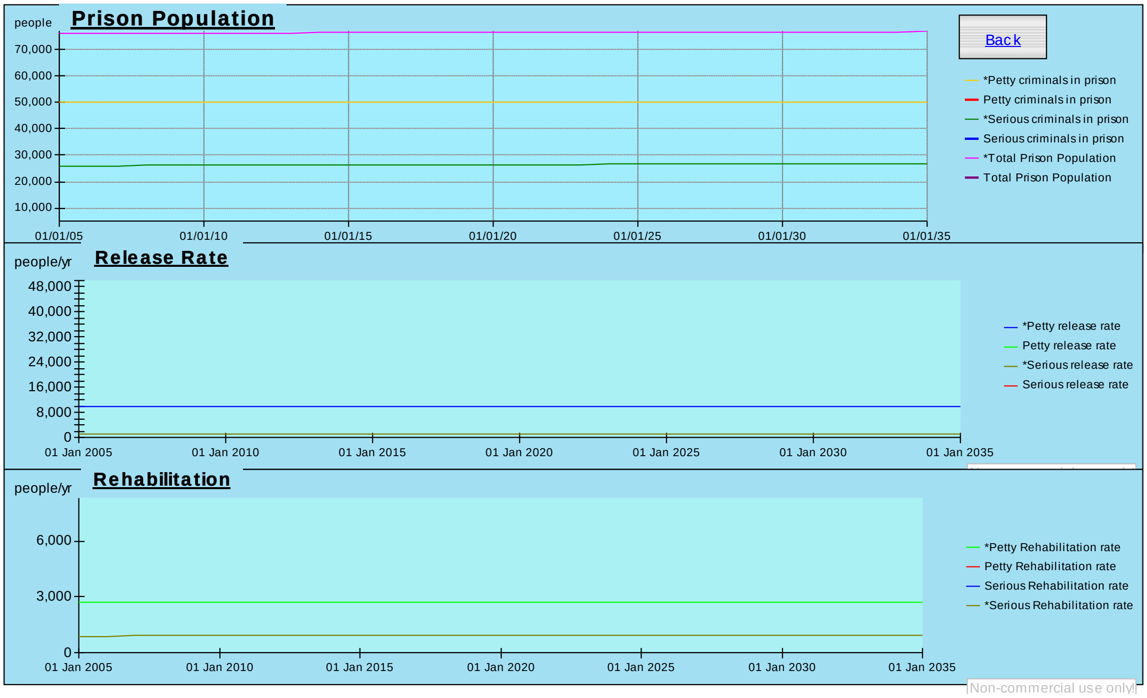


Figure 6: Zooming out graphs only

When re-running the model the user is asked to choose different parameters in line with possible scenarios (as explained in the Prison Population case study in the sub-section: Potential options for policy makers). After entering the values for each parameter in the User Interface and hit the button Run, the simulation runs until year 2035. In order to see the changes by year one can push the button “Step Run”. To re-run the model under new settings, click on the button “Reset simulation”, which will reset the clock to year 2005.

In each page, there is a “Back to Introduction” button, which re-directs the user to the Homepage. Two more pages are available, the Main model where the model that drives the numbers is included. In the Diagram page, a diagram of the model is included, which displays the relationships between the different parameters and variables of the model.

Tip: At the end of a simulation run with a satisfactory performance, the user can save the Prison model file and keep it for a later reference by giving it a new name. The saved file will keep the values entered and also the results, so long as one closes it down after saving it and opens again the original Prison Model file.

A.7 Survey questionnaire

I. Personal details

1. In what industry does your organisation work in? (Please tick)

☐ Manufacturing	☐ Financial Services	☐ Trade (Wholesale/Retail)
☐ Community & Public Services	☐ Business Services, (i.e. Consulting, Law, etc.)	☐ Transport & Communications
☐ Construction	☐ Energy & Mining	☐ Other: ____________________

2. What is the functional area you work in within your organisation? (Please tick)

☐ Human Resources	☐ Computing/IT	☐ Sales/Marketing
☐ Procurement	☐ Research & Development	☐ Production/Operations
☐ Finance/Accounts	☐ Customer Services	☐ Other: ____________________

3. In the hierarchy of your company, which level of management are you currently involved in?

☐ Executive/Managing Director	☐ Director (Middle management)	☐ Manager/Analyst
☐ Other *(Please specify)*: ____		

4. Have you ever used a simulation model before?
   A. ☐ Yes ***(go to question 5)***
   B. ☐ No ***(go to section II)***

5. If you have answered YES to question 4, which simulation package have you used?

☐ Simul8	☐ Witness	☐ Powersim
☐ Vensim	☐ Extend	☐ iThink
☐ ProModel	☐ Arena	☐ Other: _____________________

II. Opinion about the prison population simulation model

a. Model understanding& complexity

Please answer the following questions based on your experience of running the prison population model in your syndicate group. The aim of this set of questions is to assess how comprehensive the model provided was.

1. How well do you feel you understand how the model works?

Understand very little	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	Understand very well

2. Please specify to what extent you feel you understand the following parts of the model?
   a. The relationship between variables

Understand very little	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	Understand very well

   b. The structure of the model

Understand very little	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	Understand very well

   c. How to use the model

Understand very little	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	Understand very well

   d. Model outputs/results

Understand very little	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	Understand very well

3. According to the level of importance, please rank from 1 to 3, where 1 is most important and 3 least important, which of the following factors, helped you understand the model?
   a. Paper-based description of model ☐
   b. Visual display of the model ☐
   c. Animation as the model runs ☐

4. How would you rate the level of detail of the prison simulation model?

Very detailed	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	Very high level

5. What do you think is the main source of complexity in the prison simulation model?
____

b. Model validity

This section deals with your opinion about the credibility of the prison simulation model.

1. To what extent do you feel the simulation model is representative of the prison population system as described in the case study?

Very little	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	Very much

2. To what extent do you feel the model generates realistic outputs, given the case study information?

Very little	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	Very much

3. How confident would you feel in using this model for decision making?

Very little	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	Very much

c. Model usefulness

Please answer the following questions with respect to the prison population model. The aim is to identify your opinion about the usefulness of the prison simulation model.

1. To what extent do you feel using the prison simulation model enhanced your learning about the effect of various policies?

Very little	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	Very much

2. To what extent do you feel using the prison simulation model helped you think strategically about the prison population problem?

Very little	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	Very much

3. In what other contexts might a similar model be used? Please name a few. Why is it relevant?
____

4. To what extent do you feel the prison simulation model facilitated the communication of ideas and suggestions throughout your group discussions?

Very little	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	Very much

d. Opinion about the simulation results

In this section, we intend to derive your opinion about the results of the prison population model.

1. When examining the model results, which did you find most useful?
   a. Table with numbers ☐ *(Tick both, if they were equally useful)*
   b. Graphs ☐

2. How did you find the interpretation of results?

Very straightforward	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	Very difficult
Any further comments? (*Please specify):* ______________________________						

3. a. On a scale from 1 to 5, how useful did you find the graphs?

Very little	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	Very much

   b. What was the main learning point from the graphs? ______________________________

4. When examining the results, to what extent did you try to determine the underlying factors that cause changes in the results?

Very little	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	Very much

5. Which result (s) do you think was most useful in evaluating the prison system performance?
____

e. Overall opinion about the simulation model

1. Would you like to make any other comments about the prison simulation model?

   ____
   ____
   ____

Appendix B: Other materials used in the model building study

B1. Initial instructions read out to participants for the VPA exercise

In this experiment we are interested in what you say to yourself as you perform a simulation modeling exercise. In order to do this we will ask you to THINK ALOUD as you work on the problem given. What I mean by think aloud is that I want you to say out loud *everything* that passes to your head, no matter how irrelevant it may seem, as you build the model. Just act is if you were alone in the room speaking to yourself. If you are silent for any length of time I will remind you to keep thinking aloud.

I will give you a case study, which describes a particular situation in the UK prisons and I would like you to build a model using your preferred simulation tool. I would like you to talk aloud CONSTANTLY from the moment that you finish reading the case study until you have built a final version of your model. I don't want you to try to plan out what you say or try to explain to me what you are saying. Let your thoughts speak as if though you were really thinking out loud. Do you understand what I want you to do?

B2. Warm up exercises

Good, before we turn to the real experiment, we will start with some practice problems. I want you to talk aloud while you do these problems.

1. Multiplication

First, I will ask you to multiply two numbers in your head.
So talk aloud while you multiply 24 times 36!

Good!

2. Anagram

Now I would like you to solve an anagram. I will show you a card with scrambled letters. It is your task to find an English word that consists of all the presented letters. For example, if the scrambled letters are KOBO, you may see that the letters spell the word BOOK. Any questions?

Please "talk aloud" while you solve the following anagram!

"EPNHPA" or "CEPART"

Good!

3. And now another practice problem: "How many windows are there in your parents' house?" I want you to think aloud as before when you think about the question.

   Good!

4. And a last exercise: Several people are standing in a row. Ann is standing rightmost. Eve is standing to the left of Carl. Bob is standing between Ann and Carl. David is standing immediately to the right of Eve. Who is standing immediately to the right of David? I want you to think aloud as before when you think about the question.

   Good!

B3. Detailed coding scheme used in VPA

First-level code	Second-level code	Explanations
1. Problem structuring		What is the problem, modelling objectives, referring to the end users, what will the model be used for, model generalisation (re-use of model in the future and other contexts), etc. Talking about having discussions with the clients.
2. Conceptual modelling		
	i. Conceptual diagram	Do modellers develop a conceptual diagram? What is mentioned regarding the conceptual diagram?
	ii. Model contents and people	Thoughts about parts of and people in the model. It also includes the definition of variables in the model, thoughts about setting boundaries or including additional issues in the system.

	iii. Systems view	Thoughts about the system, taking either an analytic or a holistic view[21].
	iv. Feedback	Thoughts about the feedback structure of the system, causal relationships between variables or parts of the system.
	v. Considerations	Thoughts about how long to run the model for, software to use or even, considering the use of simulation. Considerations about using simulation or the specific simulation approach. It also includes considerations about improving or extending the current model.
3. Model coding		
	i. Coding	Talking about the model and creating the code (including: stocks & flows or queues and processes/activities, writing coding

[21] *Note*

- Analytic: deals with individual relationships between variables
  - ➢ Definition: separating a whole into its elemental parts or basic principles (www.thefreedictionary.com)
  - division into elements or principles (Wikipedia)
  - dealing with or treating the whole of something or someone and not just a part (Cambridge dictionary)
  - The analytic approach seeks to reduce a system to its elementary elements in order to study in detail and understand the types of interaction that exist between them. (http://pespmc1.vub.ac.be/ANALSYST.html)

  - ➢ In comparison literature:
    - breaks a system down into its constituent parts (Brailsford, 2001)
    - related to detail, involving considerably fine data, multiple attributes, stochasticity, etc. (Lane, 2000)
    - interested in particular processes (Sweetser)

- Holistic: deals with inter-relationships between variables
  - ➢ Definitions of Holistic - emphasizing the organic or functional relation between parts and the whole (www.thefreedictionary.com)
    - concerned with wholes rather than analysis or separation into parts (www.answers.com)
    - Holistic relates to, or is concerned with, complete systems rather than with the details or parts that make up the system. The term also refers to a person who tends to be good at understanding the large picture and at relating large areas of information to each other. (http://www.sil.org/lingualinks/literacy/ReferenceMaterials/glossaryofliteracyterms/WhatIsHolistic.htm)
    - consider a system in its totality, its complexity, and its own dynamics (http://pespmc1.vub.ac.be/ANALSYST.html)
  - ➢ In comparison literature:
    - emphasis on dynamic complexity (Lane, 2000)
    - focus on the analysis of the system (Sweetser, 1999)

		routines, etc.).
	ii. Flexibility	References to particular problems with creating the model (e.g. large population size, etc.)
	iii. Modelling structures	How is the aspect of decision making modelled? Types of modelling structures: theoretical (borrowed from other disciplines) or practical (derived from discussions with the client).
	iv. Interface	Modelling the interface part of the model, so that the model is clearly displayed on screen, with the end user in mind.
	v. Documentation	Thoughts about documenting the code to help the user or for future use.
	vi.Units	References to the time unit, steps, etc., or the measuring units.
	vii. Other	Issues related to the appearance of the model.
4. Data inputs		Instances where users talk about the data inputs.
	i. Supplied data	References to the already provided data inputs, including quantitative and/or qualitative data: - Quantitative: references to model parameters, statistics, making calculations, etc. - Qualitative: references to the graph provided in the case study, soft variables, etc.
	ii. Extensions	Statements referring to the need for data collection, references to missing data, anecdotal data versus concrete/observable data. Includes also assumptions made for missing data.
	iii. Randomness	Use of statistical or empirical distributions and references to them.
	iv. Relationships between variables	References to linear and/or non-linear relationships between variables in the model.
5. Verification & Validation		Instances where modellers refer to testing the model including the following activities: - Checking that the model runs as expected - Checking the code

		- Visual checks - Inspecting model results - Comparing results to the real world.
6. Verification & Validation	i. Equilibrium	References to creating a model in a steady or in equilibrium state
	ii. Model results	references to model outputs (specific qualitative or quantitative outputs of the model)
	iii. Scenarios	"What if" analysis, references to using the model to explore different possible scenarios.
7. Implementation		References to implementing the findings, putting the solution into practice or learning/improved understanding.
8. Other		Includes statements about the practice of the modeller or comments about the model. Or anything else which is not covered in the codes above.

Appendix C: Timeline plots

C.1 Data from Kolmogorov-Smirnov tests

Problem structuring							
	DES			SD		S1-S2 (Vertical distance)	
	Raw data	S1 (Cumulative		Raw data	S2 (Cum distribution)		
DES2, DES4	0, 0	0.40		-	0	0.40	
	-	0.4	SD4	6	0.2	0.20	
DES5	42	0.6		-	0.2	0.40	
DES1	53	0.8		-	0.2	**0.60**	**M**
	-	0.8	SD3	67	0.4	0.40	
DES3	125	1	SD5	125	0.6	0.40	
	-	1	SD2	160	0.8	0.20	
	-	1	SD1	431	1	0.00	

No significant difference in the verbalisations between the two groups.

Conceptual modelling							
	DES			SD		S1-S2 (Vertical distance)	
	Raw data	S1 (Cum. distribution)		Raw data	S2 (Cum. distribution)		
DES1	208	0.20		-	0	0.20	
DES4	287	0.4		-	0	0.40	
DES2	300	0.6		-	0	0.60	
DES5	808	0.8		-	0	**0.80**	**M**
	-	0.8	SD3	915	0.2	0.60	
	-	0.8	SD2	977.5	0.4	0.40	
	-	0.8	SD1	1086	0.6	0.20	
DES3	1768	1		-	0.6	0.40	
	-	1	SD5	1861	0.8	0.2	
	-	1	SD4	1911	1	0.0	

Significant difference in the verbalisations, SD modellers verbalise more.

Model coding						
	DES			SD		S1-S2
	Raw data	S1 (Cum. distribution)		Raw data	S2 (Cum. distribution)	(Vertical distance)
	-	0.00	SD2	557	0.2	-0.20
	-	0	SD1	684	0.4	-0.40
	-	0	SD5	749	0.6	-0.60
	-	0	SD3	838	0.8	**-0.80** M
DES1	1224	0.2		-	0.8	-0.60
DES4	1364	0.4		-	0.8	-0.40
DES2	1578	0.6		-	0.8	-0.20
	-	0.6	SD4	1538.5	1	-0.40
DES3	3098	0.8		-	1	-0.2
DES5	4372	1		-	1	0.0

Significant differences in the verbalisations between the two groups, DES modellers verbalise more.

Data Inputs						
	DES			SD		S1-S2
	Raw data	S1 (Cum. distribution)		Raw data	S2 (Cum. distribution)	(Vertical distance)
	-	0.00	SD3	279	0.2	-0.20
	-	0	SD2	436.5	0.4	-0.40
DES1	469	0.2		-	0.4	-0.20
	-	0.2	SD5	525	0.6	-0.40
	-	0.2	SD1	526	0.8	**-0.60** M
DES2	651.5	0.4		-	0.8	-0.40
	-	0.4	SD4	805.5	1	**-0.60** M
DES4	825	0.6		-	1	-0.40
DES5	1133.5	0.8		-	1	-0.2
DES3	1266	1		-	1	0.0

No significant differences in the verbalisations between the two groups.

Verification & Validation							
	DES			SD		S1-S2 (Vertical distance)	
	Raw data	S1 (Cum. distribution)		Raw data	S2 (Cum. distribution)		
	-	0.00	SD2	246	0.2	-0.20	
	-	0	SD5	248	0.4	-0.40	
	-	0	SD4	324	0.6	-0.60	
	-	0	SD1	333	0.8	-0.80	
	-	0	SD3	415	1	**-1.00**	**M**
DES1	507	0.2		-	1	-0.80	
DES4	1086	0.4		-	1	-0.60	
DES5	1089.5	0.6		-	1	-0.40	
DES2	1127.5	0.8		-	1	-0.2	
DES3	1734	1		-	1	0.0	

Significant difference in the verbalisations between the two groups, DES modellers verbalise more.

Results & Experimentation							
	DES			SD		S1-S2 (Vertical distance)	
	Raw data	S1 (Cum. distribution)		Raw data	S2 (Cum. distribution)		
DES4	42	0.20		-	0	0.20	
DES2	117	0.4		-	0	**0.40**	**M**
	-	0.4	SD4	176	0.2	0.20	
DES3	247	0.6		-	0.2	**0.40**	**M**
	-	0.6	SD2	309	0.4	0.20	
	-	0.6	SD5	331	0.6	0.00	
DES5	458	0.8		-	0.6	0.20	
DES1	474	1		-	0.6	**0.40**	**M**
	-	1	SD3	625	0.8	0.2	
	-	1	SD1	695	1	0.0	

No significant difference in the verbalisations between the two groups.

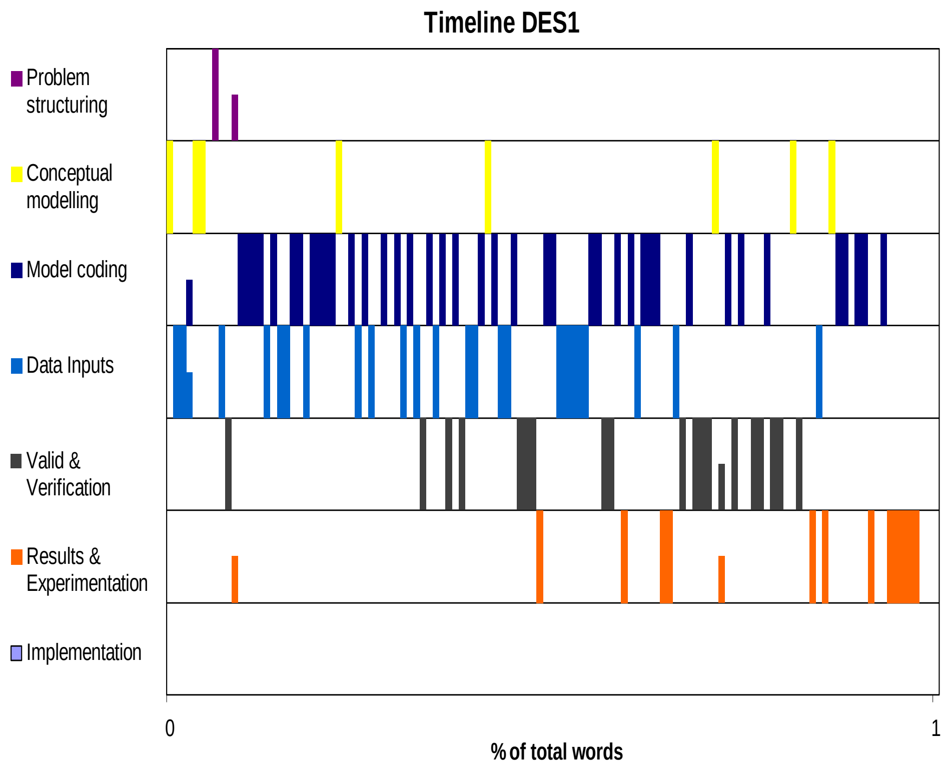
Implementation							
	DES			SD		S1-S2 (Vertical distance)	
	Raw data	S1 (Cum. distribution)		Raw data	S2 (Cum. distribution)		
DES1, DES2, DES4, DES5	0, 0, 0, 0	0.8	SD2, SD5	0, 0	0.4	**0.40**	**M**
	-	0.8	SD3	25	0.6	0.20	
DES3	38	1		-	0.6	**0.40**	**M**
	-	1	SD4	41	0.8	0.20	
	-	1	SD1	172	1	0.00	

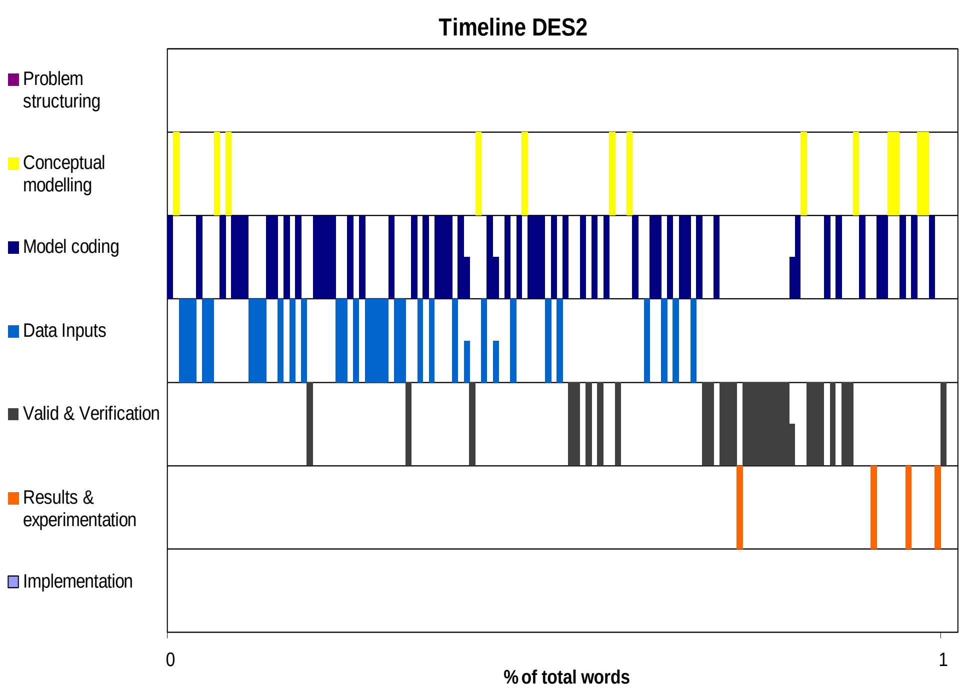
No significant differences in the verbalisations between the two groups.

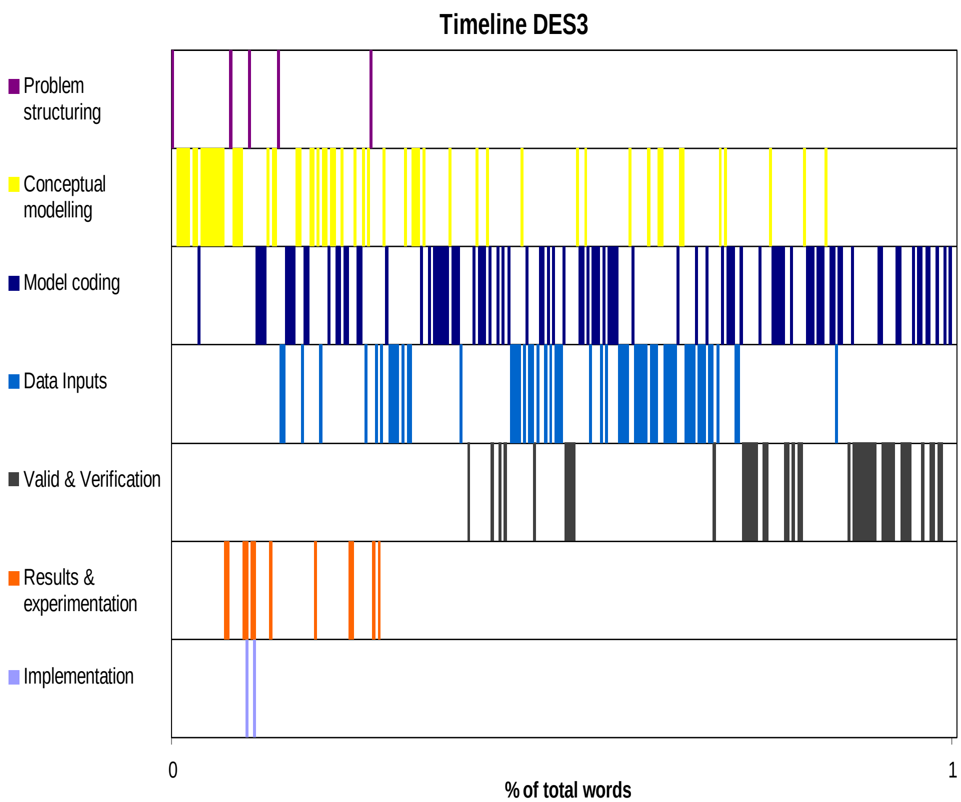
Total verbalisations						
	DES			SD		S1-S2
	Raw data	S1 (Cum. distribution)		Raw data	S2 (Cum. distribution)	(Vertical distance)
	-	0	SD2	2762	0.2	-0.20
DES1	3068	0.2		-	0.2	0.00
	-	0.2	SD3	3294	0.4	-0.20
DES4	3714	0.4		-	0.4	0.00
DES2	3877	0.6		-	0.4	0.20
	-	0.6	SD1	3960	0.6	0.00
	-	0.6	SD5	4022	0.8	-0.20
	-	0.6	SD4	4858	1	**-0.40** M
DES5	8172	0.8		-	1	-0.2
DES3	8521	1		-	1	0.0

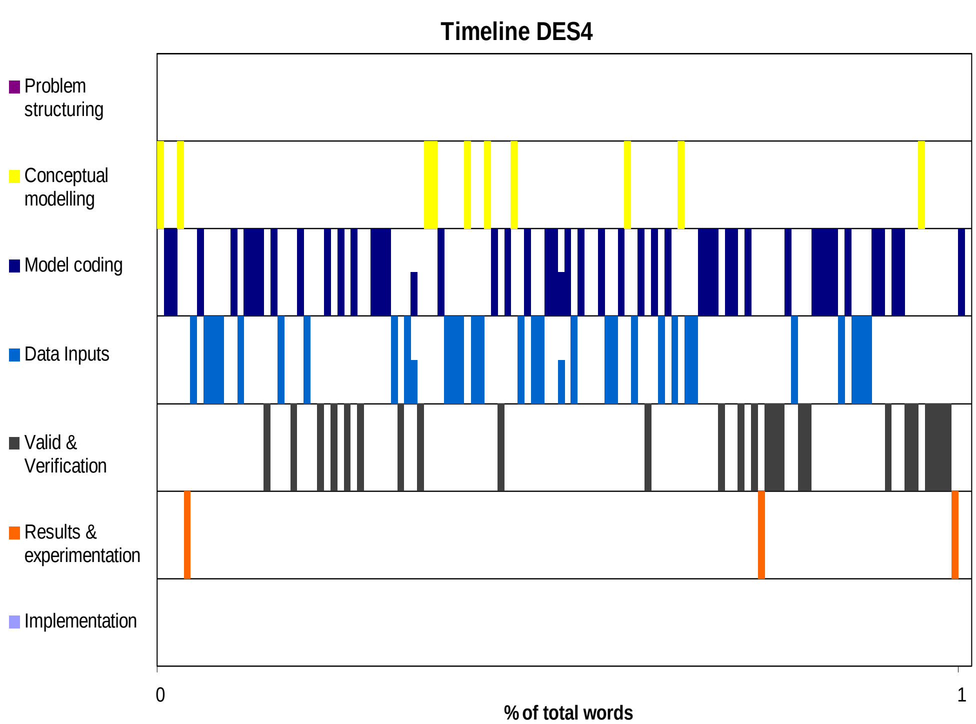
No significant differences in the verbalisations between the two groups.

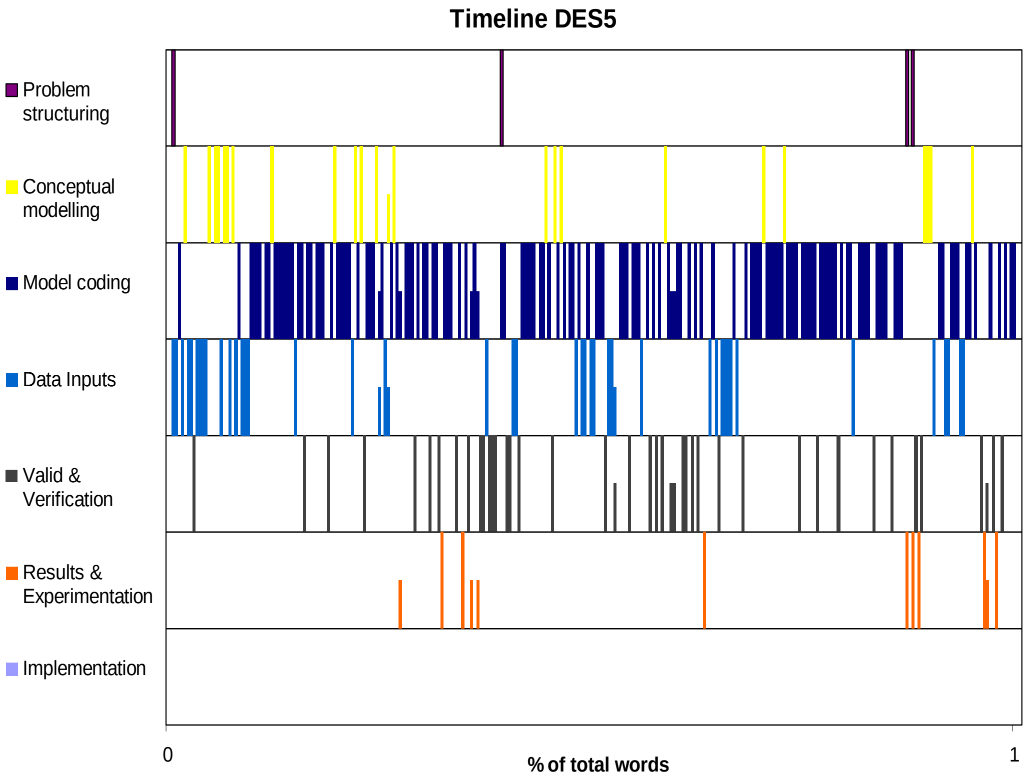
C.2 DES timeline plots



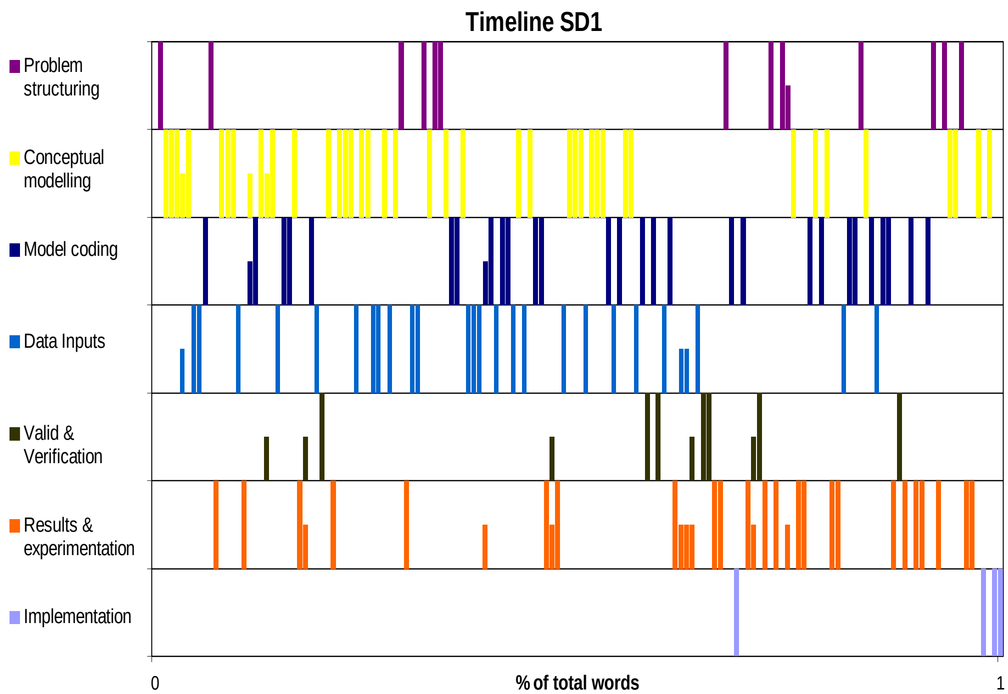


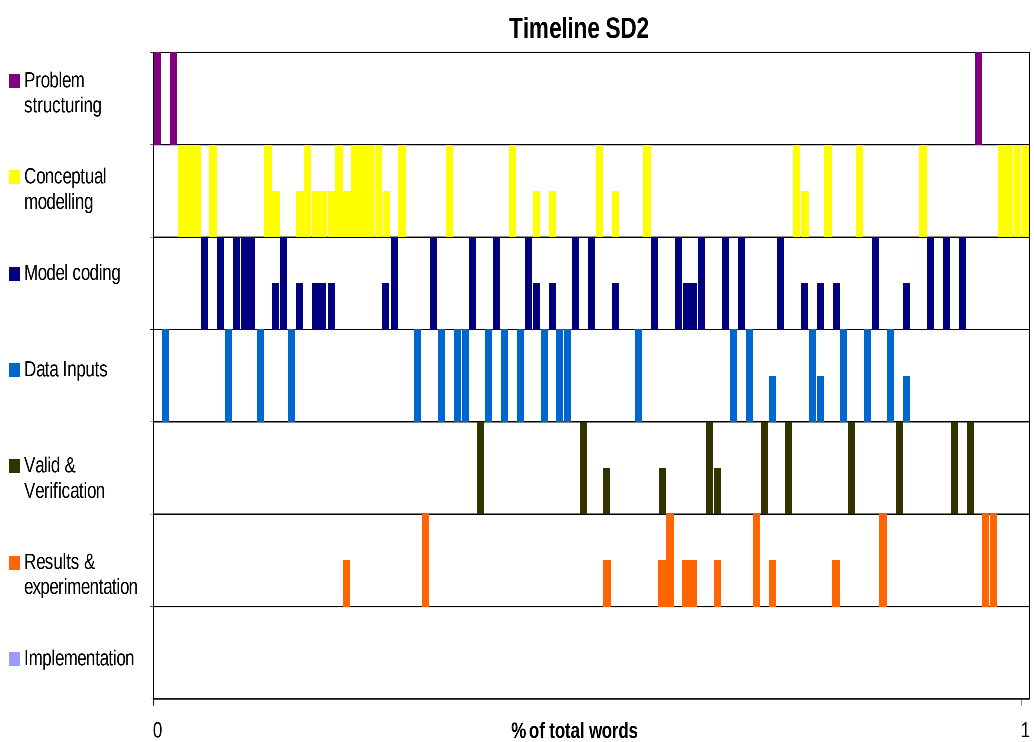

Timeline DES3
Problem structuring
Conceptual modelling
Model coding
Data Inputs
Valid & Verification
Results & experimentation
Implementation
0
1
% of total words

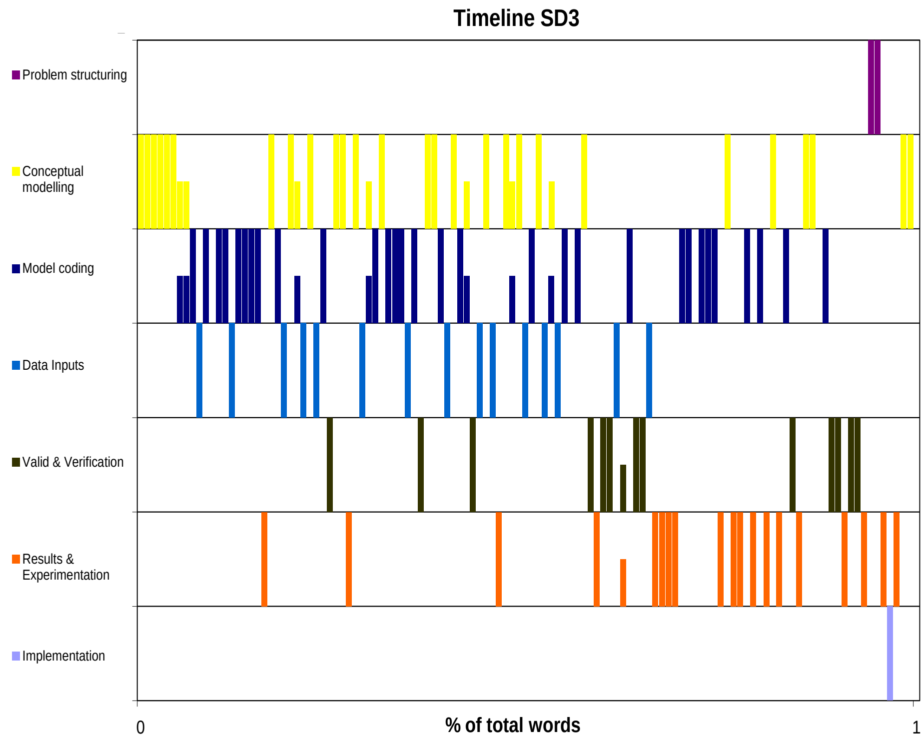

Timeline DES4
Problem structuring
Conceptual modelling
Model coding
Data Inputs
Valid & Verification
Results & experimentation
Implementation
0
1
% of total words

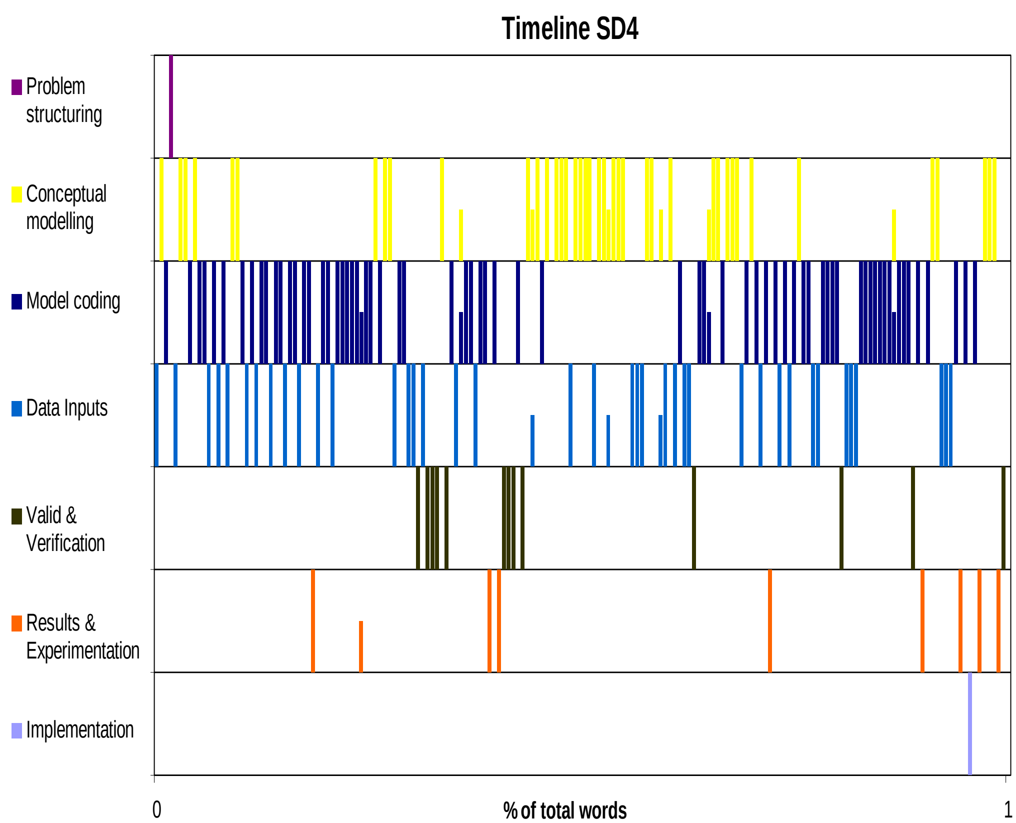

Timeline DES5
Problem structuring
Conceptual modelling
Model coding
Data Inputs
Valid & Verification
Results & Experimentation
Implementation
0
1
% of total words

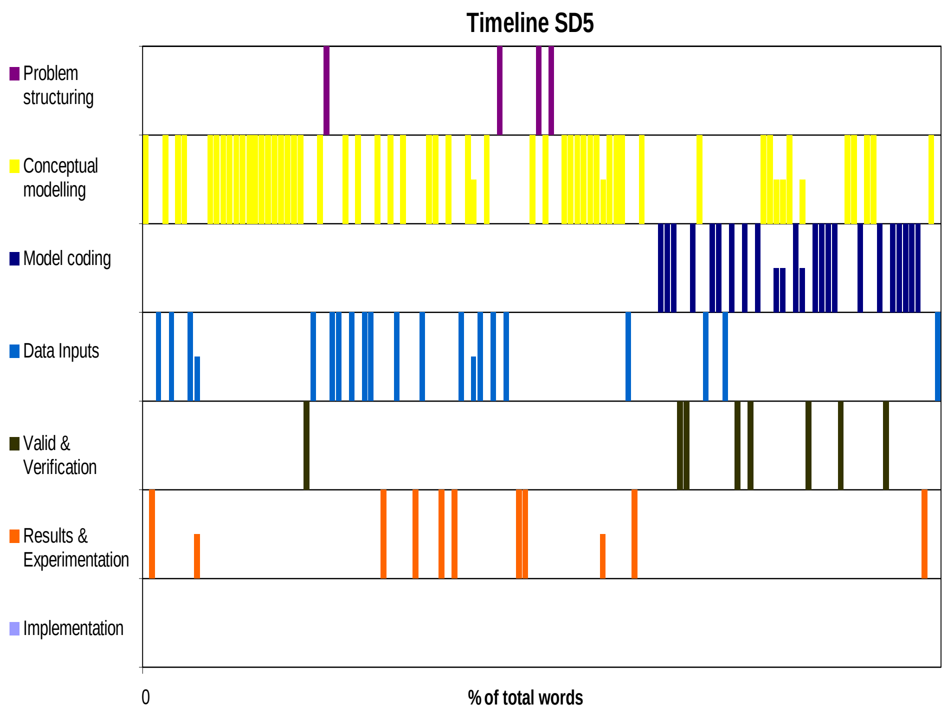
C.2 SD timeline plots






Timeline SD3
Problem structuring
Conceptual modelling
Model coding
Data Inputs
Valid & Verification
Results & Experimentation
Implementation
0
% of total words
1


Timeline SD4
Problem structuring
Conceptual modelling
Model coding
Data Inputs
Valid & Verification
Results & Experimentation
Implementation
0
% of total words
1


Timeline SD5
Problem structuring
Conceptual modelling
Model coding
Data Inputs
Valid & Verification
Results & Experimentation
Implementation
0
% of total words