View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by Lancaster E-Prints

Conceptual modelling and the project process in real simulation projects: a survey of simulation modellers

Abstract

A survey was used to obtain information on the processes and methods used by simulation experts in real projects. The 102 survey respondents answered questions about their most recent simulation project. This paper presents some of the survey results, focussing mainly on conceptual modelling and the pattern of time allocation to different topics. There are a wide range of findings that include the modellers making changes to the initial conceptual model during subsequent tasks in most of the projects usually by adding complexity, model coding taking on average about twice the time of other topics, and the topics generally occurring in single blocks of time (at the resolution of the survey data collection) but with considerable overlaps. The results give an insight into the way experts approach simulation projects and their problem solving strategies. A potential application is in training novice modellers, particularly in developing 'craft skills'. The results also provide an empirical basis for further research, especially in conceptual modelling.

Keywords: simulation; practice of OR; OR education; conceptual modelling; modelling process

1. Introduction

In an article in Operations Research, Willemain (1995) wrote: '… the important issues of how analysts proceed in practice have so far not been studied and understood in a systematic way. While the OR community presses on with its traditional business of developing and applying new methods, it must also attend to metamodeling issues and the larger subject of an epistemology of practice …', and he also quoted Gass (1987): 'We need to get away from the crutch that modeling is an art.' In order to start to address these issues, Willemain's (1995) study looked at the model formulation step in the first hour of tackling an Operational Research (OR) problem, and the approach taken was the use of think aloud protocols to capture and then analyse the thoughts of experts. This produced valuable data and novel insights into what actually happens in the first part of the modelling process. For example, one result was that there was a lot of switching between the different topics, particularly between developing the model (model structure) and evaluating it (model assessment). In other words, the experts often thought about some aspects of how the problem could be modelled, then critically evaluated the ideas, and then further revised or developed the model.

We also believe that it is very important for research in OR to include the study of the process of OR and that a good way of doing this is to understand, analyse and learn from the ways of working of expert OR practitioners. However, in the many years since Willemain's paper very little further work has been done in this area and, in our view, his comments still apply. There is a need for empirical data of what happens in practice on OR projects in order to start to understand the procedures, approaches and strategies used.

An important potential application is to use this information in the education and training of students and inexperienced modellers. This should help them to develop an effective modelling style more quickly and less painfully than by just using trial and error and learning by their mistakes. Willemain and Powell (2007) suggest that learning about the behaviour and approaches of expert modellers will give novices a 'boost up the learning curve'. Indeed, Powell taught an 'art of modelling' course informed in part by Willemain's (1995) study (Powell and Willemain, 2007). Data on expert practice should also be useful for experienced OR analysts in providing some comparative benchmark data aiding reflection on their own approaches. It might also act as a foundation and stimulus for developing better methods or general principles or an improved methodology for certain steps in the modelling process.

In the light of this research gap, the aim of the work described in this paper was to obtain data on expert practice. In order to get information from a large number of experts, a survey was used. This paper reports and analyses some of the results from the survey. Our particular interest is simulation and so the survey was about experiences on simulation projects. Aspects covered by the survey included the time spent on different tasks, the approaches used for certain tasks especially conceptual modelling, the patterns of the task timeline, and the interactions between the tasks. The emphasis on conceptual modelling is because this step is concerned with deciding what to include in the model and because it has received relatively little attention in research or textbooks. As already indicated in the quote from Gass, this task is often considered to be an art in contrast to the more scientific approach applied to the other simulation steps. Looking at conceptual modelling is also similar to the focus in Willemain's (1995) study of looking at model formulation.

The next section discusses previous literature. Then the questionnaire design and administration is explained with the survey results presented in the remainder of the paper.

2. Literature review

2.1. Willemain (1995) and related literature

Willemain's (1995) think aloud protocol experiment involved 12 OR experts. Four artificial but realistic OR problems were used, each of a slightly different type. Four experts did all four problems with the other eight doing one each. There were therefore 24 exercises, and in each one the expert spoke aloud their thoughts for an hour as they started to tackle the problem. The sessions were recorded and a transcript produced. The transcripts were then coded by segments of text (down to part sentence if appropriate) according to which one of six 'topics' (problem context, model structure, model realisation, model assessment, model implementation, other) the expert was considered to be addressing. Willemain uses the term topic although in various other literature tasks, phases, stages, or steps are used. These terms are considered to be synonymous in this context and are used inter-changeably in this paper. The topics were chosen by Willemain and are explained in detail in his paper but, briefly, context is about understanding and formulating the problem, structure is developing the model itself including initial analysis of data, realisation is calculations or methods to obtain parameter values for the model or results from the model, assessment is evaluations of the model, implementation is how the client would use the model in the problem.

The codings were analysed in various ways such as the percentage of transcript lines categorised as each topic, the timeline of the topics, and the transitions between topics. As already mentioned there was a lot of switching between topics, and these patterns are a particularly interesting result. Willemain describes this as 'The experts … conduct thought experiments involving the other stages as an integral part of the formulation stage.' The six most common transitions were between model structure and one of model assessment, problem context or model realization, with the transitions to and from model assessment having the two highest frequencies (together 37% of the total).

The paper tabulates the 1451 transitions from the 24 exercises. Although not calculated in the paper, this means that there was an average of 60.5 transitions per exercise, and so they occurred on average about one minute apart. In other words the mean time spent on one topic before switching to another topic was about one minute (including the end of the exercise as the end of the last topic and assuming an exact duration of 60 minutes, the precise calculation is 60/(60.5+1) minutes).

Some of the other findings were that although model structure was the main focus (as would be expected in a model formulation exercise), considerable time (about 40% of the transcript lines) was spent on the other five topics, the median positions of topics were in the expected order, and there was some indication of different modelling styles amongst the experts.

Willemain also carried out a survey with the same 12 experts (Willemain, 1994) consisting mainly of 37 Likert self-description questions (all 12 responded) and questions asking them to list what they considered to be the important qualities of an effective modeller, model, and modelling process, and desirable qualities in a modelling client (11 respondents). A notable result was that for the effective modeller qualities only a fairly small proportion of responses were about technical skills and knowledge. Instead, most of the responses were categorised as either mind-set (43%) with several mentions of creativity and related qualities, or nontechnical expertise (26%) such as communication, experience, team working, and time management. This emphasises the importance of what Willemain terms ‘craft skills’ in modelling, which, for example, are needed for problem structuring and conceptual modelling.

The Willemain (1995) study was analysed further and extended in three papers published in the Journal of the Operational Research Society. Waisel et al. (2008) analysed the sketches made by the four experts who did all four problems (i.e., 16 exercises in total) relating them to the parts of the transcript when they were being drawn. The modellers used sketches more often while working on the model (topics of structure and realisation) than while understanding the problem or evaluating the model (topics of context and assessment). The use of sketches tended to start earlier in the harder and more complex problems.

Willemain and Powell extended the work by carrying out a similar experiment using novice modellers (Powell and Willemain, 2007; Willemain and Powell, 2007). The participants were 28 MBA students who each took part in two sessions about 10 weeks apart. Each session used think aloud protocols as the students tackled two OR problems spending 30 minutes on each problem. One problem from the first session was repeated in the second. The protocol transcripts were analysed qualitatively (Powell and Willemain, 2007) and quantitatively (Willemain and Powell, 2007).

In the qualitative analysis, Powell and Willemain identified five main ways in which the novices’ behaviour was not, in their opinion, good modelling practice, which were: ‘over-reliance on available data’, ‘taking shortcuts to an answer’, ‘insufficient use of abstract variables and relationships’, ‘ineffective self-regulation’, and ‘overuse of brainstorming relative to structured problem solving’.

The quantitative analysis was broadly similar to that in Willemain (1995) and the results were compared with those of the experts in Willemain (1995). The novices spent a much higher percentage of time than the experts on problem context and a lower percentage of time on each of the other topics. Similarly, for the transitions the most common pair for the novices was model structure and problem context whereas for the experts it was model structure and model assessment.

The paper reports 883 transitions for the 112 novice exercises (3360 minutes in total), although transitions to or from the ‘other’ topic are not included. This compares with 1195 transitions (excluding transitions involving ‘other’) for the 24 expert exercises (1440 minutes). As with Willemain (1995) the transition rate was not calculated in the paper but, in fact, this is a clear difference between the novices and the experts. Including the end of each exercise as a transition and excluding the ‘other’ topic, there were 0.296 transitions per minute for the novices compared with 0.847 for the experts. Measured in mean time between transitions this is 203 seconds for the novices and 71 seconds for the experts. The experts therefore switch between topics much more frequently than the novices.

Comparing the behaviours, the novices seem to be often thinking about the model and then going back to the problem description whereas the experts tend to think about the model and then critically evaluate the idea. The lower transition rate also implies that the novices tend to spend longer on each train of thought before changing their thinking mode.

The quantitative analysis also included carrying out textual analysis of both the novice and expert transcripts to compare keyword usage. One difference between the groups was that the novices used certain keywords about numerical calculations more often than the experts and keywords about some more abstract concepts less often. A similar qualitative finding was that the novices had a tendency to attempt to perform calculations on any data provided (sometimes being irrelevant or nonsensical) rather than thinking in more general abstract terms.

One of the aims of the papers was to investigate the effect of the ‘art of modelling’ course taught by Powell. Four of the participants took the course between the two sessions and there was some indication in both the qualitative and quantitative analysis that the course enabled the students to alter their performance to be a bit more like the experts compared with the change in performance of the other students in the experiment. However, the sample size of four was very small and the papers comment that there is therefore a lot of uncertainty and the evidence is quite weak.

Wang and Brooks (2007) compared expert and novice behaviour in the area of simulation. They obtained data on the time spent on different topics by one expert and nine novice groups whilst they carried out real simulation projects. Some data collection was by the modellers themselves and some by observation by the researcher. Differences in behaviour included more overlapping of topics on the expert project. However, it was difficult to know the extent to which this was due to the higher level of expertise as there were significant differences in the nature of the expert and novice projects. A novice expert comparison for real projects is valuable potentially but there are great practical difficulties in setting this up and collecting detailed data for the same or similar real projects.

2.2. Conceptual modelling

Conceptual modelling was given an emphasis in our survey due to the lack of guidance on how to approach this task. Other simulation tasks such as data analysis, model building, verification and validation, and output analysis have well established methods based on mathematics, computing, statistics or logic. There is plenty of advice for these tasks (e.g., Law 2007) with standard methods for many typical situations. Specialist software is available to assist some of these tasks, such as distribution fitting for data analysis. The current state of conceptual modelling methodology is quite different with most textbooks devoting only a few pages to the topic and providing only a few general guidelines. Conceptual modelling is however recognised as a challenge with, for example, Law (1991) stating '… after having been involved in numerous large-scale simulation projects, I now strongly feel that the most difficult aspect of a study is that of determining the appropriate level of model detail'.

The term conceptual modelling itself can cause confusion because of its different uses in different areas of science and also because there is no agreed definition in the simulation literature (Robinson, 2008a). This paper follows the definition of Brooks and Robinson (2001) that a conceptual model is 'a software independent description of the model that is to be constructed'. Conceptual modelling consists of creating the design of the virtual world of the simulation model, typically determining the entities that it contains and all the interactions, rules, and equations that determine their behaviour. This is a different task to actually building the model, which comprises implementing the conceptual model usually by building and coding it in a simulation software package. It is possible for conceptual modelling and model building to take place together, but it is still important to consider them as separate tasks since their nature and objectives are quite different. Conceptual modelling is the choices as to which model to build and use in the project and therefore affects all the subsequent tasks. Consequently, it is a really important part of simulation and one that should have more research work devoted to it.

Much of the general advice on conceptual modelling in simulation and in OR has emphasised the advantages in keeping the model as simple as possible (e.g., Ward, 1989; Salt, 1993; Chwif et al., 2000). Typical advantages suggested are that simpler models are easier to build, change, maintain, understand, and experiment with, they require less data, and they enable the complete list of their assumptions to be identified more easily. As a result they are more likely to be accepted by the client and to lead to a successful outcome for the modelling project

(Tilanus, 1985). However, if a model is too simple and omits important factors then it will probably not have sufficient validity. In these discussions in the literature the meaning of the complexity or the level of detail of a model is not defined clearly which means that these statements are somewhat imprecise. Model complexity is really quite a broad concept that can be broken down into several different factors, and Brooks and Tobias (1996) suggested three such factors termed ‘size’ (number of elements or components), ‘connectedness’ (number of connections between the elements) and ‘calculational complexity’ (complexity of other calculations such as determining the routes). A small scale experiment with Masters students indicated that understanding may be affected mainly by size and connectedness, with model building time depending more on calculational complexity (Brooks, 2011).

Recently, there have been efforts to increase the interest in conceptual modelling research in simulation. Robinson (2006) proposed a research agenda for conceptual modelling of: definition, requirements, guidelines for model development (e.g., principles, frameworks), representation and communication (e.g., documentation), validation, and teaching. In Robinson (2008b) he set out a framework for conceptual modelling and provided an example of how it applied to a Ford Motor Company project. There has been a special issue on conceptual modelling in the Journal of Simulation (volume 1 issue 3, 2007), one day conceptual modelling workshops following the last four biennial UK Operational Research Society Simulation Workshops, and a book with contributions from 29 authors covering recent work in various aspects of conceptual modelling (Robinson et al., 2011). However, what is lacking is empirical data as there are very few studies that provide any data that is relevant for these research issues. An important starting point should be to understand what currently happens on simulation projects, and our survey aims to contribute to this.

3. Survey methodology

The questionnaire was composed of 24 questions, and had a mixture of multiple choice, Likert-scale, and open-ended questions. One of the main objectives was to obtain data about specific real projects, as it was considered that this would give more precise and richer data than general questions on the overall modelling approach preferred. Therefore, most of the questions asked about the last modelling project that the respondent had completed, with the project background, project outcome, conceptual modelling, model coding and modelling process being addressed. The unit of analysis is therefore the simulation project. Each project is done by a different respondent and so any differences will be due to a combination of project characteristics and the respondent’s modelling approach.

Some of the questions were based on previous literature with some initial questions being taken from Willemain (1994) and a question on the timeline of the projects inspired by Willemain's (1995) data. This paper focuses on the questions about conceptual modelling and the modelling process. Space limitations mean that results for some questions particularly reasons for the success or failure of the project and the most difficult task are not included. The main research questions that the part of the survey included in this paper aimed to address were:

- What are the objectives of simulation projects?
- When does conceptual modelling take place in comparison to the other topics?
- How is the conceptual model developed and documented?
- When and what kind of changes are made to the initial conceptual model and why?
- What is the relative time spent on the different tasks in a simulation project?
- What is the pattern of the time allocation to the tasks during the project?
- What are the interactions between tasks in the project timeline?

Given the confusion that seems to exist as to what is meant by conceptual modelling the questionnaire included at the start the following explanation of our meaning: 'We are particularly interested in conceptual modelling, by which we mean all the decisions about which parts of the real system to include in the model: for example, which elements and factors to include and exclude, the relationships between the elements, the rules of behaviour of the elements and the modelling assumptions and simplifications. The conceptual model is a specification of the model that, in principle, is software independent. We distinguish conceptual modelling from model coding, with model coding being the process of writing the code to implement the conceptual model in the particular software. We also distinguish conceptual modelling from problem structuring, which is understanding the real system and the problem.'

The inclusion of this statement could bring an element of bias to the responses in leading to the respondents thinking about and emphasising conceptual modelling more. This was considered preferable to the alternatives of not including a definition, in which case the results would be difficult to interpret as it would not be known what meaning the respondents were using, or asking the respondents for their own definition, which would also have emphasised conceptual modelling and could have produced a wide range of responses limiting the analysis and interpretation. Our subjective observations were that the respondents completed the questionnaire seriously and conscientiously.

The questionnaire was administrated in a period of one and half months from 25th March 2006 to 15th May 2006. Questionnaires were handed out to attendees at the 2006 UK Operational Research Society Simulation Workshop with 30 responses being received from 60 contacts, and at the 2006 WITNESS User Conference with 10 responses received from 68 contacts. Authors from the proceedings of 2005 Winter Simulation Conference were emailed and asked to fill in the questionnaire online, generating 62 responses from about 600 contacts. Overall this gave a total of 102 responses. From observations at the conferences, it took about 20 minutes to complete the questionnaire. Despite being a long questionnaire the respondents rarely missed or skipped a question unless it was not relevant. There appeared to be considerable interest in this topic with 70% of respondents wanting to receive a summary of the results and 43% being willing to be contacted again for further discussion, as well as enthusiasm being expressed in informal discussions.

The groups targeted were chosen as a convenient and efficient way of getting a sizable number of responses from experienced simulation users. The OR Society Simulation Workshop was held over two days and provided good opportunities to approach the attendees personally, resulting in an excellent response rate. All three conferences are specialist simulation events, with the Winter Simulation conference being the world's largest simulation conference, and so most of the questionnaire respondents will be specialists in simulation. Overall, a good proportion of world's simulation experts should therefore have been contacted. This type of approach of convenience sampling and voluntary response sampling through targeting suitable groups, rather than random sampling from the entire population, has been the usual approach in simulation research surveys (e.g., Cochran et al., 1995; Bell et al., 1999; Melao and Pidd, 2003) and other OR surveys (e.g., O'Brien, 2011). However, as a consequence we cannot claim that the results are entirely representative of the population of experienced simulation modellers and, indeed, it would be very difficult to define or list such a population. The good number of responses from the OR Society Simulation Workshop means that there will be some bias towards UK and Europe. Although the WITNESS User Conference is mainly focussed on simulation users in industry, overall, academia is likely to be over represented. Nevertheless, the results should give a valuable insight into the working methods of simulation specialists.

4. Results for respondents' background and modelling style

The initial part of the questionnaire asked about the respondent's background and asked some Likert scale questions about their general modelling style. All 102 respondents answered the question on the type of organisation they worked for. Exactly half (51) worked at a university, with the number of responses for the other types of organisation being: consultancy: 19; manufacturing: 14; government: 11; service sector: 8; R&D / research institute: 3; oil and gas / energy: 3; defence and aerospace: 2; distribution: 1; software: 1; network operator: 1; (nine respondents listed more than one organisation type, with 114 responses in total). The bias towards the academic sector reflects the sources used for obtaining contacts, but there is still good representation from industry.

The respondents were asked to write the number of years of modelling experience that they had and the answers ranged from one year to forty two years with an average of twelve. We summarised the responses into intervals as follows: 1 – 5 years: 39; 6 – 10 years: 23; 11 – 15 years: 10; 16 – 20 years: 13, 21 – 30 years: 11; more than 30: 6. In the least experienced category of 1 – 5 years the actual answers were: 1: 6; 1.5: 1; 2: 13; 3: 8; 4: 5; 4.5: 1; 5: 5. The respondents were also asked to select from a list of six categories (which in error included some overlaps) the number of simulation modelling projects they had carried out and the results were: 1 – 5: 32; 6 – 10: 12; 10 – 20: 21; 20 – 30: 7; 30 – 50: 10; more than 50: 20. As would be expected, there is a relationship between the number of projects and experience, with the average experience increasing for each category of number of projects from 3.7 years for 1 – 5 projects up to 23.1 years for more than 50 projects. Overall, most respondents have considerable experience and so we can be confident that the group of respondents include a high proportion of simulation experts.

Whilst the main focus of the questionnaire was on the most recent project, five initial questions were asked on the respondent's general modelling style. Willemain (1994) used a seven-point Likert scale format to ask 12 experts a total of 37 'self-description' questions covering various characteristics of themselves, their models, the problems they model and the way they model. Five of these questions from the latter category that we considered the most relevant for our survey were used with the respondents asked: 'To describe the way you model, please locate yourself on the 7-point scale for the following items'. The results are shown in Table 1. Simply looking at the total proportions giving some preference, the strongest results are for 'start small and add' (24 + 44 + 13 = 81 responses out of 101 = 80%), and 'always draw / doodle' (74%), which were also amongst the strongest results in Willemain (1994). Starting with a small model and adding detail is sometimes suggested in the literature as a good modelling approach (e.g., Pidd, 1999), whilst 'always draw / doodle' highlights the creative element in modelling. For the other three questions, a majority of respondents selected 'systematic process' (66%) and 'look at data first' (61%), whilst there is a fairly even split between 'make a single model' and 'make alternative models'. The indication therefore is that the modellers have some standard procedures or methods and tend to look at data fairly early in the process. Willemain's much smaller sample of 12 experts also had a small majority for 'look at data first', but there was a fairly even split between 'systematic process' and 'unique, ad hoc process', and a majority for ‘make alternative models’. Interestingly, for each question there are some selections for every value in our responses indicating a wide range of modelling styles.

Table 1 Modelling style (101 respondents for each question)

Start small and add	24	44	13	9	3	6	2	Start big and subtract
Never draw/doodle	3	6	6	11	18	28	29	Always draw/doodle
Make single model	6	19	19	19	13	16	9	Make alternative models
Look at data first	21	26	15	19	10	8	2	Look at data last
Systematic process	15	24	28	15	6	9	4	Unique, ad hoc process

5. Results for the most recent project

The respondents were asked to answer the remainder of the questions for their ‘most recent completed simulation project’.

5.1. Project background, objectives and success

The first question about the project asked the respondents to select the problem area from a supplied list (with an ‘other’ option where they were asked to specify the answer). A few selected more than one area and there were 115 responses, as follows: logistics, transportation and distribution: 30; manufacturing: 25; military: 18; service sector: 10; health care: 10; business process modelling: 7; call centre: 4; other: 11 (one each for public sector, hydro, construction, delinquency, project scheduling, planning, computer system, network, government, homeland security, not specified).

The respondents were then asked to select the project objectives from a list provided and number them in order of importance if there was more than one, and the results are shown in Table 2. There was one no response and 10 respondents selected more than one response but did not specify the order of their responses (5 gave 2 responses, 3 gave 3 responses, and 1 each gave 4 and 6 responses). Most of the projects had multiple objectives (78 of those who gave an order plus 10 who did not specify an order, i.e., 88 / 101 = 87%), with 30% having at least 5 objectives. More than half of the respondents selected ‘improve the understanding of the system’ as an objective, and as well as being the objective with the highest overall frequency it was also the most common main objective (i.e., importance 1). This is also reflected in the simulation literature, which often emphasises the importance of using simulation for gaining insight and understanding and not just treating it as a black box model that produces a prediction. The next most common main objectives were maximise profit / minimise cost and improve throughput. The prevalence of multiple objectives is probably partly due to the list in the questionnaire including items that could be different aspects of one overall objective, such as maximising profit by improving throughput and reducing work in progress. However, it also could indicate simulation models being used for several different purposes within a project.

Six categories were provided for the length of the project, with the responses indicating a wide variation in duration: less than a week: 6; 1 – 4 weeks: 10; 2 – 3 months: 19; 4 – 6 months: 14; 7 – 12 months: 33, over a year: 20. Just over half of the projects (52%) had a duration of over 6 months and so there are a considerable number of substantial projects. Those specifying over a year were asked to state the length, and 14 responses were provided, categorised as: between 1 and 2 years: 5; 2 years: 5; 3 years: 2; 4 years: 1; 1 response unclear. Only the duration was asked for and so these responses do not necessarily equate to the total person time for the project. With hindsight, asking the respondents to state the total person time may have been more useful, although perhaps harder for the respondents to estimate.

Table 2 Project objectives (101 respondents)

	Importance											*No*	*Total*	*% of*
Project objective	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*Order**		*respondents*
Improve understanding of the system	19	15	10	4	4	3	1	1	0	1	0	3	61	60%
Improve throughput	13	10	5	5	5	0	1	0	0	0	0	5	44	44%
Design a new system	10	11	7	3	4	1	1	0	0	0	0	4	41	41%
Maximise profit / minimise cost	17	7	6	3	1	1	1	0	0	0	1	3	40	40%
Reduce queueing time	8	3	8	5	2	1	0	4	0	1	0	3	35	35%
Reduce resources required	2	11	5	4	1	4	3	0	0	0	0	3	33	33%
Develop a long term plan	4	9	5	2	5	0	0	0	1	1	0	4	31	31%
Improve schedule	2	3	5	7	2	1	2	0	2	0	0	2	26	26%
Improve facility layout	3	3	3	7	1	2	1	0	0	0	0	2	22	22%
Reduce work in progress	3	3	3	3	1	2	0	1	0	0	0	0	16	16%
Other: various	8	3	4	1	0	0	0	0	0	0	0	0	16	16%
Training	2	0	4	1	3	2	2	0	1	0	0	0	15	15%
Total number	91	78	65	45	29	17	12	6	4	3	1	29*	380	
% of respondents	90%	77%	64%	45%	29%	17%	12%	6%	4%	3%	1%	10%*		

* 10 respondents (10% of respondents) did not specify an order. They gave a total of 29 responses.

The respondents were asked to rate the success of the project on a 7 point integer scale from 1 (successful) to 7 (unsuccessful). All but one of the respondents provided an answer and Table 3 shows the 101 responses. The project was given the highest success rating by 20% of the respondents to this question, and most (93%) selected 1, 2 or 3 indicating that the project was successful to some extent, at least in the opinion of the respondents. Although this is only the views of the modellers it does lend additional support to the survey results representing good practice.

Table 3 Evaluation of project outcome (101 respondents)

Successful		→			*Unsuccessful*	
1	*2*	*3*	*4*	*5*	*6*	*7*
20 (20%)	57 (56%)	17 (17%)	2 (2%)	4 (4%)	1 (1%)	0 (0%)

5.2. Conceptual model: Development of the conceptual model

Generating the conceptual model requires obtaining an understanding of the real system (problem structuring) and then deciding which aspects of the real system to include in the model and how to represent them (conceptual modelling). The survey questions asked about how the respondents gained their understanding of the real system, the methods used for developing the conceptual model, the relationships of the conceptual model with data and coding, and how the conceptual model was documented.

The respondents were asked to choose from a provided list the methods used to understand the real system and the problem and number them in order of importance if there were more than one. The results are shown in Table 4. There were 101 respondents who answered this question, although a few gave more than one answer with the same level of importance and eight listed more than one method without specifying the order of importance. Just 10 respondents selected one method only and so on most projects several methods were used. From the responses, ‘analyse system data’ (74%) is the most widely used method, with ‘talk to the management’ a close second (69%), although ‘talk to the management’ is the one with the highest frequency as the most important method. It was expected that these would be used widely since understanding the problem requires understanding both the system and the viewpoint and objectives of the client. It is perhaps more surprising that in 31% of projects the modeller didn’t talk to the management as part of understanding the system and the problem. Perhaps this could be because the project was actually commissioned by someone at a different level in the organisation. Observing the system was used by 53% but this means that in nearly half the projects the respondent did not observe the system directly. It is not clear whether this is because it was considered unnecessary or because it was not feasible (e.g., the system didn’t exist). Talking to system operators also occurs often and each of the top four methods in Table 4 were used by a majority of respondents. There is some usage of problem structuring methods indicating combining soft and hard OR approaches. The ‘other’ category includes four who used documents or literature, as well as a variety of other methods.

Table 4 Methods used for understanding the real system and the problem (101 respondents)

Method for understanding the problem	*Importance*							*No Order*	*Total*	*% of respondents*
	1	*2*	*3*	*4*	*5*	*6*	*7*			
Analyse system data	16	23	17	8	2	2	0	7	75	74%
Talk to the management	28	14	11	5	3	2	0	7	70	69%
Observe the system	16	15	14	4	2	0	0	3	54	53%
Talk to system operators / servers	14	14	8	9	0	2	0	5	52	51%
Talk to customers	11	5	10	3	3	1	0	2	35	35%
Problem structuring method	5	11	6	2	5	0	1	0	30	30%
Other	10	6	1	2	0	0	0	0	19	19%
Total	100	88	67	33	15	7	1	24	335	

There are 127 different possible combinations of seven methods (choosing between one and seven methods ignoring the order) and none of the specific combinations has a large frequency. A more useful analysis is therefore to calculate the frequency of each pair and each group of three and four methods either on their own or with other methods (again ignoring the order). Pairs used by least 30% of respondents are analyse system data and talk to management (56%), analyse system data and observe the system (46%), analyse system data and talk to system operators (46%), talk to management and observe the system (40%), talk to management and talk to system operators (40%), observe the system and talk to system operators (33%). Groups of three used by at least 30% are analyse system data and talk to management and observe the system (37%), analyse system data and talk to management and talk to system operators (36%), analyse system data and observe the system and talk to system operators (31%). The most common group of four is analyse system data and talk to management and observe the system and talk to system operators with 26%. Not surprisingly these groups are all combinations of the first four methods in Table 4 with the ordering within each group of two, three and four mainly following the sum of the frequencies in Table 4. Considering just these four methods, 93% of respondents used at least one, 77% used at least two, 52%, used at least three, and as already noted 26% used all four.

The respondents were asked to choose which methods, if any, they used for developing the initial conceptual model and number them in order of importance if there was more than one. Four methods were provided including ‘any other formal method, please specify’, and in addition there was a box to indicate that none of the methods were used. All 102 participants answered this question, although 5 respondents listed more than one method without specifying the order of importance, and two respondents each gave two answers with importance 1. There were 6 respondents who selected the answer that they didn’t use any of the methods (i.e., they didn’t use a formal method). The responses are summarised in Table 5. This shows that the most common method is to use previous experience of a similar problem (71%), emphasising the importance of experience in conceptual modelling. About two thirds used some form of preliminary analysis of the system, and so this is also a very common technique. This is probably something that novices rarely do and find quite difficult. Many respondents used brainstorming, indicating the creative nature of conceptual modelling. As already noted, in Willemain's (1994) survey there were several mentions of creativity as an important quality of an effective modeller. For 'any other formal method' the respondents were asked to state the method and the 11 responses were (one respondent gave two methods and one gave none): UML: 2; process mapping: 2; systems and software theory: 1; conical methodology specification: 1; data model: 1; interviewing: 1; internal assembly line process: 1; SAD: 1; freedom for creativity: 1.

Table 5 Methods used for developing the initial conceptual model (102 respondents)

Conceptual modelling method	*Importance*				*No order*	*Total*	*% of respondents*
	1	*2*	*3*	*4*			
Previous experience of modelling a similar system or problem	39	20	8	0	5	72	71%
Preliminary analysis of the system (e.g. simple analytical model)	36	21	4	1	3	65	64%
Brainstorming session	14	16	11	0	3	44	43%
Any other formal method	4	4	3	0	0	11	11%
None of the above methods	6				0	6	6%
Total	99	61	26	1	11	198	

Considering combinations, for pairs of the main three methods on their own or with other methods the percentages are: previous experience and preliminary analysis: 47%, previous experience and brainstorming: 33%, preliminary analysis and brainstorming: 31%. The percentage using all three methods is 25%. Overall, 66 respondents (65%) used more than one method. The results highlight the difficulty of conceptual modelling. There isn't a standard formal procedure to follow and the methods that are used are hard generally requiring previous experience, analysis or creativity. Based on the survey results, most modellers use some combination of the methods.

Respondents were asked when they developed the conceptual model in relation to model coding with the results of 'before model coding': 62 responses (61%); 'started before model coding, and finished while coding': 33 (33%); 'entirely during model coding': 6 (6%). Prior to the survey, it was considered that the ease of use of simulation software might mean that coding and conceptual modelling often take place together in an iterative manner. However, based on these results, modellers normally develop the conceptual model before model coding. In other words, the model is fully planned before being built.

The survey asked about the impact of data on the conceptual model and the 100 responses were ‘data altered the conceptual model’: 61 (61%); ‘data had no impact on the conceptual model’: 28 (28%); ‘conceptual model was not developed until after data collection’: 11 (11%). There was one no response and one respondent didn’t select one of the answers provided but wrote ‘the model is designed based on the available data’. Robinson (2004) suggests that ideally the conceptual model should be developed without being constrained by whether the data is available, with the choice of model driving data collection rather than the other way round. Problems with obtaining data can sometimes be handled by changing the conceptual model or by methods for dealing with unavailable data such as treating it as an experimental variable. These responses provide support for this approach with very few respondents completing data collection before developing the conceptual model. However, in many cases data then causes changes in the conceptual model. A result from the survey not included in this paper is that data collection and analysis is often the hardest task and so it seems that difficulties with data collection are common and these can cause the model to change.

The respondents were asked to select the methods used for documenting the conceptual model from a provided list and all 102 gave a response, with the results given in Table 6. Many respondents (66%) used more than one method with the process flow diagram (PFD) and a list of assumptions and simplifications (LAS) being the two most popular. Most projects (81%) used at least one of PFD and LAS, with 38% using both methods. The logic diagram was combined with PFD in 27 out of the 32 cases in which it was used. Other pairs of methods (on their own or with other methods) with at least 15% of respondents are: LAS and logic diagram 20%, PFD and component list 18%, LAS and component list 17%. The only group of three methods with a frequency of more than 15% is PFD and LAS and logic diagram with 17%. The ‘other’ responses included four answers that indicated the use of a text description or Word document.

PFD is probably the representation that is most meaningful to the customer as it can be related most directly to the real system. However, the more simulation-based representations such as the activity cycle diagram and UML still have some usage. The importance of identifying explicitly the assumptions and simplifications made during the modelling process is borne out by the widespread use of LAS. Most projects had some form of documentation of the conceptual model with only four respondents selecting ‘none’.

Table 6 Documentation methods for the conceptual model (102 respondents)

Documentation method	*Total*	*% of respondents*
Process flow diagram	64	63%
List of assumptions and simplifications	58	57%
Logic diagram	32	31%
Component list	22	22%
Activity cycle diagram	19	19%
UML (the unified modelling language)	14	14%
Other	15	15%
None	4	4%
Total	228	

5.3. Conceptual model: changes in the conceptual model

The survey asked two questions about changes to the conceptual model. The respondents were asked to choose one of three statements about the development of the conceptual model and the results from 101 responses were: ‘developed one conceptual model with no further changes during the project’ 2 (2%), ‘developed one conceptual model with some changes during the project’ 84 (83%), ‘developed several alternative conceptual models’ 15 (15%). For the latter option respondents were asked to state the number of models with answers of between 2 and 8 conceptual models with an average of 3.5. Overall, this indicates that in almost all projects the conceptual model does change during the project.

In a detailed question, the respondents were asked to select the type of conceptual model changes made during the four subsequent modelling ‘stages’ of data collection (DC), model coding (MC), verification and validation (VV), experimentation (EX), and the reasons for the changes. There was also a box to tick for each stage if no change was made. The format of the question means that the responses show whether each type of change and reason occurred in the project during these stages. It could be that each type of change or reason consists of several separate items.

The total number of responses for each option are shown in Table 7. The results include a few responses where changes were selected but not a reason (6, 6, 7, 5 respectively for the four stages) and two where a reason was selected but not a change (1 each for MC and VV). However, these are a very small proportion of the overall responses.

Table 7 shows the total responses and so includes many respondents who selected more than one change for the stages. A further analysis was done to examine the results on a project basis to look at the type of conceptual model changes made. This splits the changes into those making the conceptual model more complex (addition of entities, more complex for inter-relationships and logic) or simpler (deletion of entities, simpler for inter-relationships and logic).

For each respondent (i.e., each project) the changes were categorised as to whether they were all making the conceptual model more complex, all making the conceptual model simpler, or a mixture. The results are shown in Table 8. The difference in number of respondents from Table 7 is due to the two who selected a reason but not a change (as mentioned in the previous paragraph).

This question gives very good data on the development of the conceptual model during simulation projects through the analysis of results in Tables 7 and 8. An initial finding is the percentage of projects in which a change was made to the conceptual model. From Table 8 the percentages are 66%, 69%, 54% and 44% respectively for the four stages DC, MC, VV, EX. The percentage for DC is similar to that for the data and conceptual modelling question in the section 5.2 (61%) indicating good consistency in responses. The results show that conceptual model change is very common at each of the stages. There were only 8 projects (9%) in which there were no changes in all four stages (out of 88 respondents who answered for all four stages). As with the question discussed in the first paragraph of this section, this implies that it is very unusual for a conceptual model to remain the same throughout the project. Most textbooks on simulation or OR comment on the iterative nature of the project process, and this evidence of expert practice emphasises the importance of being prepared to revise the model throughout the project. The main difference between the stages is the lower proportion of projects with changes at the VV and EX stages. It was anticipated that there would be far fewer changes in the later stages of the project and it is actually the large proportion of projects with changes at the VV and EX stages that we found surprising. The statistical significance of the differences at each stage was tested using a chi-square test. This was carried out on the values for no change and any change for the four stages (i.e., 33 and 63 for DC, 29 and 66 for MC, etc.). The chi-square p value is 0.002 and so the differences are significant at the 1% level. Therefore there is good evidence of the percentages varying across the stages.

For those projects where there is a change in the conceptual model the proportions with all more complex, all simpler, or a mixture are very similar for the four stages (e.g., all complex for DC is 40/63 = 63%). A chi-square test on just the change values confirmed a lack of evidence for a difference in this aspect between the stages (p value of 0.82). The percentages for the total of the values for the four stages are: all more complex 68%, all simpler 14%, mixture 19%. Therefore, there is a strong tendency for the conceptual model to become more complex as it develops, which is consistent with much modelling advice (e.g., 'be parsimonious, start small and add' Pidd, 1999). For each stage, when some of the changes simplified the model this is more often in conjunction with other changes adding complexity (mixture) rather than all the changes making the model simpler.

Table 7 Conceptual model changes during the project (DC= Data collection, MC= model coding, VV= verification and validation, EX= experimentation).

Modelling stage	*No. of respondents*	*No change*	*Conceptual model change*							*Reason for change*				
			------- *Entities* -------		*Inter-relationship*		------ *Logic* ------							
			Addition	*Deletion*	*More complex*	*Simpler*	*More complex*	*Simpler*	*Total*	*Client's requirement*	*Problem situation changed*	*Model not realistic*	*Better information about real system*	*Total*
DC	96	33	30	7	30	17	39	13	136	9	8	17	38	72
MC	96	29	35	12	37	14	38	13	149	16	11	19	35	81
VV	95	43	18	3	25	10	27	12	95	11	4	18	22	55
EX	88	49	16	4	18	5	28	3	74	9	9	8	19	45
Total			99	26	110	46	132	41	454	45	32	62	114	253
% of changes			22%	6%	24%	10%	29%	9%	100%	18%	13%	25%	45%	100%

Table 8 Overall conceptual model changes for each respondent.

Modelling stage	*No. of respondents*	*No change*	*All changes more complex*	*All changes simpler*	*Some changes simpler, some more complex*
Data collection	96	33 (34%)	40 (42%)	11 (11%)	12 (13%)
Model coding	95	29 (31%)	43 (45%)	9 (9%)	14 (15%)
Verification and validation	94	43 (46%)	35 (37%)	7 (7%)	9 (10%)
Experimentation	88	49 (56%)	30 (34%)	3 (3%)	6 (7%)
Total		154 (41%)	148 (40%)	30 (8%)	41 (11%)

Looking at Table 7, there are similar proportions of occurrences of the different types of changes (entities, inter-relationships and logic) at each stage, with about 75% of the responses being making the model more complex. The respondents often selected more than one type of change for the particular stage but sometimes only gave one reason, which is why the total number of changes (454) exceeds the total number of reasons (253). On average, therefore, each type of reason produces 1.79 types of change. The most popular reason for making a change is better information about the real system (45%) although the other reasons also occur reasonably frequently. There is no significant evidence of a difference in the proportions of each reason between the stages (chi-square *p* value of 0.46).

The data was also analysed to see if there is a difference in the type of changes for the four different reasons. Table 9 shows the total number of changes for each reason. This contains some duplicate entries in that some changes are in more than one category where the respondent selected several reasons. The proportion of each type of change is similar for the different reasons and there is no significant evidence of a difference (chi-square *p* value of 0.51). Therefore it appears that the type of reason has little effect on the types of changes made.

Table 9 Conceptual model changes for each type of reason (refer to Table 7 for the full text for the reasons).

Reason	*No. of reason responses*	*Conceptual model change*						
		------- Entities -------		*Inter-relationship*		*------ Logic ------*		
		Addition	*Deletion*	*More complex*	*Simpler*	*More complex*	*Simpler*	*Total*
Client	45	26	3	32	7	32	6	106
Problem	32	14	7	19	12	20	9	81
Model	62	25	6	31	11	41	10	124
Information	114	65	15	72	23	88	15	278
Total		130	31	154	53	181	40	589
% of changes		22%	5%	26%	9%	31%	7%	100%

In a separate question the respondents were asked what the final model looked like in terms of structure compared with the initial conceptual model and were provided with five options. The 99 responses were: more complex (significant change): 22 (22%); more complex (minor change): 44 (44%); the same: 12 (12%); simpler (minor change): 14 (14%); simpler (significant change): 7 (7%). Together with the results for the individual changes in Tables 7 and 8, this provides evidence of modellers tending to start with a simple model and adding detail as it becomes necessary during the project. This is also consistent with the respondents' assessments of their modelling style in Table 1 with most preferring 'start small and add' (80%) to 'start big and subtract' (11%).

Whether the project experienced a significant conceptual model change or a minor change / no change was compared against the time that the conceptual model was developed in relation to model coding. A considerably lower proportion of those who completed the conceptual model before model coding made a significant change to the conceptual model (12/59 = 20%) than those developed it partly or entirely during model coding (17/39 = 44%). A chi-square test gives a p value of 0.014 showing that there is quite strong statistical evidence of a difference. A possible explanation is that fully developing the conceptual model before embarking on model coding is an indicator of a well-planned project that has good information about the problem and the system at an early stage, resulting in a better initial conceptual model with less likelihood of later changes.

5.4. Whether the best model was chosen

The survey asked if, with benefit of hindsight, the best model was chosen for the project and the 99 responses were: yes: 87 (88%); no, too simple: 6 (6%); no, too complex: 6 (6%). Therefore a very high proportion of respondents felt that the best model was used. Respondents were asked to state how it could have been improved if they answered 'no'. There were comments from four of the respondents who said the model was too simple which referred to better data analysis and validation, more parts and a larger dataset, using OR algorithms, and building the model from scratch or using a different base model. All six respondents who said the model was too complex provided comments about how to improve it which referred to simplifying the model, choosing a better simulation tool with higher speed, the model having unnecessary features, a complex system being added that provided little insight, limited access to stakeholders causing difficulties in learning about the system and building the model, and a comment that sometimes it is difficult to say what model is the best until after it has been developed.

5.5. Time allocation to the topics: percentage of time on each topic

Two questions were asked about the time spent on the different 'topics' (i.e., the stages or tasks in the modelling process). One simply asked the percentage of time spent on each topic and the other asked the respondents to draw a Gantt chart. Our own preferred list of topics was used, which is problem structuring (PS), conceptual modelling (CM), data collection and analysis (DC), model coding (MC), verification and validation (VV), experimentation (EX), and report writing / presentation of results (RW). There was also an 'other' (OT) category.

There were 92 responses for the percentage of time on each topic. The mean, sample standard deviation, maximum, and coefficient of variation for each topic are shown in Table 10 and the mean values are plotted in Figure 1. This shows that the average percentage of time is fairly similar for all the topics except for model coding which has about twice as much time as each of the other topics. In Cochran et al.'s (1995) survey the 138 respondents gave their time allocations for 10 simulation stages (from Pritsker (1986)). Matching the stages up with ours using the descriptions in Pritsker (1986), the main differences are in model coding (their stage being 'model translation') where their respondents' average time allocation was much lower at 13.9%, and experimentation (their three phases for this being 'strategic and tactical planning', 'experimentation', 'analysis of results') where their respondents average allocation was much higher at 23.2%. The time allocations for other stages are very similar to our results. There does not appear to be an obvious reason for the differences with our survey responses for the average time for model coding and experimentation, although there is over 10 years between the surveys and perhaps changes in software and the types of problems tackled may be part of the reason.

Table 10 Mean and variability statistics for the percentage of time spent on each of the modelling topics.

	PS	*CM*	*DC*	*MC*	*VV*	*EX*	*RW*
Mean	11.2%	11.7%	15.9%	26.7%	13.1%	12.7%	8.0%
Standard deviation	7.8%	7.4%	10.2%	15.9%	8.3%	10.0%	4.7%
Maximum	50.0%	40.0%	50.0%	80.0%	40.0%	50.0%	25.0%
Coefficient of variation	0.70	0.63	0.64	0.60	0.64	0.79	0.58

The responses were analysed to produce a distribution of the percentage allocations. Almost all of the values entered by the respondents were a multiple of 5%, with only 29 out of the 736 values not being such a value. Also, the frequency of each multiple of 10% is much larger than the frequency for the following multiple of 5%, with the exception of the pair 0% and 5%. In view of this pattern, the data was analysed by dividing the allocation values into the intervals 0% - 7.5%, 7.5% - 17.5%, 17.5% - 27.5%, etc. so that each interval contains a pair of values of multiple 5%. For each of the topics, the average value of the responses was calculated for each interval. The relative frequencies for the intervals were then plotted against these average values. Intervals with zero frequency use the 10% value as the x-axis point (generally being the most common point in the intervals). The result is the relative frequency polygon shown in Figure 2. Using the averages for the x-axis values rather than the mid-points of the intervals gives a slightly better representation of the data and also results in data points that produce the correct overall mean for each topic.

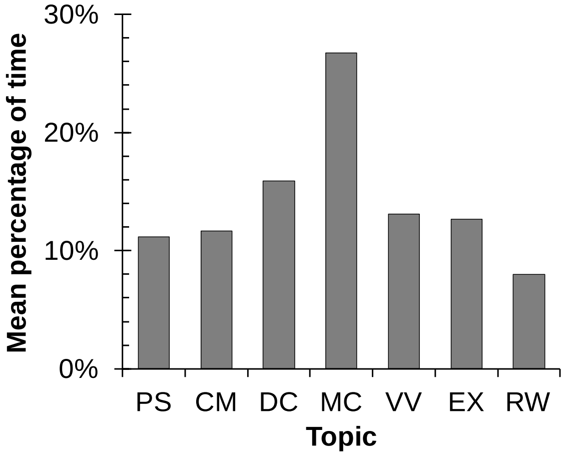

Figure 1 Mean percentage of time allocated to the topics.

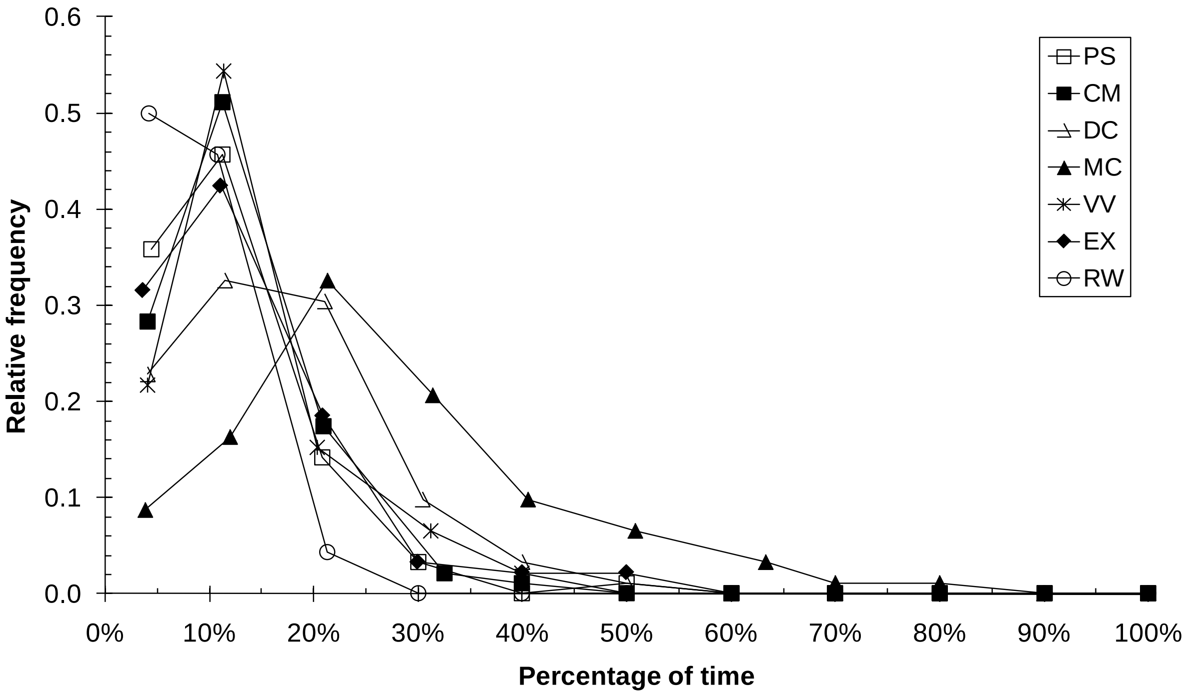

Figure 2 Relative frequency polygon of the time spent on each of the modelling topics.

The distributions for PS, CM, VV and EX are very similar with each having most of their values in the first two points plotted in Figure 2 (i.e., most values 0%, 5%, 10% or 15%). The distribution for DC differs slightly, with the third value having a higher relative frequency of 0.30 (i.e., in about 30% of projects either 20% or 25% of the time was reported to be data collection). Report writing (RW) has more values in the first interval and fewer high values (i.e., often only about 5% of time is used for this). The shape of the distribution for MC is similar to PS, CM, VV and EX, but the scale is different with the x-axis values being rescaled by a factor of about two. Therefore its mean and standard deviation values are about twice the size. Further evidence of the similarity in the shapes of the distributions is the similarity in the coefficient of variation values in Table 10, which are all between 0.58 and 0.79.

The 'other' category is not included in Figures 1 and 2 and Table 10 because it was used by only 6 respondents, and the descriptions and time allocations for these were 'client liaison' 5%, 'custom manual' 15%, 'talking with client' 25%, 'demonstration' 10%, 'stakeholder management' 3%, no description 15%. Hence the average percentage of 'other' across the 92 respondents is 0.8% (and hence the rounded mean percentages in Table 10 add up to 99.3%).

The relationships between the topic times were explored by calculating the Pearson correlations between each pair of topics. The difficulty is that the topic times are percentages and so this is compositional data. A negative correlation would be expected since a higher proportion of time spent on one topic leaves less time for the other topics. Therefore correlation is difficult to interpret and can be misleading (Aitchison, 1986). There are 28 pairs of topics although the correlations with 'other' (7 pairs) are inevitably all close to 0 since almost all the 'other' values are 0%. Considering the remaining 21 pairs, the correlation is negative for 17 out of the 21 pairs, and the mean of these correlation values is -0.14. There are 18 values between -0.34 and 0.06, with the other 3 correlations being MC and EX -0.53, EX and RW 0.18, PS and CM 0.21.

The correlations between MC and each of the other topics (apart from 'other') is a fairly large negative value (all -0.18 or less). This is presumably partly due to the magnitude of the MC values being larger than the other topics and so there is a greater effect of higher values of MC reducing the time available for the other topics. This probably partly explains the large negative correlation between MC and EX.

The likely reason for the positive correlation of 0.21 between PS and CM is that these are related topics with both tending to occur early in the modelling process. For example, an unusual and complex problem will tend to present more modelling options than a standard simple problem and so will require spending a relatively high proportion of time on both topics. Once the model is defined the time for the other topics may be closer to a normal project. There is also a positive correlation between EX and RW of 0.18 which may simply be because the experiments carried out and their implications typically receive the most emphasis in modelling reports.

The amount of variation in the times for different topics within each project was explored by calculating the standard deviation (population version) of the seven or eight topic percentages for each project (the 'other' category only included if it was used by the respondent). The average standard deviation is 10.2% and the histogram of standard deviations is shown in Figure 3. Many of the projects have a small standard deviation resulting from a fairly even allocation between the topics. Since the topic percentages sum to 100%, there is a very strong relationship between the standard deviation and the maximum of the topic values with the correlation between them being 0.97. The regression relationship (r^2 = 0.94) is:

standard deviation = -1.833 + 0.352 × maximum value

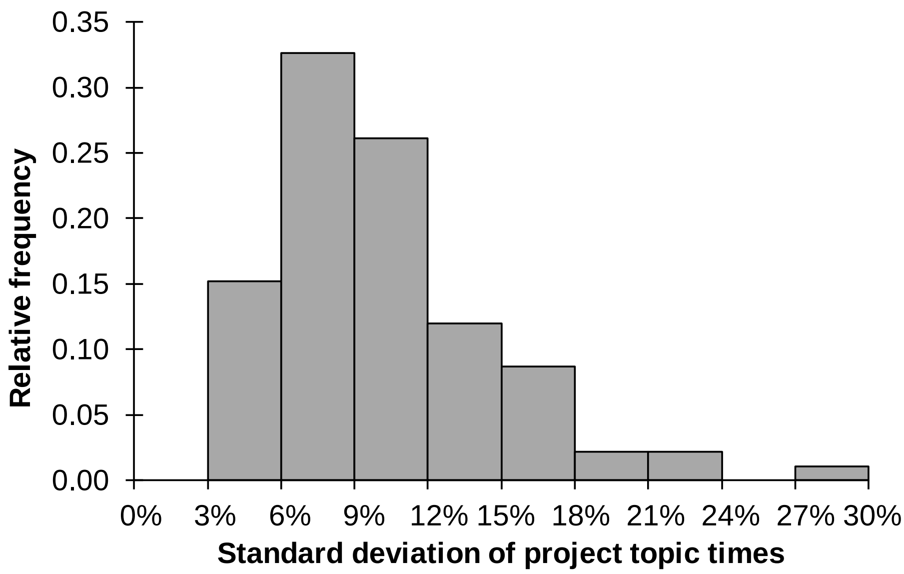

Figure 3 Histogram of the standard deviation of project topic times.

The number of times each topic is the maximum value (or one of the maximum values) over the 92 responses is (with the number in brackets giving the adjusted value by allocating 1 / n for projects where n topics are the equal maximum): PS: 8 (4.8); CM: 10 (5.6); DC: 24 (16.4); MC: 54 (44.5); VV: 14 (9.1); EX: 15 (10.8); RW: 1 (0.3); OT: 1 (0.5). MC is the maximum far more than the other topics and the correlation between MC and the standard deviation is 0.65. The maximum values are all multiples of 5 except for one value of 37% and Figure 4 shows the relative frequency of different maximum values for the same intervals as Figure 2 and using the average value within each interval as the x-axis value. The high correlation between the maximum and the standard deviation means that this chart has a similar shape to Figure 3 with most projects having a maximum of 35% or less.

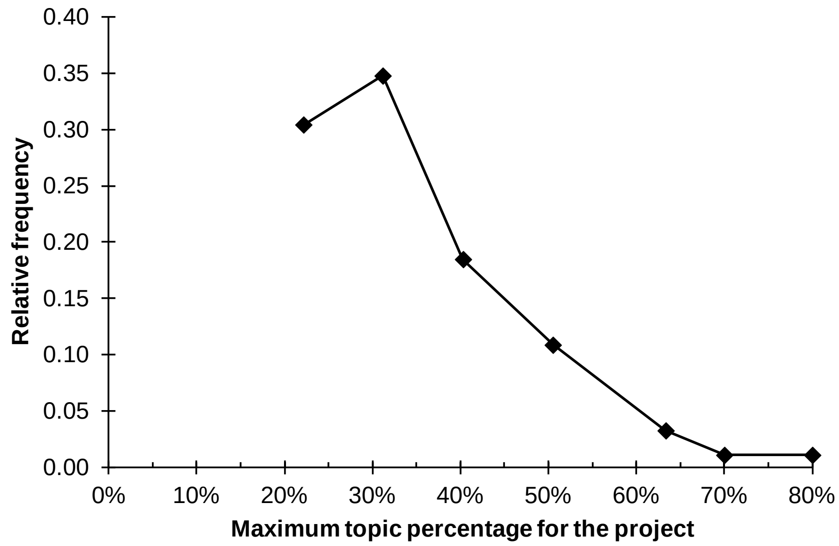

Figure 4 Relative frequency polygon of the maximum project topic times.

There are a few projects with one topic with a high percentage, which is usually MC. In particular, there are 16 projects with a maximum topic percentage of 45% or more (45%: 1, 50%: 9, 55%: 1, 60% 1, 65%: 2, 70%: 1, 80%: 1), and in 12 of these projects it is MC that is the maximum. The average topic percentages for these 12 projects and the rest of the projects are shown in Table 11. It is clear that a small proportion of projects have the characteristic that the time required for model coding is much higher than the time required for the other topics.

Table 11 Mean topic percentages for the 12 projects with MC ≥ 45% and for the other 80 projects.

	PS	*CM*	*DC*	*MC*	*VV*	*EX*	*RW*	*OT*
MC ≥ 45%	5.3%	6.1%	10.0%	57.5%	8.5%	5.8%	5.9%	0.8%
Other projects	12.1%	12.5%	16.8%	22.1%	13.8%	13.7%	8.3%	0.8%

5.6. Time allocation to the topics: Gantt chart timeline

The second question on the topics in the modelling process asked the respondents to complete a Gantt chart showing when the topics were worked on during the project. For the paper-based questionnaire handed out at the Simulation Workshop and WITNESS Conference, the respondents drew a horizontal line for each modelling topic on a timescale (no units were specified on the scale). For the online questionnaire, the project timescale was divided into 20 using a row of 20 check boxes for each of the 7 topics (i.e., a grid of 20 × 7 check boxes) so that the data could be entered easily. For each of the 20 time intervals the respondents could select which topics were worked on and could select more than one topic for each interval. Overall, there were 88 responses for this question of which 35 were from the paper-based questionnaire. In some cases no topic was allocated to a few of the time intervals.

In order to make the two types of Gantt chart compatible and to analyse the topic switching behaviour quantitatively, the Gantt charts for the paper-based questionnaires were converted into the same format as the online ones. This was done by dividing the timescale into 20 equal intervals and then reading off which topics were being worked in each interval. To be consistent with the online questionnaire the topics were simply recorded as either occurring or not in each interval. In most cases this was straightforward but where the line of a topic spanned only part of an interval, the topic was considered to occur if it covered at least half of the interval (judged subjectively).

The number of topics selected for each of the 1760 intervals (88 responses) were: 0 topics: 69 intervals (3.9%), 1 topic: 716 (40.7%), 2 topics: 538 (30.6%), 3 topics: 306 (17.4%), 4 topics: 100 (5.7%), 5 topics: 31 (1.8%). As the first step in analysing the Gantt charts, in each time interval in which a topic occurred the value assigned to it was 1 / n, where n is the number of topics worked on in that interval. This assumes (in the absence of any other information) an equal allocation of the time in an interval with multiple topics. Figure 5 is a graphical representation of the Gantt chart results for one of the projects (a representation format based on Figures 3 – 6 in Willemain, 1995) where the height of the bars is the value for each topic. The horizontal lines are 1 unit apart and so if only one topic was worked on in a period then the bar is of full height and reaches the horizontal line above. In Figure 5 the first time interval is just PS, the second is PS and CM, the third is CM and DC, and so on. Using this assignment and then summing the number of intervals for each topic for all the respondents, the topic percentages are PS: 10.4%, CM: 10.2%, DC: 15.9%, MC: 22.7%, VV: 13.8%, EX: 14.1%, RW: 12.9%. These are similar to the topic percentages in Table 10 and so this gives some confidence in the consistency of the responses to these two questions and in the reasonableness of the allocation assumption for multiple topics.

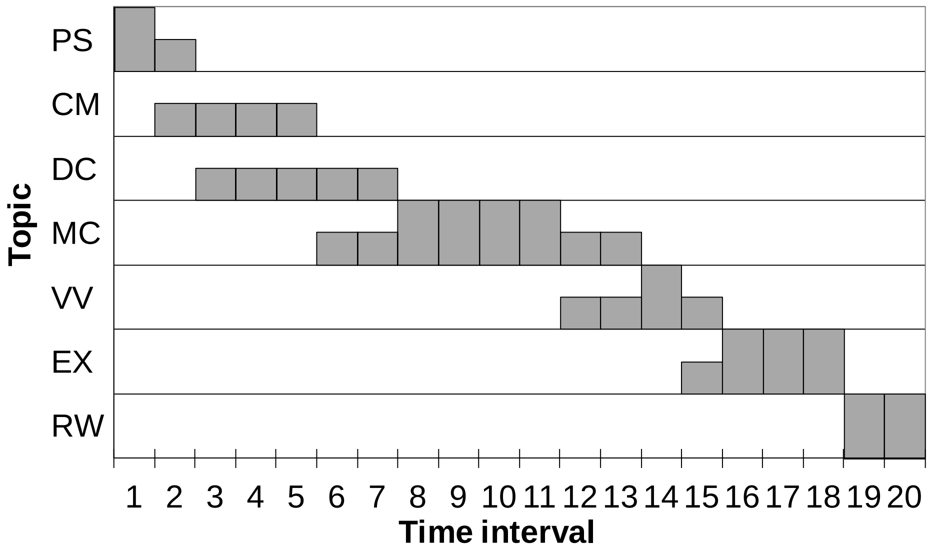

Figure 5 Example of Gantt chart data from one of the respondents.

The positions of the topics in the Gantt charts were analysed by calculating the average value for each topic in each of the 20 time intervals (ignoring intervals where no topic was chosen) and the results are shown in Figure 6. Labels are included above the peak value for each topic to make it easier to identify the topics. Each topic has a clear single peak with the peaks occurring in the expected order of the topics. Most of the topics have a considerable range of time periods over which they have a fairly high proportion.

Figure 6 provides the detailed average project pattern over the 20 time intervals. For example, across all the projects the first period is mainly problem structuring, the second period is mainly problem structuring, conceptual modelling and data collection, and so on. For this data, the intervals containing the median position for each topic are: PS: 2, CM: 4, DC: 6, MC: 10, VV: 14, EX: 16, RW: 19.

Most of the Gantt charts have some intervals in which several topics were worked on. The number of intervals with multiple topics was counted for each project and Figure 7 shows the frequency distribution. As can be seen in Figure 7, only eight projects had no such intervals. There is a fairly even distribution across the range of values with five projects having multiple topics in all 20 intervals. The average number of intervals per project with multiple topics is 11.08 (55% of intervals).

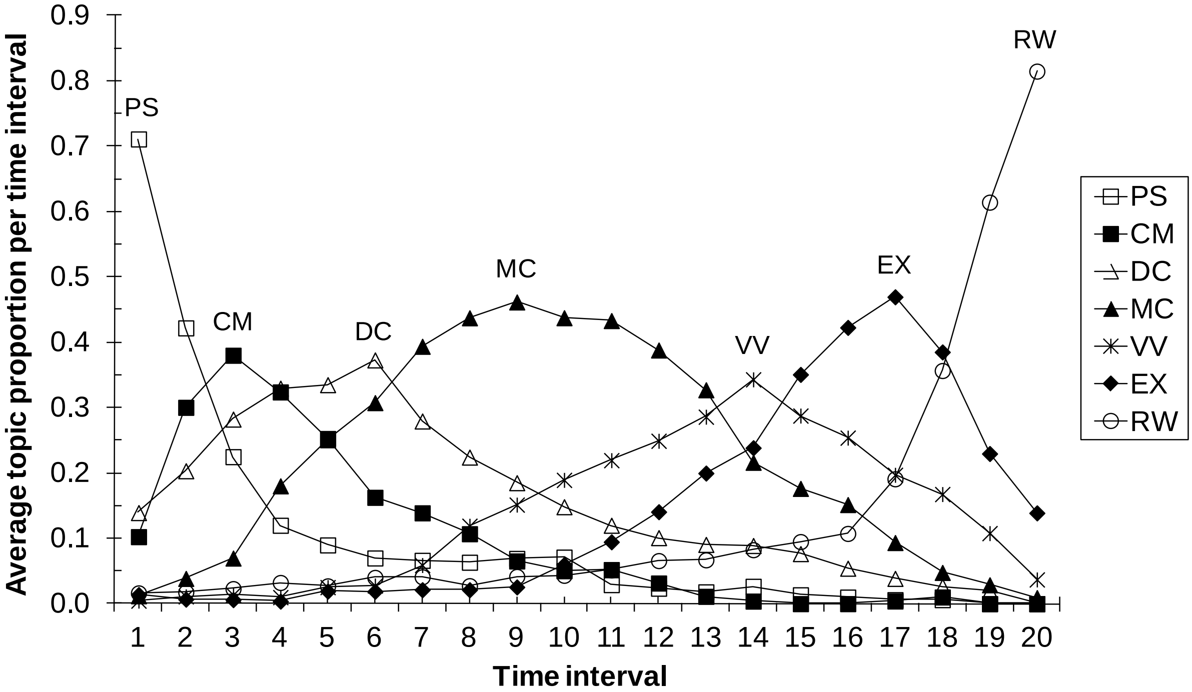

Figure 6 Average topic proportion for the Gantt charts in each of the 20 time intervals.

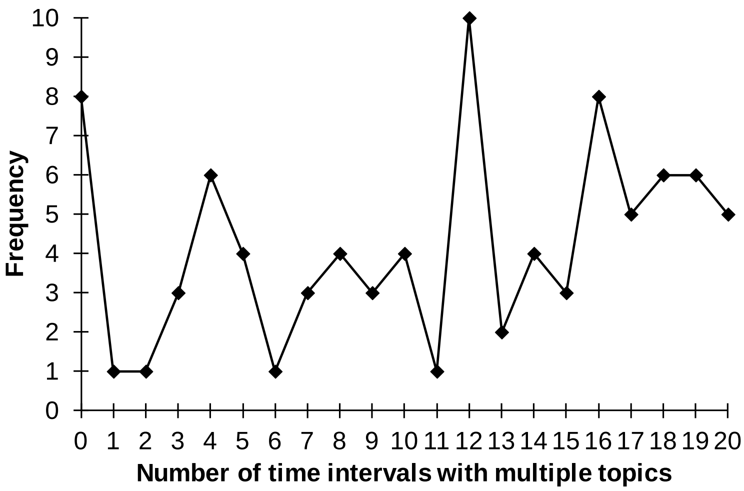

Figure 7 Frequency distribution for the number of intervals with multiple topics for each project.

Analysis of the intervals with multiple topics gives some insight into the relationships between topics. An interval with a pair of topics could be due to the topics being worked on in parallel, or a single transition between the topics, or several transitions between the topics. These alternatives apply to an interval with more than two topics but there is also the possibility of no direct interaction if the topics are separated in time by a third topic. However, it is assumed that in most cases each pair of topics occurring in the same interval means some interaction even when the number of topics in the interval is more than two. The frequency of occurrence of each pair was counted and the average number of intervals per project for each pair is shown in Table 12. For each topic the most common pair is with a neighbouring topic (in the usual order) with the number of pairs with other topics generally decreasing as the topic get further away. Topics nearer the middle of the project have more pairs (except for RW being higher than EX). Figure 8 is a chart of the data in Table 12 which shows the patterns more clearly. Each line is a topic with the values being the number of intervals for each pair that include that topic. The other topic in the pair is referred to by its relative position (using the order PS CM DC MC VV EX RW) as shown on the x-axis.

Table 12 Average number of intervals per project containing each pair of topics (total for topic is the total for all pairs involving that topic).

	PS	CM	DC	MC	VV	EX	RW	Total for topic
PS		1.92	1.82	1.17	0.59	0.22	0.35	6.07
CM			2.40	1.60	0.47	0.15	0.34	6.88
DC				3.64	1.38	0.52	0.92	10.67
MC					2.70	1.00	1.28	11.40
VV						1.73	1.25	8.11
EX							1.44	5.06
RW								5.59

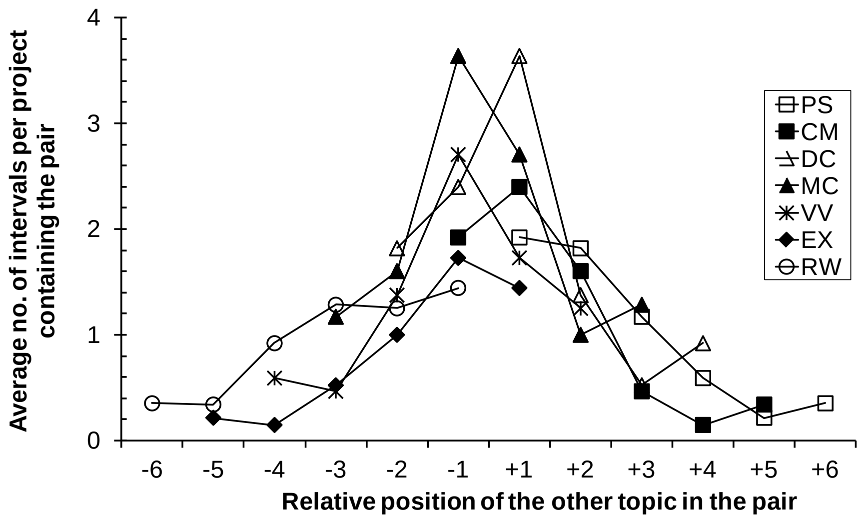

Figure 8 Chart of the data in Table 15 of the average number of intervals per project containing each pair of topics.

The total number of intervals represented by the data in table 12 is 26.89 (half the total per topic since each pair is counted twice). There are only 20 intervals in a project but some intervals have several pairs (intervals with 3 topics have 3 pairs, 4 topics have 6 pairs, 5 topics have 10 pairs).

The percentage of projects in which each pair of topics occurs at least once was also calculated and these values are given in Table 13. This shows a similar pattern to Table 12 with the highest percentages for each topic being the pair with a neighbouring topic. The percentages for pairs of successive topics are very high with such pairs occurring in most projects. One difference to Table 12 is that the percentages decrease slightly for later topics with the highest percentage being for the pair of the first two topics of PS and CM. Therefore, the earlier successive pairs occur in slightly more projects than the later successive pairs.

Table 13 The percentage of projects in which the pair of topics occurs.

	PS	CM	DC	MC	VV	EX	RW
PS		75%	59%	23%	10%	9%	10%
CM			73%	48%	13%	7%	10%
DC				69%	36%	20%	22%
MC					64%	32%	24%
VV						55%	35%
EX							49%
RW							

In Willemain (1995) the frequency of transitions between topics gave particularly interesting results in showing that modellers often alternated between formulation and assessment during the initial stages of formulating the model. In our survey it is not possible to identify all the transitions with certainty since the respondents could select more than one topic for each interval. However, estimates of the transitions were made using some assumptions and approximations, as follows: if a topic started or finished in an interval with multiple topics then it was assumed that the transition occurred in that interval; if several topics could be the transition topic then they were each allocated an equal fraction of the transition. Tables 14 and 15 illustrate this for some example situations.

Table 14 Illustration of the rules applied for calculating start transitions. The table illustrates four situations for topics denoted A, B, C and shows what would be counted as the start transition for A (topic A starts in interval *n*)

Interval *n* - 1	Interval *n*	Start transition to A from:
B	A	B
B, C	A	0.5B, 0.5C
Any except A	A, B	B
Any except A	A, B, C	0.5B, 0.5C

Table 15 Illustration of the rules applied for calculating finish transitions. The table illustrates four situations for topics denoted A, B, C and shows what would be counted as the finish transition for A (topic A finishes in interval *n*)

Interval *n*	Interval *n* + 1	Finish transition from A to:
A	B	B
A	B, C	0.5B, 0.5C
A, B	Any except A	B
A, B, C	Any except A	0.5B, 0.5C

The average numbers of start and finish transitions per project calculated using this method are shown in Tables 16 and 17. For example, the first value in the top row of table 16 shows that on average there were 0.499 start transitions per project from PS to CM. This means that when CM started the previous topic was PS. The columns in Table 16 therefore show the start transitions to each topic. In Table 17 the first value in the top row shows that on average there were 0.516 finish transitions per project from PS to CM. This means that when PS finished the next topic was CM. The rows in Table 17 therefore show the finish transitions from each topic.

In most (but not all) cases a start transition is also a finish transition and so the values in the two tables are very similar. For example, a transition from PS to CM is often a start transition for CM and a finish transition for PS in which case it will be included in the PS to CM cell in both tables.

As with the pairs of topics the most common transitions for each topic are from or to the neighbouring topic. The total number of start and finish transitions for each topic (column totals in Table 16 and row totals in Table 17) are fairly close to 1, except for the PS start transitions and RW finish transitions. This means that the topics generally just have one transition and therefore the topics take place in a single block of time in our Gantt format rather than several episodes separated by intervals without the topic. The much lower value for the start transitions for PS is because it is usually the first topic in the project (and so doesn't have a start transition), and similarly the much lower value for the finish transitions for RW is because it is usually the last topic.

Table 16 Average number of start transitions per project.

From \ To	PS	CM	DC	MC	VV	EX	RW	Total
PS		0.499	0.215	0.101	0.044	0.019	0.027	0.904
CM	0.042		0.366	0.301	0.044	0.011	0.032	0.796
DC	0.045	0.211		0.483	0.181	0.101	0.097	1.118
MC	0.019	0.033	0.102		0.705	0.202	0.152	1.213
VV	0.027	0.018	0.037	0.049		0.603	0.191	0.925
EX	0.008	0.015	0.033	0.034	0.060		0.638	0.788
RW	0.008	0.019	0.019	0.054	0.092	0.154		0.346
Total	0.148	0.795	0.773	1.023	1.125	1.091	1.136	6.091

Table 17 Average number of finish transitions per project.

From \ To	PS	CM	DC	MC	VV	EX	RW	Total
PS		0.516	0.272	0.122	0.047	0.017	0.037	1.011
CM	0.110		0.473	0.345	0.045	0.021	0.041	1.034
DC	0.086	0.170		0.503	0.150	0.064	0.107	1.080
MC	0.041	0.020	0.082		0.644	0.170	0.099	1.057
VV	0.023	0.016	0.083	0.140		0.592	0.203	1.057
EX	0.015	0.006	0.081	0.080	0.094		0.598	0.875
RW	0.013	0.032	0.040	0.104	0.040	0.044		0.273
Total	0.288	0.759	1.032	1.295	1.019	0.908	1.085	6.386

Overall, therefore, the general average pattern is of the topics taking place in blocks of time in the expected order but with quite a bit of overlapping between successive topics. The project shown in Figure 5 is a good example of such a pattern. A review of the patterns in each project supports this conclusion. There are of course differences in the patterns between different projects, with a few having some unusual characteristics. An attempt was made to identify and classify different patterns by manual comparison but this is rather subjective and it was hard to identify clear groups. Further work could perhaps involve the use of clustering algorithms.

5.7. Whether the project is typical

In the final question in the survey respondents were asked whether the project was typical of most projects they are involved with and, if not, in what ways. The results as to whether the project was typical were: yes: 68 (71%), no: 28 (29%). Out of those who said 'no' 24 provided a reason and Table 18 shows our subjective categorisation of the responses. Four of the responses were considered to contain comments in two of the categories giving a total of 28 responses in Table 18.

Table 18 Categorisation of reasons for the project not being typical.

Aspect that is different to normal projects	*No. responses*
Nature of project (different scope, client, project arrangements or objectives)	9
Larger or more complex	4
Smaller, simpler or shorter timescale	4
Re-using an existing model rather than developing it from scratch	3
All projects vary	3
Data collection	2
Model coding / software	2
Verification, validation and experimentation	1
Total	28

There were 64 'typical' and 26 'not typical' respondents who also completed the topic percentages. The average percentages for the two groups are shown in Table 19. The main difference is the difference in model coding (MC), being 23.8% for typical projects and 32.7% for non-typical ones. The statistical significance of this difference was tested by selecting 26 responses randomly from the 90 overall responses and calculating the largest topic difference between the 26 chosen and the remaining 64. Repeating this 100000 times, in only 1454 occasions did the maximum topic difference exceed the value from the actual data of 8.90% (i.e., a p value of 0.01454). This means that there is quite strong evidence of a difference between the topic percentages of the two groups. The main impact of tackling a relatively unfamiliar type of project therefore appears to be that it requires a higher percentage of time in building the model. It may be that in a non-typical project all topics take longer but the effect is greatest on the time for model coding. This could be useful information when planning for an unfamiliar project.

Table 19 Mean topic percentages for typical and non-typical projects.

	PS	*CM*	*DC*	*MC*	*VV*	*EX*	*RW*	*OT*
Not typical	10.0%	10.8%	14.5%	32.7%	12.3%	11.2%	7.7%	1.0%
Typical	11.9%	12.2%	16.5%	23.8%	13.4%	13.3%	8.1%	0.8%
Difference	-1.9%	-1.4%	-2.1%	8.9%	-1.1%	-2.1%	-0.5%	0.2%

5.8. Effect of experience

We wished to investigate whether there was evidence in the survey results of modelling styles or behaviours changing with experience. Most of the data is categorical and so this requires dividing the respondents into categories by level of experience. The survey asked the respondents for both their number of years of simulation experience and the number of projects carried out. We decided that the number of years was the better measure of experience. We particularly wanted to see if the more inexperienced modellers would have different approaches and so we chose a cut-off of 3 years experience. Therefore we divided the respondents into those with three years or less experience (28 respondents) and those with more than three years experience (74 respondents). This also gives a reasonable sample size for the less experienced group (of between 24 and 28 depending on the number answering the particular question). The number of years of experience of this group is 1 year: 6; 1.5 years: 1; 2 years: 13; 3years: 8. Each of the results in this paper was then split into the results for the two groups. However, for none of the questions was a statistically significant difference found at the 5% significance level between the two groups. Therefore, there is no strong evidence in the survey of experience having an effect on behaviour for this group of simulation modellers. This may be because even the least experienced of the respondents are still simulation specialists who have had enough experience and training to behave in a reasonably expert manner. It could also be that many of the least experienced respondents were supervised by or worked as part of a team with more experienced colleagues

6. Discussion

6.1. Conceptual modelling

There is very little previous data on what experts do for conceptual modelling and so this study provides new information on this. We will try and summarise here the main results and the most common methods and ways of working.

Understanding the system generally involves some combination of analysing system data, observing the system and talking to staff. In developing the initial conceptual model previous experience was used in many cases. In addition, preliminary analysis, such as a simple analytical model, was also widely used. Considering these two stages, the data analysis and preliminary analysis tend to imply experts being keen to get a good feel for the important factors and the main interactions in the system at an early stage in the project.

In almost all the projects (96%) the conceptual model was documented in some way. Given that there isn't a standard documentation format and also the lack of information on conceptual modelling in most textbooks, it is perhaps surprising that this is such a high percentage. This presumably indicates the importance that experts attach to documenting what is done on the project and emphasises that developing better documentation methods would be a useful area of conceptual modelling research. More than half the projects had a list of assumptions and simplifications and so the experts presumably feel that it is important to state these explicitly. In general in modelling it is easy to make implicit assumptions and for these not to be apparent to the modeller (particularly inexperienced modellers) or for them to be forgotten in the subsequent interpretation of the modelling results.

In most projects one conceptual model is developed but this is then changed during the project. The conceptual model is nearly always started before model coding and is often finished before coding starts. Changes to the conceptual model are common in each of the four subsequent stages DC, MC, VV, EX. Even at the experimentation stage (EX), change occurred in 44% of projects. There is a strong tendency for the conceptual model changes to add complexity and the most of the experts also had a preference towards 'start small and add' when describing their modelling style. Changes are most often a result of better information about the real system, although the other three reasons provided in the questionnaire also occur. It is commonly said in OR and simulation textbooks that the project process is iterative rather than linear, and these results provide evidence of what that actually means in practice on simulation projects for conceptual modelling.

6.2. Time allocation to topics

The time allocation questions were inspired Willemain (1995) and to an extent can be compared with those results. However, it is important to recognise that timeline data for a project will depend on the resolution of the data collection. In our case the project was split into 20 time intervals which is quite a coarse resolution but this was necessary to cover the whole project. Willemain's (1995) experiment was very contrasting data since this was obtained by using think aloud protocols and analysing the data in parts of sentences spoken. Therefore it is probably the most detailed resolution possible but only covered the first hour of thinking about an OR problem. As a result the pattern was very different with lots of switching between topics. Examining any part of a project at such a micro level may well identify a similar pattern. This is the likely explanation for differences in results between two of the questions in our survey. The respondents stated that in the later stages in the project a change in the conceptual model was often made (Tables 7 and 8) and yet few projects have an overlapping interval or transitions between conceptual modelling and these topics in the timeline data. In particular in 44% of projects the conceptual model changes during experimentation (Tables 7 and 8) but only 7% of projects have an interval in which these topics overlap. This is presumably because the time spent on revising the conceptual model was too short to be included in the Gantt chart.

The survey provides timeline data on the whole project. The average or typical pattern was MC taking about twice the time as the other topics, and the topics appearing mainly in single blocks in the expected order in the timeline but quite widely spread and with many overlapping intervals. The extensive occurrence of overlaps may again be a result of and evidence for the iterative nature of the project process.

6.3. Implications for novices and for teaching simulation

The literature review section described the previous work by Willemain and then Willemain and Powell in studying the behaviour of experts and novices in formulating a model. A key aspect of the experts' approach and problem solving strategy was the continual critical evaluation of the model and their willingness to alter the model and their approach. By contrast, novices often continued along a particular track even when it should have been clear that it was not working and that they were not making progress in tackling the problem.

We interpret many of the results from our survey as indicating this approach taking place throughout the project in all the tasks, and so reinforcing and extending this message about how experts work and their mindset. They evaluate what they do and are prepared to revise their approach and revisit previous steps. One aspect of this is changing the conceptual model. This starts at the initial stages of understanding the problem and developing the initial conceptual model where the experts will often do some analysis and simple modelling. This preliminary model and the understanding gained are used to develop the initial conceptual model. The experts then usually revise the conceptual model during the following simulation stages and are prepared to do this for various different reasons. The experts like to document the conceptual model and we would suggest that the documentation and particularly listing the assumptions and simplifications are part of the critical evaluation of the model. Considering other aspects of the project, the average timeline pattern is topics being widely spread with lots of overlaps. Again this probably indicates critical evaluation and being prepared to go back and revise previous tasks.

Various other results may be picked up as useful information for novices. For example, the modelling style questions show evidence of both a systematic approach and creativity and we would suggest that an ideal style needs to combine these two attributes. The tendency is to start with a small model and add detail. Projects generally have multiple objectives but improving understanding of the system (i.e., not just taking a black box modelling approach) is usually one of these and quite often the most important objective. Identifying and documenting assumptions and simplifications is important. Task times clearly can vary on different types of project but model coding typically takes the most time. The coding is also planned with the conceptual model usually specified fully before coding starts. In addition to specific points, the detailed results for each survey question should help novices to picture what happens on a simulation project and how a typical project progresses.

7. Conclusions

The survey results provide a detailed picture of expert practice that can inform simulation education and practice, and provide an empirical basis for research on conceptual modelling and the simulation project process. More research is needed on all aspects of conceptual modelling but the survey at least provides evidence of current practice, for example in the methods for developing the model and in the documentation used. In the task time data and analysis the paper gives an overview of the whole project showing the time spent, the timeline and the interactions between tasks. In further work it would be interesting to examine some of the topics at a micro level, similar to the Willemain (1995) experiment on the model formulation task.

Acknowledgements

We are very grateful to all the questionnaire respondents for their assistance. Some of the text and results in this paper were previously included in two conference papers: Wang W and Brooks RJ (2008), Conceptual modelling and the modelling process in simulation projects: a survey of simulation modellers, In Kotiadis K, Currie C, Robinson S and Taylor S (eds), Proceedings of the 2008 OR Society Simulation Workshop, pp. 107-115; Wang W and Brooks RJ (2007), Empirical investigations of conceptual modeling and the modeling process,

In: Henderson SG, Biller B, Hsieh M-H, Shortle J, Tew JD and Barton RR (eds), Proceedings of the 2007 Winter Simulation Conference, IEEE, pp. 762-770.

References

Aitchison J (1986). *The Statistical Analysis of Compositional Data*. Chapman and Hall: London.

Bell PC, Anderson CK, Staples DS and Elder M (1999). Decision-makers' perceptions of the value and impact of visual interactive modelling. *Omega, The International Journal of Management Science* **27**: 155-165

Brooks RJ and Robinson S (2001). *Simulation*, with Inventory Control (author C. Lewis), Operational Research Series. Palgrave: Basingstoke.

Brooks RJ and Tobias AM (1996). Choosing the best model: Level of detail, complexity, and model performance. *Mathematical and Computer Modelling* **24**(4): 1-14.

Brooks RJ (2011). Complexity, level of detail, and model performance. In: Robinson S, Brooks RJ, Kotiadis K, van der Zee D-J (eds). *Conceptual Modeling for Discrete-Event Simulation*, pp. 31-56.

Chwif L, Barretto MRP and Paul RJ (2000). On simulation model complexity. In Joines JA, Barton RR, Kang K and Fishwick PA (eds). *Proceedings of the 2000 Winter Simulation Conference*, IEEE, pp. 449-455.

Cochran JK, Mackulak GT and Savory PA (1995). Simulation Project Characteristics in Industrial Settings. *Interfaces* **25**(4): 104-113.

Gass SI (1987). Managing the modeling process: a personal reflection. *European Journal of Operational Research* **31**(1): 1-8.

Law AM (1991). Simulation models' level of detail determines effectiveness. *Industrial Engineering* **23**(10): 16-18.

Law AM (2007). *Simulation modeling and analysis*, 4th edition. McGraw-Hill: NY.

Melao N and Pidd M (2003). Use of business simulation: A survey of practitioners. *Journal of the Operational Research Society* **54**(1): 2-10.

O'Brien F (2011). Supporting the strategy process: a survey of UK OR/MS practitioners. *Journal of the Operational Research Society* **62**(5): 900-920.

Pidd M (1999). Just Modeling Through: A Rough Guide to Modeling. *Interfaces* **29**(2): 118-132.

Powell SG and Willemain TR (2007), How novices formulate models. Part I: qualitative insights and implications for teaching. *Journal of the Operational Research Society* **58**(8): 983-995.

Pritsker AAB (1986). *Introduction to Simulation and Slam II*. Systems Publishing Corporation: West Lafayette, Indiana.

Robinson S (2004). *Simulation: the practice of model development and use*. John Wiley and Sons: Chichester.

Robinson S (2006). Conceptual modeling for simulation: Issues and research requirements. In Perrone LF, Wieland FP, Lawson BG, Nicol DM and Fujimoto RM (eds). *Proceedings of the 2006 Winter Simulation Conference*, IEEE, pp. 792-800.

Robinson S (2008a). Conceptual modelling for simulation Part I: definition and requirements. *Journal of the Operational Research Society* **59**(3): 278-290.

Robinson S (2008b). Conceptual modelling for simulation Part II: a framework for conceptual modelling. *Journal of the Operational Research Society* **59**(3): 291-304.

Robinson S, Brooks RJ, Kotiadis K and van der Zee D-J (2011). *Conceptual Modeling for Discrete-Event Simulation*. CRC Press, Taylor & Francis Group: Boca Raton.

Salt JD (1993). Keynote address: Simulation should be easy and fun! In Evans GW, Mollaghasemi M, Russell EC and Biles WE (eds). *Proceedings of the 1993 Winter Simulation Conference*, IEEE, pp. 1-5.

Tilanus CB (1985). Failures and successes of quantitative methods in management. *European Journal of Operational Research* **19**(2): 170-175.

Ward SC (1989). Arguments for constructively simple models. *Journal of the Operational Research Society* **40**(2): 141-153.

Wang W and Brooks RJ (2007). Improving the understanding of conceptual modelling. *Journal of Simulation* **1**(3): 153-158.

Waisel LB, Wallace WA and Willemain TR (2008). Visualization and model formulation: an analysis of the sketches of expert modellers. *Journal of the Operational Research Society* **59**(3): 353-361.

Willemain TR (1994). Insights on modeling from a dozen experts. *Operations Research* **42**(2): 213-222.

Willemain TR (1995). Model formulation: what experts think about and when. *Operations Research* **43**(6): 916-932.

Willemain TR and Powell SG (2007). How novices formulate models. Part II: a quantitative description of behaviour. *Journal of the Operational Research Society* **58**(10): 1271-1283.