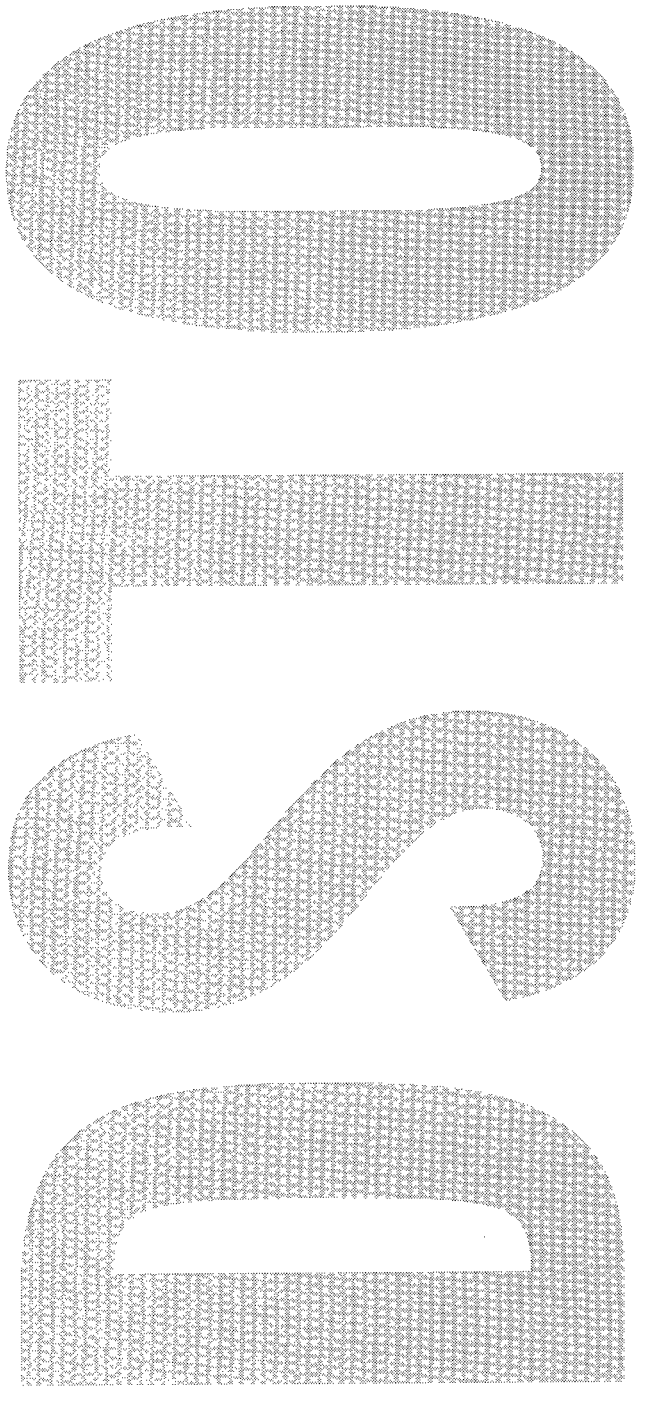
DEPARTMENT OF DEFENCE
DEFENCE SCIENCE & TECHNOLOGY ORGANISATION
DSTO



Software Design Reviews Using the Software Architecture Analysis Method: A Case Study

Gina Kingston

DSTO-RR-0170

Software Design Reviews Using the Software Architecture Analysis Method: A Case Study

Gina Kingston

Information Technology Division
Electronics and Surveillance Research Laboratory

DSTO-RR-0170

ABSTRACT

Software reviews, including design reviews, are conducted on most software-intensive Defence projects and are an important component of the software acquisition process. However, software reviews are often conducted in an *ad hoc* manner, and many are inefficient. This report investigates an alternative process for reviewing software designs that is based on the Software Architecture Analysis Method (SAAM).

The SAAM review process is driven by the identification of scenarios that capture how the system might be used or modified. This report describes a case study of the SAAM review process. According to the results of the study, the SAAM review process offers potential benefits over the traditional design review process in the identification and clarification of requirements, but was less effective at identifying conflicts and trade-offs. Consequently, it is recommended that projects continue to use traditional review processes, and where appropriate, supplement these reviews with SAAM reviews to clarify and identify requirements.

RELEASE LIMITATION

Approved for public release

DEPARTMENT OF DEFENCE
DEFENCE SCIENCE & TECHNOLOGY ORGANISATION
DSTO

Published by

DSTO Electronics and Surveillance Research Laboratory
PO Box 1500
Salisbury South Australia 5108 Australia

Telephone: (08) 8259 5555
Fax: (08) 8259 6567

AR-011-217
February 2000

APPROVED FOR PUBLIC RELEASE

Software Design Reviews Using the Software Architecture Analysis Method: A Case Study

Executive Summary

Software reviews are conducted on most, major, software-intensive, Defence projects and are an important component of the software acquisition process. However, software reviews are often conducted in an *ad hoc* manner, and many are inefficient. This report investigates an alternative process for the review of software designs that is based on the Software Architecture Analysis Method (SAAM).

The SAAM process uses a facilitator and scenario-driven review process. That is, participants identify many scenarios that describe the potential uses of system and select some of these scenarios for further evaluation during the review.

This report describes a case study that was conducted to characterise the SAAM review process and to allow a preliminary comparison between SAAM reviews and traditional software reviews. The reviews were characterised according to the goals and roles of the participants, the perceived benefits of using scenarios and a facilitator, the performance of the review, and the negotiation stumbling blocks encountered by the participants.

The case study focused on the SAAM review of the EXC3ITE OCP/IMAD architecture that was developed to explore the feasibility and benefits of incorporating OCP technology into Concept Technology Demonstrators, such as IMAD. Four sources of information were used to characterise the EXC3ITE review: a pre-review questionnaire; observation of the review; a post-review questionnaire; and meeting artefacts, such as the minutes of the review meeting.

The results of the case study were compared with the results of previous field (Project Llama, JP2030) and laboratory studies of software reviews. These studies focused on software reviews conducted in different environments – eg within a laboratory or industrial setting rather than a research environment. Therefore, the results only provide an indication of the relative strengths and weaknesses of SAAM reviews and traditional reviews.

It appears that SAAM reviews offer benefits over traditional software reviews in clarifying and refining requirements. Therefore, SAAM reviews might be beneficial in Defence projects, such as those procured using Evolutionary Acquisition, that do not have a complete set of clearly defined requirements. Other potential benefits of SAAM reviews, which need to be confirmed by additional studies, include increased

participation from meeting attendees and greater retention, or more detailed records, of the issues discussed during the meetings. However, SAAM reviews provide only a limited ability to identify conflicts and trade-offs. Furthermore, the SAAM participants perceived that some of the SAAM procedures, such as the voting process, were biased. It is possible that modifications to the voting process, and increased facilitator training could address both these limitations of SAAM reviews. However, until additional studies are conducted, SAAM reviews cannot be recommended as a replacement for traditional design reviews. It is recommended that Defence continue to conduct traditional reviews and that, where appropriate, these reviews are supplemented by SAAM reviews.

Author

Gina Kingston
Information Technology Division

Gina has been employed in the Software Systems Engineering Group of the Information Technology Division of the Defence Science and Technology Organisation (DSTO) since graduating from the University of Tasmania with a BSc with First Class Honours in Mathematics in 1990. Her research interests include process improvement and empirical software engineering. She has conducted research into software costing, the analysis of software, and software acquisition. She has recently submitted a PhD in Software Reviews through the School of Information Systems at the University of New South Wales.

Contents

List of Tables

List of Figures

1. Introduction

Joint software reviews are conducted on most major, software-intensive, Defence projects, and they form an important component of the software acquisition processes (MIL-STD-498 1994; ISO/IEC 12207 1995). Defence invests considerable resources in the conduct of reviews, including the review of software-related artefacts. Regular reviews are conducted on major acquisitions, often involving tens of people for days at a time. However, ad-hoc evidence collected during interviews by the author in 1997 reveals that the participants in many joint software reviews consider them to be inefficient. That is, the review participants believe that it takes a long time to identify issues or defects using joint software review techniques. Nevertheless, it is acknowleged that joint software reviews are required because of the benefits of early issue detection.

The negotiation literature (eg Luthans 1985; Shea and Guzzo 1987), and formal studies of software reviews (Kingston et al. 1999c) indicate that the joint software reviews are inefficient and that the poor performance of joint software reviews arises because of differences in the goals of the review participants. In the presence of different or conflicting goals, the participants in meetings, such as joint software reviews, encounter more negotiation stumbling blocks when their goals are different or conflicting, than when they are similar (Foroughi et al. 1995; Kingston et al 1999c).

This report describes a study of an alternative review process with the potential to align the goals of participants and to reduce the negotiation stumbling blocks encountered by the participants. The alternative process is based on the Software Architecture Analysis Method (SAAM) proposed by the Software Engineering Institute (SEI) (Kazman et al. 1996). SAAM reviews are based around scenarios, which the participants select as the focus for the review. SAAM reviews also use an independent facilitator to coordinate the review. Section 2 describes the SAAM review process in more detail.

The proponents of SAAM claim that the approach helps elicit the goals of the stakeholders and clarify the goals for the review (Bass et al. 1998). The negotiation literature suggests that facilitators can be used to reduce negotiation stumbling blocks. For example, Lewicki (1992) identifies several techniques that facilitators have used to reduce the impact of negotiation stumbling blocks. These include reducing the impact of conflict by ensuring that discussions are based on facts and not suppositions, by identifying areas on which the negotiators agree, and by identifying alternative options that stimulate the negotiation. This report describes a study that investigated the validity of these claims in the context of joint software reviews.

Section 3 defines the research objectives for the study and a model showing the importance of negotiation stumbling blocks on the review process. The study was designed not only to compare SAAM reviews with other software reviews, but also to

characterise SAAM reviews. The characterisation was designed to capture information about the activities conducted during the review, the participants' goals, the negotiation stumbling blocks encountered during the review, and the performance of the review. This characterisation provides valuable information about the review process, and could also form the basis for additional, more detailed studies of SAAM reviews.

The study was conducted on the review of the architecture for the EXperimental C3I Technology Environment (EXC3ITE) OCP/IMAD system. The prototype architecture aimed to investigate the integration of the Object Computing Platform (OCP) tool suite, developed by Object-Oriented Propriety Limited (OOPL), with the Image Management and Dissemination (IMAD) testbed developed by DSTO. Section 4 discusses the research method and explains why a case study of the EXC3ITE OCP/IMAD system was conducted. The study design uses questionnaires, observations and analysis of the meeting minutes to provide information about SAAM reviews. These three techniques were used to provide information about three distinct areas. The first area concerns the participants' goals, and how the *goals* were managed during the SAAM review (Section 5). The second area concerns the *performance characteristics* of the review (Section 6). The third area relates to the *negotiation stumbling blocks* encountered during the review (Section 7). Sections 5-7 each discuss the detailed research questions and hypotheses related to the area, the variables and analysis approach used to investigate the area, and the relevant results from the study.

Section 8 considers the general implications of the results from all three areas given the strengths and limitations of the study design, and Section 9 summarises the implications, and provides recommendations, for Defence.

2. SAAM Reviews

The SAAM review process is a modern review technique that aims to aid the evaluation and understanding of software architectures, and 'to address quality concerns such as maintainability portability, modularity, reusability, and so forth" [Kazman et al, 1994]. SAAM reviews are a modern review technique that is becoming increasingly popular. The main features of the SAAM review process are the use of a facilitator and a special review process that is driven by scenarios. The participants in a SAAM review identify, and select, scenarios to drive discussions about the software architecture. Section 2.1 describes the SAAM review process in detail.

A *scenario* is a "brief description of a single interaction of a stakeholder with a system" (Bass et al. 1998). Table 1 gives some examples of SAAM scenarios that are taken from (Bass et al. 1998). Note that other software review techniques, such as Porter et al's (1995) scenario-based inspection approach, use different definitions of the term 'scenario'. SAAM scenarios are developed to evaluate and explore the architecture or design of a specific system.

SAAM scenarios were proposed as a means of capturing the goals of the stakeholders and then selecting a subset of these goals as the goals for the review of the software architecture (Bass et al. 1998). If successful, this approach should improve the performance of joint software reviews by aligning the goals of the review participants – the presence of different or conflicting goals has been shown to reduce the number of issues raised during a software review (Kingston et al. 1999c).

Table 1: SAAM scenarios. Adapted from Bass et al. 1998.

System	The KWIC system takes sentences as input and outputs permutations (circularly shifted) of those sentences in alphabetical order. The KWIC system could be required to operate in an incremental or a batch fashion. An incremental version would work by accepting one sentence at a time and producing an alphabetical list of all permutations of all sentences that had been given as input to date.
Scenarios	Make the KWIC program eliminate entries beginning with "noise" words.
	Change the internal representation of sentences (eg compressed versus uncompressed).
	Change the internal representation of intermediate data structures (eg either change the shifted sentences directly or store an index to shifted words).

The proponents of the SAAM approach have provided anecdotal evidence that scenarios can be used to align the goals of the participants (Bass et al. 1998). Weidenhaupt et al (1998) also found that scenarios – similar to SAAM scenarios – could help align the goals of participants involved in capturing the requirements for a software system. They studied 15 organisations and spent up to one day interviewing some of the developers (usually the project leader) about the characteristics, strengths and weaknesses of scenarios. They did not interview other stakeholders in the development, such as the clients or users of the systems. They found that the developers believed that scenarios helped align the goals of stakeholders by focusing them on particular aspects of a system under discussion. This allowed the stakeholders to agree on some areas while other areas were left unresolved. Scenarios might allow the same alignment of goals in SAAM reviews. However, a scenario may be included in a SAAM review without all of the participants agreeing on the scenario. This situation arises because of the voting scheme used in SAAM reviews. Thus, the ability of scenarios to align the goals of the participants in a SAAM review requires further investigation.

2.1 SAAM Review Process

The main difference between SAAM reviews and traditional software reviews is the variety of activities conducted during a SAAM meeting. The documented (Bass et al.

1998) approach to SAAM reviews consists of nine steps. However, there was only limited control over the process used during the investigation, because the study was conducted on a live industrial review. The facilitator, who was experienced at resolving conflicts during negotiations, varied the process slightly from the SAAM approach as described by its proponents in (Bass et al. 1998), given in Figure 1. The process used for the EXC3ITE OCP/IMAD review is shown in Figure 2. Details of each of the process steps and how they vary from the formal SAAM approach are now described.

Preparation

The participants received 255 pages of documentation. Some of the documentation was received two working days before the review and some of the documentation was received one working day before the review. The documentation consisted of the main 28-page document that summarised the architectural requirements, concepts, solution and guidelines; 182 pages of more detailed information about the architecture and its components; and 45 pages of supporting material. The supporting material described the SAAM review process and provided guidelines for the participants. The participants were asked to familiarise themselves with the documentation and to identify scenarios before the review meeting. The facilitator was given the same information as the other participants, but also received a half-day briefing on the planned series of EXC3ITE review, with particular emphasis on the EXC3ITE OCP/IMAD review. The review organisers arranged the briefing, which covered the review process, including the main objectives of the review, and possible barriers to the success of the review.

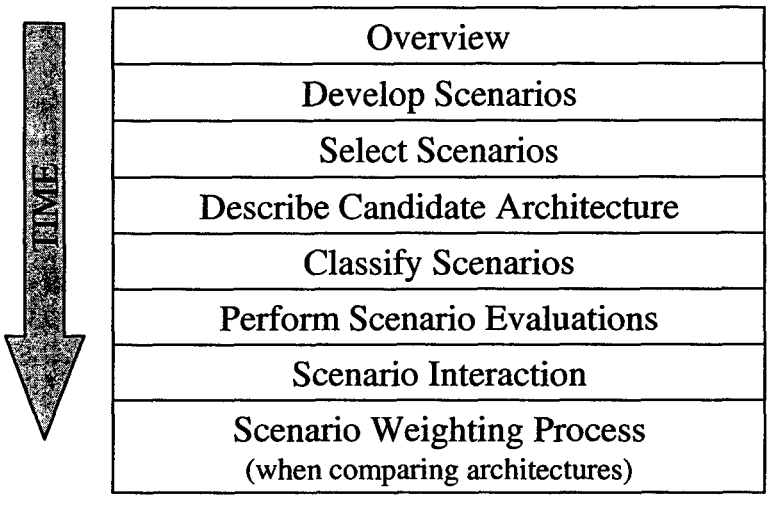


Figure 1: Nominal steps in the SAAM review process.

Overview

The meeting commenced with a thirty-minute overview during which the participants were introduced, the SAAM process was outlined and the scope of the review was defined. The facilitator also chaired a needs analysis session during the overview. The review participants had about five minutes to discuss their objectives and expectations

for the review. The facilitator then summarised input from all the participants in a diagram on a whiteboard.

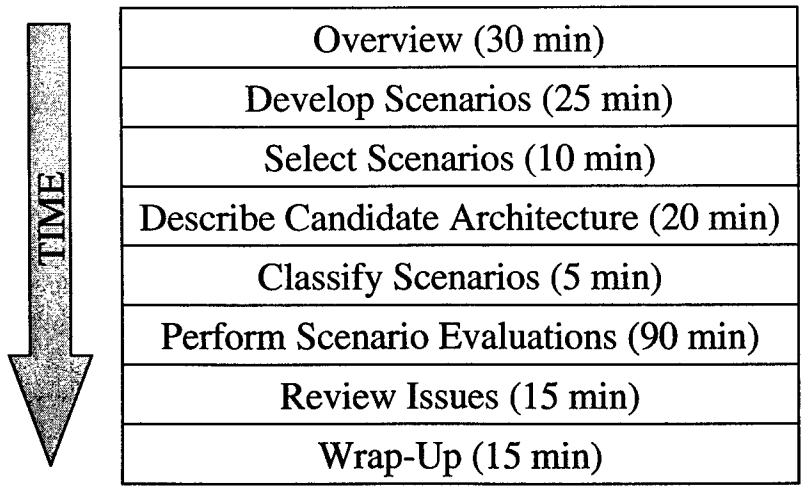


Figure 2: Steps in the SAAM review process as planned and conducted.

Develop scenarios

The scenario development process is usually a brainstorming session followed by an elaboration session (Bass et al. 1998). During the brainstorming session, participants gave short (<1 minute) descriptions of the scenarios they identified, and the facilitator recorded the scenarios using a key phrase or title of three-to-five words from the description. The purpose of the elaboration session is to clarify the descriptions of the scenarios and to combine, group and discard scenarios.

The participants in the EXC3ITE OCP/IMAD review attempted to combine, group and discard scenarios without providing clarification and detailed descriptions of any of the thirty-five scenarios that were identified during the brainstorming session. However, none of the scenarios were combined or identified as similar before commencing the next step, despite similar scenarios being identified later. This is discussed further in Section 5.4, but is probably because very little time was spent discussing or elaborating the scenarios. This might have been due to time constraints (30 minutes was allocated for scenario development, and 25 minutes was actually used) or it might have been due to the facilitator's and participants' inexperience with SAAM reviews.

Select scenarios

The SAAM review process recognises that it is not possible to address all possible scenarios within a particular review. Therefore, the participants have to select the 'most important' scenarios as the focus for the review. In the EXC3ITE OCP/IMAD review, five scenarios were selected from the thirty-five scenarios identified during the brainstorming session using a voting process. The voting process was determined by two of the review organisers based on a SAAM training session that they had just attended. The voting was conducted over two rounds with the intent that each

stakeholder had 10 votes in each round that they could allocate to any of the scenarios. However, one of the stakeholders was accidentally omitted during the first round of voting. (The research observers did not interfere with the review process. Therefore, while they noted the omission they did not inform the facilitator. One of the reviewers noted the omission during the second round, and the omitted stakeholder was then given an opportunity to vote.) The facilitator, the recorder, the architect, and the review observers did not get any votes. Participants were asked to vote on the scenarios that were most important to them, and that would help them to address their goals. During the second round they were given the opportunity to adjust their voting patterns based on the results of the first round. The voting process is described in more detail in Section 5.4.

Describe candidate architecture
After the scenarios had been selected, the architect gave a 20-minute introduction to the system's architecture and its rationale.

Classify scenarios
The selected scenarios were then classified as direct scenarios or indirect scenarios. Direct scenarios are supported by the current architecture, while indirect scenarios require modifications to the architecture before they can be supported (Bass et al. 1998). There was some confusion over the definition of a direct scenario[1]. The participants were not sure whether a direct scenario was a scenario that *was* supported by the architecture, or one that *should be* supported by the architecture. Furthermore, the participants were not sure whether a scenario that involved components that were neither *included in*, nor *excluded by*, the architecture was *supported* by the architecture. That is, does the architecture support a scenario, if the architecture does not currently implement the scenario, but can easily be modified to implement the scenario (if required). Despite this confusion the participants classified one scenario as direct, and said that it should be included in the architecture. The other four were classified as indirect. This classification was used to determine which scenario to evaluate first (the direct scenario). No other use was made of the scenarios' classifications, and they were not recorded in the minutes of the meeting.

Perform scenario evaluations
During this step, the ability of the architecture to support the five selected scenarios was discussed. The discussions were based on a description of the scenario provided by the person who first proposed the scenario, and a description of how the architecture, or a modification of the architecture, would support the scenario. All of the review participants were allowed to ask questions and participate in the discussions. After the discussions, the members of the review group were asked to summarise any issues that had been raised during the discussions. The issues were

[1] Note that since the EXC3ITE OCP/IMAD review, the proponents of the SAAM review process have modified the process, and now identify three types of scenarios (Kazman 1999). Use cases reflect the current (intended) system use; growth scenarios reflect anticipated changes; and exploratory scenarios are intended to stress the architecture.

then discussed to determine whether or not they had been adequately resolved during the discussions.

Review issues
The next step was to review the issues. In reviewing the issues, the recorder indicated the issues that had been identified and whether or not further action was required. Issues that had been successfully addressed within the meeting were also identified.

The SAAM review process as described by Bass et al (1998) does not include a review of the issues raised during the review. However, it usually contains one or two additional steps. First, a scenario interaction step can be included (Bass et al. 1998). During this step, the components that would need to be modified to address each of the scenarios are identified and compared. Direct scenarios should require no modification to the architecture. Similar indirect scenarios should result in modifications to similar components, while very different indirect scenarios should result in the modification of different components. The facilitator omitted this stage from the process. Second, a scenario-weighting step can be included (Bass et al. 1998). This step is normally included when two alternative architectures are being compared. It can be used to assist in determining which architecture is more appropriate in a given situation. This step was not included because only a single architecture was being investigated.

Wrap-up
Finally, the facilitator conducted a wrap-up session where the strengths and weaknesses of the review were discussed and summarised on a white-board. The implications of this activity on the research method are discussed in Section 4.3 and the results of this stage are summarised in Section 4.4.

3. Research objectives

The study of SAAM reviews had four objectives as given in Table 2. Using the terminology of Yin (1984), the first line in each row of the table indicates the 'research questions', while the subsequent lines indicate the 'propositions'. That is, the first line indicates the objective of the study, while the subsequent lines indicate how the research objectives will be achieved, as well as what information is required to achieve them.

Table 2: Research questions about SAAM reviews.

RQ1.	Can the performance of joint software reviews be improved through the use of the SAAM review process?
RQ2.	What are the characteristics of SAAM reviews? This will consider the SAAM voting scheme, the negotiation stumbling blocks encountered during the review, and the performance characteristics of the review.
RQ3.	How do the characteristics of SAAM reviews compare to the characteristics of traditional software reviews? This will consider the ability of SAAM reviews to align the goals of the review participants, compare the negotiation stumbling blocks encountered in SAAM reviews and traditional reviews, and consider the participants perceptions of the use of scenarios, the voting scheme and the facilitator.
RQ4.	Can the SAAM review process be improved?

Two of the objectives are concerned with comparing SAAM reviews with traditional Joint Software Reviews (RQ1 and RQ3). However, the case study also offers an opportunity to study SAAM reviews in their own right. The SAAM review process is relatively immature and it is possible that an improved SAAM review process would offer benefits over the traditional review process even if the original SAAM review process does not. Therefore, research questions RQ2 and RQ4 were also included.

The case study is exploratory rather than explanatory, and as such the research questions are general and do not lead to specific hypothesis. The development and testing of detailed hypothesis is left for later studies. The detailed results in this study could be used for comparison with the results of such studies.

The characteristics considered in this study – goal alignment, negotiation stumbling blocks and performance – are based on evidence that these factors affect the performance of software reviews eg (Macdonald et al, 1996; Kingston et al 1999c). Goal alignment captures the similarities and differences between the goals of the review participants both at the start of the review (an independent variable) and during the review (a dependent variable). The degree of goal alignment affects the negotiation stumbling blocks encountered during the review, which in turn affect the performance (or outcome) of the review. Kingston et al (1999a) developed a model showing the relationship between these factors in traditional software reviews. Figure 3 presents an adaptation of this model for SAAM reviews. The model shows the main steps in the identification of scenarios and issues during the SAAM review process, and where the presence of negotiation stumbling blocks can affect the performance of the review in terms of the number or quality of scenarios or issues raised and recorded. The model does not capture the impact of specific negotiation stumbling blocks, because their impact on the SAAM review process was not known *a priori*. The negotiation stumbling blocks considered in this report are derived from the list provided by Nunamaker et al

(1991) and include Conformance Pressure, Cognitive Inertia, Production Blocking, Domination, Communications Breakdown and Freeloading.

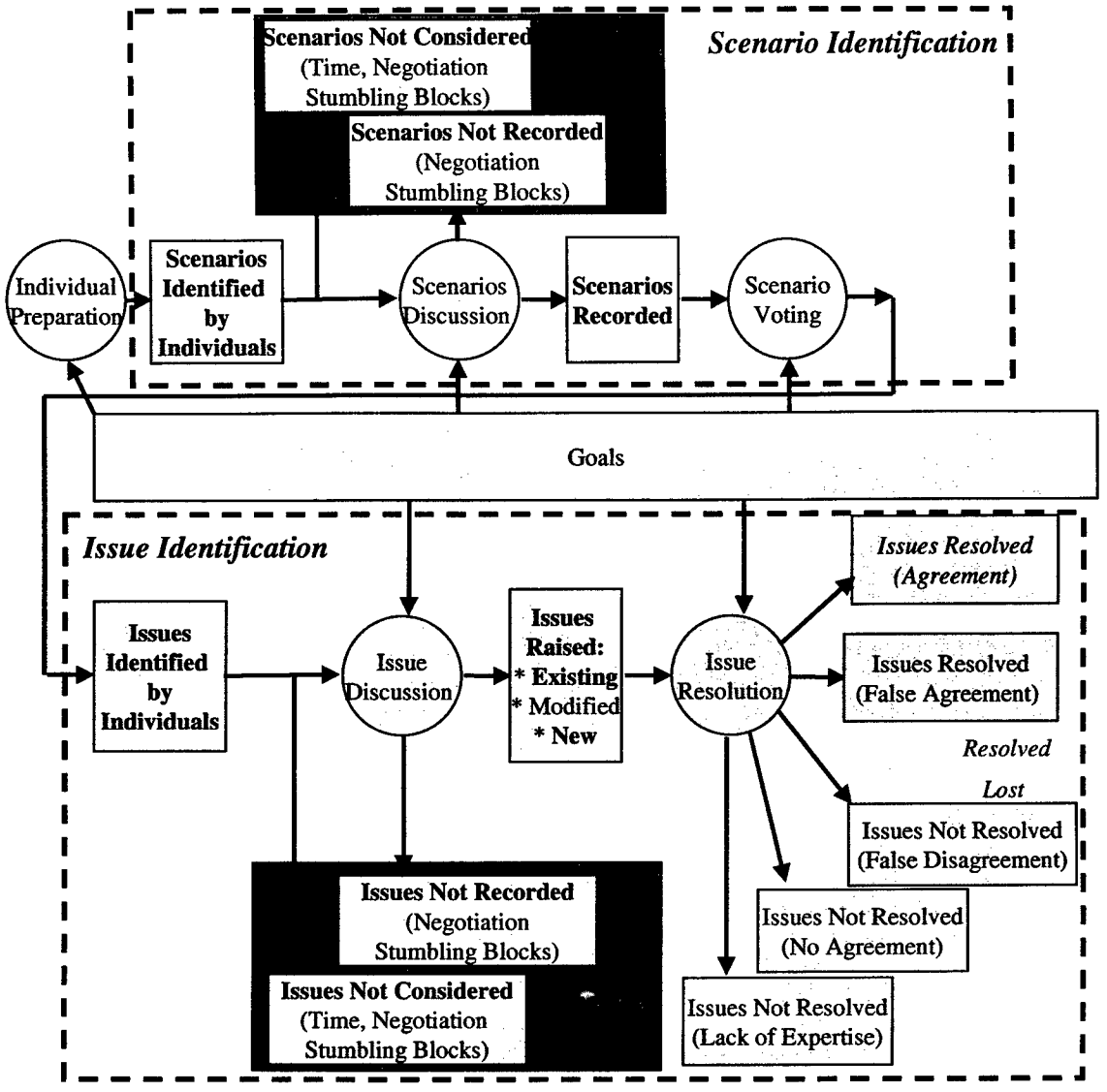


Figure 3: A model explaining how goals and negotiation stumbling blocks affect the performance of software reviews.

4. Method

The research questions presented in Table 2 were addressed by studying a SAAM review conducted on the EXC3ITE OCP/IMAD Architecture (EXC3ITE 1999). The reasons for choosing this review are discussed in Section 4.2. The results from the SAAM review were compared with the results of laboratory studies on traditional-style reviews of software designs (Kingston 1999a; Kingston et al. 1999c), and with the

results from an earlier study of the review of the architecture of Project Llama (Kingston 1999a; Kingston 1999b). Section 4.1 summarises the designs of these studies, and Section 4.3 discusses the design of the case study in more detail. Section 4.4 provides a link between the design of the study and the results. It describes the general characteristics of the review – such as the number of participants. The results are discussed in Sections 5 to 7.

The discussions in Sections 4.2 to 4.4 presuppose the use of a case study approach. However, two alternative empirical approaches were considered and ruled out. These were a survey approach and an experimental approach. A survey approach was ruled out because the ADO had not previously conducted SAAM reviews. Furthermore, there was no evidence that any SAAM reviews had been conducted in Australia, either within or outside Defence. Thus, there was nobody with experience in SAAM reviews that could have been surveyed. The case study approach also has added benefit of allowing the detailed observation and analysis of areas, which have not previously been explored in the context of SAAM– such as the voting mechanism.

The case study approach also allows more detailed observations than an experimental approach. However, there were also two practical reasons for not using an experimental approach. First, the participants were required to have a detailed understanding of their goals for the review, so that they could determine and vote on scenarios that were appropriate to their goals. This means that student participants were not appropriate for this study. Therefore, any experimental study would have had to use professional software engineers. It was believed that it would be difficult to find sufficient professional software engineers for a significant study. Second, SAAM reviews also require the use of experienced facilitators. Had an experimental approach been used, a large number of (paid) facilitators would have been required. This was not feasible given the resource constraints. Thus, a case study of the EXC3ITE OCP/IMAD Architecture was conducted.

4.1 Software Review Studies

This section summarises the characteristics and design of three studies of software reviews that provide the basis for comparing the SAAM review of the EXC3ITE OCP/IMAD architecture with traditional reviews of software designs or architectures. These studies consist of the Project Llama review (Kingston 1999a; Kingston 1999b) and two laboratory studies (Kingston 1999a; Kingston et al. 1999c).

Project Llama
Project Llama was acquired by the Australian Defence Organisation (ADO) and developed by Australian Defence Industries (ADI) as a replacement for the situation monitoring (tracking) sub-system of the Joint Command Support Environment[2] (JCSE) project, JP2030 (Hay 1997; ADO 1998; Quin-Conroy 1999). The Project Llama study

[2] The Joint Command Support Environment project has since been renamed the Joint Command Support System (JCSS) project.

investigated the properties of the traditional review process that are related to the goals of the participants. The researcher acted as a participant-observer in a team that provided input to the review (the ITD Expert Review Team), and conducted a post-review survey of review participants and a post-review interview with the person who participated in both the preparation meetings and the review meeting. The relationship between these activities is shown in Figure 4.

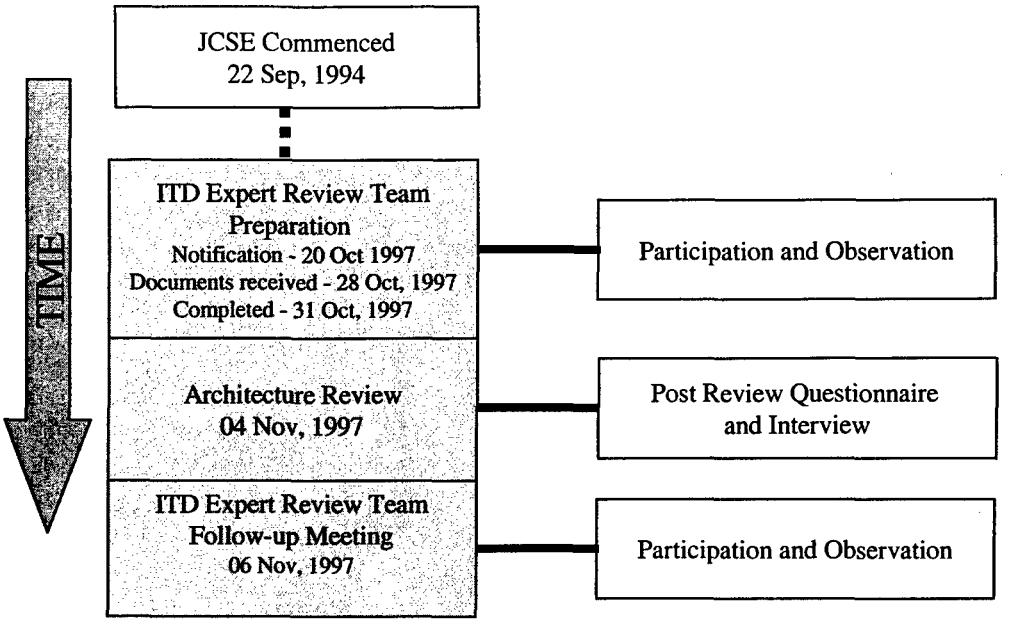


Figure 4: Project Llama time-line and data collection techniques.

The Project Llama review was selected as representative of the traditional review process because the ADI and the ADO have been conducting joint software reviews since the JCSE project's inception in 1994 (ADO 1998). The development of project Llama has since been completed. The entire system was developed within 12 months using an average of 25 developers. It consists of 62 main classes and it is smaller than the system it replaces - 128,000 lines of Java rather than 750,000 lines of Ada, Unix and X11/Motif code (Quin-Conroy 1999).

As shown in Figure 4, the Architecture Review was attended by one of the 8 members of the ITD Expert Review Team, 4 Client representatives and 10 Developer representatives. That is, 15 people from 3 organisations were present at the review. Five of those 15 responded to the questionnaire. Thus, the response rate to the questionnaire was only 33%. However, the key stakeholders (the developer and client responsible for the project) both completed and returned the questionnaire. Furthermore, one of the developers indicated that ideally there would have been fewer developers at the review. From the information on the developer's form, and the comments of the Interviewee, it appears that five of the developers attended primarily to observe the review or to answer questions if called upon. The response rate

improves to 63% if these observers are excluded from the analysis. It is argued that the low response rate poses only a limited threat to the validity of the questionnaire.

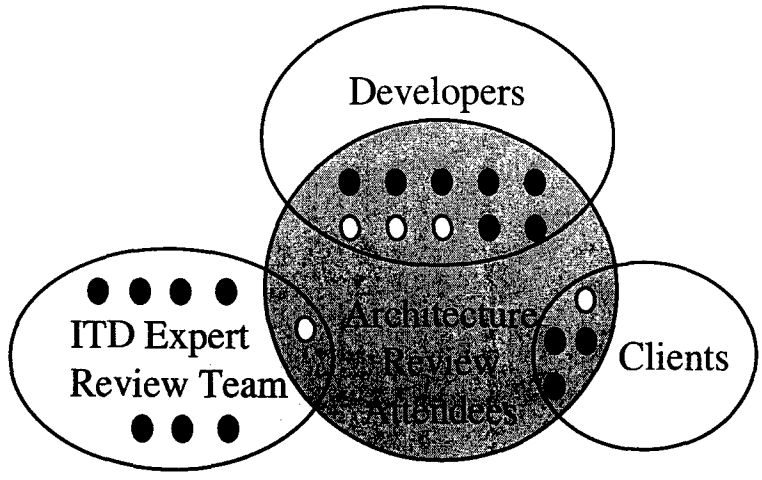


Figure 5: The relationships between the participants in the Architecture Review and the participants in the ITD Expert Review Team. (Each dot represents one person. White dots indicate participants in the Architecture Review meeting, who also responded to the questionnaire, while black dots indicate people who did not respond to the questionnaire or did not attend the meeting.)

The study identifed the participant's goals for the review and uncovered evidence of the negotiation stumbling blocks Conformance Pressure, Freeloading and Production Blocking. The detailed results, and descriptions of these negotiaton stumbling blocks, are given at appropriate points in the text. The implications of the low response rate to the questionnaire are also considered.

Laboratory Studies

The two laboratory studies were designed to investigate the impact of goal conflict on the performance of software reviews, and on the negotiation stumbling blocks encountered by review participants. The studies were based on two-person reviews of a sub-section of the design of a conference management system. They used a common between-subjects, fully randomised design with two experimental treatments and a control. The control was a software review where both participants had the same goals. The treatments were different goals (different, but not mutually satisfiable goals) and conflicting goals. The three treatments were obtained by manipulating the goals of the participants. The participants were randomly assigned to treatment groups. Data from 144 participants (72 review groups) was analysed in the studies.

The performance of the review participants was analysed using issue collection forms completed by the study participants during the reviews. The negotiation stumbling blocks Conformance Pressure and Communications Breakdowns were analysed using a post-review questionnaire. Details of the results are presented where relevant in Sections 6 and 7.

4.2 EXC3ITE OCP/IMAD architecture

The EXC3ITE project is a Capability Technology Demonstrator – a DSTO research and development project – that is being run as a formal ADO project – JP2061. EXC3ITE will provide a platform for research and development into the next generation of command support systems (ADO 1998). Thus, there could be differences between the EXC3ITE project and other projects that affect the performance of SAAM reviews. For example, the participants in the EXC3ITE review had more technical expertise and less military experience that the participants in a typical joint software review conducted by Defence. The implications of these differences are discussed further in Section 8.1.

Kitchenham et al (1995) have argued that case studies should be selected based on their representativeness. That is, a case study project should be similar to other projects of interest. However, Stake (1994) argues that case studies should be selected on the basis of the benefits they offer. The EXC3ITE project offers a unique opportunity for the study of SAAM reviews in Australia and thus meets Stake's criterion for case study selection.

Single case studies can also be used to test theories, either because they offer an extreme example of the theory (Yin 1984), or because comparisons are possible within the single case (Lee 1989). The choice of the EXC3ITE project means that considerable care is required in making comparisons. It is planned to conduct all of the joint software reviews (architecture) for EXC3ITE using the SAAM methodology. Thus, there is no opportunity to compare SAAM reviews with traditional joint software reviews within the EXC3ITE project. However, as already stated, there are some differences between the EXC3ITE project and other ADO projects. The main difference between EXC3ITE and other ADO projects in similar areas is that the EXC3ITE project is being jointly run by DSTO and the Defence Acquisition Organisation (DAO) while other projects are run exclusively by the DAO. This could result in differences that affect the ability to directly compare the EXC3ITE review with other ADO projects' reviews. Therefore, comparisons were made between the EXC3ITE review and both the Project Llama study and the laboratory studies discussed in Section 4.1. It should be noted that this approach is only intended to give an indication of the possible benefits of the SAAM review process and, where necessary, the limitations of the comparisons will be clearly spelt out.

One objective of this case study is to characterise the SAAM review process. Basili (1986) claims that characterisation is a valid purpose for an empirical software engineering study. This is also supported by Yin (1984). Where possible, the characteristics used in previous studies of software reviews are used for consistency. However, differences in the processes used mean that some changes were necessary to the definitions of the measures used in the analysis. These changes are noted at the appropriate points in the text. Furthermore, the ability to directly observe the EXC3ITE review means that additional characteristics could be observed and alternative methods for characterising the review process could be used.

The SAAM review studied was the first of a series of planned reviews for the software architecture for EXC3ITE. The EXC3ITE architecture should support the integration of several concept demonstrators (systems) developed by DSTO. It is planned to base the EXC3ITE architecture around a series of object-oriented components that are part of a product – Object Computing Platform (OCP) – supplied by Object-Oriented Proprietary Limited (OOPL). To test this approach, a single product – an image management system called IMAD – was chosen for integration with a subset of the OCP. This combined system was intended as a proof of concept for the approach to be taken with the EXC3ITE architecture. Thus, the review focused not just on the current architecture, but also its ability to allow the integration of other systems and capabilities in the future.

4.3 Design

Case study information can be collected from a variety of sources such as: the documents and artefacts produced by an activity; observational techniques; and questionnaires (Lethbridge et al. 1998). Kaplan and Duchon (1988) claim that using both qualitative and quantitative information is beneficial. Using multiple sources of information to address the same research questions, within the same case study, is called triangulation (Stake 1994). Stake (1994) argues that results supported by triangulation information are more believable than results that are obtained from a single information source because the different sources of data have different potential biases. For example, self-reported evidence from the participants might be biased if the participants cannot accurately recall information or if they wish to make an impression on the researcher through their responses. Evidence from the researcher's observations could be biased according to the outcome expected by the researcher. Thus, the approach taken in the Case Study was to obtain information from a variety of sources. These sources were 1) a pre-review questionnaire, 2) a post-review questionnaire, 3) observation of the review and 4) the minutes for the review. Figure 6 shows how these information sources relate to the review.

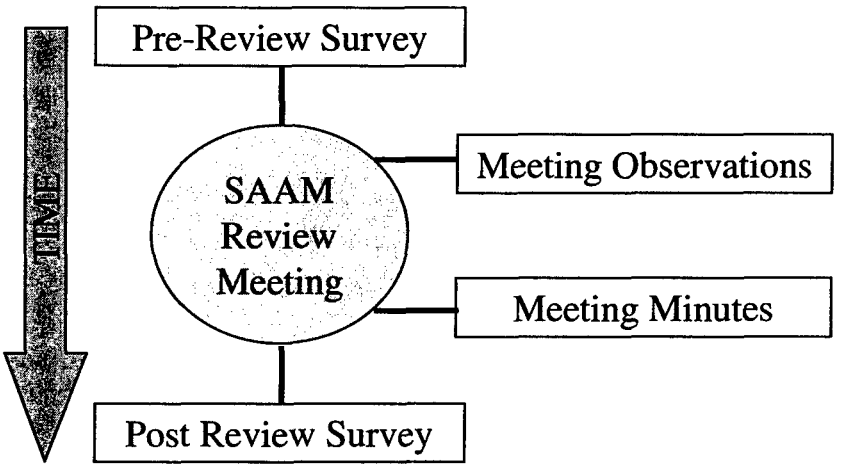


Figure 6: Data collection activities performed before, during and after the SAAM review.

Questionnaires

The pre-review questionnaire (Appendix A) was administered immediately before the review and the post-review questionnaire (Appendix B) immediately after the review. Ellis (1994) identifies four sources of inaccurate information from self-assessments: 1) failure to understand a question, 2) failure to recall the answer to a question, 3) indecision, and 4) dishonesty. The first potential problem was addressed by having the questionnaire reviewed prior to the study. The researcher was also available to answer questions while the questionnaires were completed. The second potential problem, failure to recall an answer, usually occurs when there is a substantial time lag between the events that the subject is being asked to recall and the completion of the questionnaire. Because the questionnaires were completed immediately before and after the review, any limitations due to failure to recall an answer were minimised. The third potential problem occurs when people have mixed feelings about an issue, or change their mind about an issue during the course of the study. According to Ellis, "There is nothing you can do about these opinion changes, but it is important to be aware that they occur". In this survey, the subjects had the option of replying with a neutral response. The fourth potential problem occurs when subjects "lean towards responses considered the most 'desirable' or least embarrassing" (Ellis 1994). The survey questions were designed to reduce the impact of the fourth potential problem: questions were carefully worded, multiple questions (some phrased in the negative) were used for most topics to improve the reliability of the responses, and related questions were distributed throughout the questionnaire. There is a small possibility that the inclusion of the 'wrap-up' stage in the review process could have affected the results of the post-review survey. However, the 'wrap-up' session focused on the identification of strengths and weaknesses of the process, while the questionnaire asked more specific questions about the process and the review outcomes.

The design of the pre-review questionnaire, and part A of the post-review questionnaire are discussed at relevant points of the text. Part D of the post-review questionnaire consisted of two open-ended questions designed to catch any additional information that the participants wished to provide. The other parts of the post-review questionnaire consisted of statements. The study participants had to decide to what degree they agreed or disagreed with the statement. The same response options were used for all the questions. The subjects could select one of five values (1-5), where 1 means Strongly Agree and 5 means Strongly Disagree. This is a circled end-anchored response option (Ellis 1994).

It is similar to the Likert scale (Likert 1932), but does not anchor each of the responses[3]. That is, it uses a limited set of ordered responses where meanings are assigned to only the first and last responses. One advantage of both the Likert scale and circled end-anchored response options is that they are easy to code, and the possible responses are known in advance (Ellis 1994). Likert-type scale has been used in other

[3] In the Likert scale, each response is associated with a phrase rather than a number. The usual response options are: Strongly Agree, Agree, Neutral, Disagree and Strongly Disagree

Software Engineering surveys. For example, Chung and Guinan used a seven point Likert scale in their study of participative management (Chung and Guinan 1994). However, the Likert scale has a disadvantage over end-anchored response options: the Likert scale is more subjective (Ellis 1994). All of the response options need to be interpreted, while only the extreme options need to be interpreted with end-anchored response options. Consequently, responses generated using the Likert scale will only be ordinal. However, it is reasonable to assume that the end-anchored response option provides a set of equally spaced options, because integers are used for all of the options and meanings are not assigned to the intermediate options. Thus, the results from an end-anchored response option can be treated as if they are of an interval scale (Ellis 1994; Fenton and Pfleeger 1996). That is, the mean (or average) of the values can be calculated, and standard statistical tests, such as the Students t-test, can be used on the results. In contrast, analysis of data collected using the Likert scale should be restricted to calculation of the median – the middle value when the results are ordered. To avoid these limitations, the questions used end-anchored response options on a scale of 1 to 5. The results of the questionnaire were converted before analysis. The results from negative questions were converted to reflect the degree to which a property *was* held and all the results were converted to a scale of –2 to +2. A result of +2 indicates strong agreement that a property was present while a result of –2 indicates strong disagreement.

Minutes

The meeting records were composed during the review, and displayed using an overhead projector during the final stages of the review. The final minutes prepared by the recorder were used to determine the number of scenarios that were raised, the number of issues that were raised, and the number of issues that were resolved in the review.

Observations

The observations have the most potential to introduce biases due to the researcher's preconceptions. Two methods were used to reduce potential biases. First, the information to be recorded was clearly identified before the review and a database was designed to enable the information to be readily recorded during the review. The information was based on verbal discussions, material written on whiteboards, and indications of approval and disapproval – such as nodding of heads. Verbal communications were characterised by the time of the communication activity, the person speaking and the type of communication – eg social, discussion or the statement of an issue or a scenario. More details of the classification scheme are described in Appendix G. Second, an independent research observer also attended the review. Adler (1994) claims that the age and gender of the observer can affect how they record events. Thus, the second observer was of a different gender to the researcher, and there was a significant difference in the age of the researcher and the age of the second observer. The second observer also had a technical (workshop) rather than a professional (academic) background. The review was also recorded on video tape, so that any differences in the two sets of observations could be resolved. The research made no other use of the video tape.

Triangulation

The four sources of information were used to address three different research areas: 1) the interaction between the scenario approach and the participants' goals for the review, 2) the performance characteristics of the review, and 3) the negotiation stumbling blocks encountered during the review. However, not all of the four information sources were equally valuable for addressing each research area. Figure 7 shows the relationship between the information sources and the research areas that they were used to address.

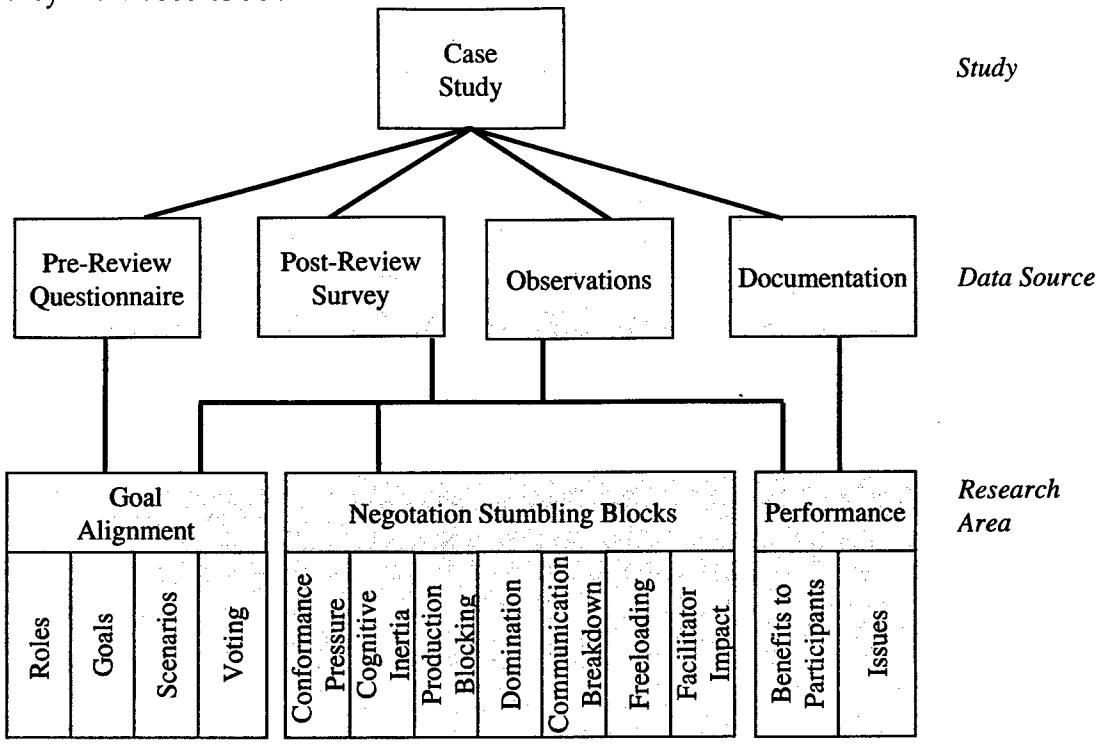


Figure 7: Data collection sources and related research areas.

Details of how the information sources were used to address each of the research areas are discussed in Sections 5 to 7. However, to interpret the results in each of those sections, an understanding of the review process and of the general characteristics of the review is required.

4.4 Review characteristics

The EXC3ITE OCP/IMAD review was the first SAAM review conducted by the ADO. Therefore, the organisers of the review made a number of decisions that affect the external validity of the study. For example, the organisers of the review deliberately chose to conduct a short review of an architecture that was designed as a 'proof of concept' for a larger architecture. The review was conducted over 4 hours, including a 15-minute break. Consequently, the review had severe time constraints and insufficient time was available for some activities. The impact of the time constraints was compounded by the participants' lack of prior experience with SAAM reviews. The

facilitator had no prior training in SAAM reviews and only three of the participants had any training in SAAM reviews. Other characteristics of the review and their effect on the validity of the study are now discussed.

Participants

Eleven people were invited to participate in the review. They came from three organisations: nine participants were invited from DSTO, one participant was invited from OOPL and an independent facilitator was also invited. Three of the people were given specific roles within the review. One was the independent facilitator, one person from DSTO was invited to act as the recorder, and the person from OOPL was invited because he was the prime architect of the product under review. Seven of the remaining eight participants were reviewers invited by the client because they were stakeholders in the product under review. The final participant invited from DSTO was one of the review organisers. He was planning to observe the review. The review planners envisaged that the stakeholders would have different interests and different roles. These are given in Table 5, Section 5. Most of these roles reflect the current positions of the people invited to participate in the review. However, one person – with the appropriate domain knowledge – was asked to assume the role of the systems manager for EXC3ITE.

Twelve people actually attended the review. One stakeholder with an interest in Capability Technology Demonstrators (CTDs[4]) that would use the EXC3ITE infrastructure was unable to attend the review meeting. Instead, he sent two proxies. Unfortunately, the role of the proxies was not clear either to them, or to the facilitator. Thus, they both acted as observers rather than reviewers. A total of three participants attended the review as observers. Where necessary, they will be called review observers to distinguish them from the research observers.

All twelve of the review participants completed both the pre-review questionnaire and the post-review questionnaire giving a response rate of 100% for both questionnaires. Thus, there are no threats to the internal validity of the study due to low response rates.

Environment

There are threats to the validity of any case study from the context in which the study is conducted. Software architectures are reviewed against a particular purpose and often against a set of requirements. In modern software development and acquisition, the requirements often undergo constant evolution. Therefore, it is important to characterise the state of the requirements before the review, so that the results of the study can be correctly interpreted. As shown in Table 3, the participants believed (mean less than 0) that the requirements were poorly documented and not understood before the review. Thus, the results of this study will be most applicable to other studies where the requirements are not understood before the review.

[4] DSTO plans to develop a number of systems (CTDs) that will build on the EXC3ITE infrastructure.

Table 3: Quality of requirements documentation as perceived by the review participants.

Variable	Question	Mean	Number of Responses	Standard Deviation
R	Requirements were well documented and I understood them prior to the review.	-0.9	12	0.8

Perceptions

The review participants were given two opportunities to comment on their impressions of the review. The facilitator asked the participants to consider the strengths and weaknesses of the review, and the post-review questionnaire asked the participants to consider how the review might be improved, and how it compared with traditional software reviews. Some of the questions concerned specific research questions and are discussed in the relevant areas of the results. This section discusses the results from the more general questions in the post-review questionnaire and the opinions expressed during the review.

Most of the limitations and the improvements suggested concerned the limited time available to discuss the scenarios. The participants believed that insufficient time was spent on scenario elaboration. Some participants also believed that the most important scenarios were not selected. This is discussed further when the voting scheme is discussed in Section 5.5. However, the participants who had also been involved in traditional joint software reviews still believed (mean greater than 0 in Table 4) that SAAM reviews were better than traditional joint software reviews.

Table 4: Comparison of SAAM reviews with traditional joint software reviews as perceived by review participants.

Variable[5]	Question or Calculation	Mean	Number of Responses	Standard Deviation
$-E_1$	SAAM reviews offered no benefit over traditional joint software reviews.	0.8	6	0.4
E_2	SAAM reviews are better than traditional joint software reviews because: SAAM reviews are more structured than ordinary reviews.	1.2	6	0.9
E	(E2 – E1)/2	1.0	6	0.4

5. Goals, roles and scenarios

[5] Negative variables indicate that the responses to the *inverse* of the questions were analysed. Thus a mean response of 0.8 to –E1, means that the mean response to the statement was –0.8. Mean responses to a statement that are greater than 0 indicate agreement with a statement. The initial response of –0.8 indicates that the participants disagreed with the statement.

This section discusses the main characteristics of SAAM reviews – the use of a facilitator and scenarios – and their ability to align the goals of the review participants. Section 5.1 discusses the roles of the participants. It compares the roles anticipated by the review organisers with the stated roles of the participants and the observers. The role of the facilitator is discussed in more detail in Section 5.2. Section 5.3 describes the goals of the participants and discusses apparent changes in the goals of the participants. These changes could be due to the use of scenarios. Scenario generation is discussed in Section 5.4, while Section 5.5 discusses the selection of scenarios and their perceived strengths and weaknesses. The discussions identify the information sources, the specific variables, and the questions used to investigate each area, as well as the results of the study.

5.1 Roles

Information about the participants' roles in the review was taken from two sources. First, the documentation supplied for the review contained a list of the review participants and their anticipated roles and interests. The review organisers generated the list and used it while planning the review. Second, the pre-review questionnaire contained a list of roles that the participants could have in the review. The list included functional roles – such as observer, reviewer and recorder as well as roles related to specific interests – such as the development of CTDs. The roles in the list were derived from the roles used by the review organisers. These included the roles planned for the EXC3ITE OCP/IMAD review and roles that the participants might have in future EXC3ITE reviews. Additional roles were also included after consultation with the review organisers. The participants were asked to indicate all of the listed roles that they were taking in the review. The list is given in the post-review questionnaire in Appendix G.

The information about the participants' roles was collected to characterise the review. It was anticipated that the participants' stated roles would be similar to those planned by the review organisers. However, there were three differences worth discussing further (Table 5). First, one of the participants was not able to attend the review. This participant and the main review organiser (the project manager) had agreed that a replacement would take the participant's role as a reviewer and a stakeholder interested in CTD development. However, neither of the replacements acted as a reviewer – both acted as observers. Second, several of the participants – including the observers just discussed – did not believe that they were attending the meeting as reviewers. Third, many of the participants believed that they had a role as a research user or a technical expert in addition to the role anticipated by the review organisers. It is believed that this affected the evaluation of the scenarios. Much of the discussion was focused at the technical level rather than at the level of user-requirements intended by the review organisers. The performance characteristics of the review are discussed further in Section 6.

Table 5: Functional roles and specific interests as anticipated by the review organisers and as stated by the review participants.

Participant	Functional Role		Specific Interest	
	Expected and Documented by Review Organisers	**Stated by Participant**	**Expected and Documented by Review Organisers**	**Stated by Participant**
1	Architect / Presenter	Architect / Presenter		
2	Facilitator	Facilitator		
3	Recorder	Recorder		
4	Reviewer	Reviewer	EXC3ITE System Manager	Alternative interests stated: • EXC3ITE Developer • Research User / Technical Expert
5	Reviewer	None stated	EXC3ITE Project Manager	EXC3ITE Project Manager
6	Reviewer	Reviewer	Research User / Technical Expert	None Stated
7	Reviewer	None stated	Research User / Technical Expert	Research User / Technical Expert
8	Reviewer	None stated	CTD Developer	CTD Developer
	Reviewer	Did not attend	CTD Developer	Did not attend
9	Reviewer	None stated	EXC3ITE Developer	EXC3ITE Developer Also stated: • Research User / Technical Expert
10	Observer	None stated		Also stated: • Research User / Technical Expert
11	Observer	None stated		Also stated: • Research User / Technical Expert
12	Observer	Observer		

5.2 Facilitator

The facilitator had an important role in the EXC3ITE OCP/IMAD review. He had to direct the review according to the SAAM review process. He was also asked to encourage participation from all the reviewers and to control the recording of issues. That is, he had to determine whether or not the participants were satisfied with the descriptions of issues before they were recorded. After issues were recorded, the facilitator had to negotiate a course of action to address each issue, which satisfied all of the participants.

The impact of the facilitator was evaluated through observations of the review and through the post-review questionnaire. The research observers recorded the following information: when the facilitator solicited input, when the facilitator encouraged more discussion, when the facilitator suggested alternatives, when the facilitator stopped one person from dominating discussions, when the facilitator stopped people from socialising and when the facilitator changed the topic of the discussion.

From the observations, it appeared that the facilitator was more involved during the early phases of the review process than during the later phases of the review process. This might be due, in part, to the time constraints imposed on the review or it might be due to the facilitator's lack of experience with the SAAM review process.

The facilitator encouraged participation from all of the participants during the early phases. During the overview, the facilitator asked each of the participants to introduce themselves and to contribute two or three items to a needs analysis session. The facilitator encouraged participation from all of the participants during the scenario generation phase. He asked each reviewer in turn to state one or two scenarios and then solicited more contributions from the reviewers. During this stage, the facilitator indicated who was to state the next scenario, thus encouraging participation from all of the reviewers and preventing one reviewer from dominating the scenario generation process. The facilitator also indicated who was to vote next during the voting phase, although he did miss one of the reviewers during the first round. (A more detailed discussion of the voting process is given in Section 5.5.)

The contribution of the facilitator was less obvious during the other phases of the review, particularly during the scenario elaboration phase and the scenario evaluation phase. During the scenario elaboration phase, the facilitator unsuccessfully requested general input from the reviewers. It is believed that the facilitator needs to be trained in a more detailed elaboration process. The process should address the problems with the scenario elaboration phase, which are described in more detail in Section 5.5. The facilitator's role would then include ensuring that the more detailed process steps were followed.

During the scenario evaluation phase, 15 minutes were allowed for the discussion of each scenario. Each scenario was described by its proposer before the architect discussed the changes to the architecture (if any) required to handle the scenario. The facilitator kept the presenters informed of how much time they had remaining. Other review participants were allowed to contribute to discussions during the presentations. After the presentations, the facilitator asked the reviewers to indicate any issues that had been identified during the discussions. Sometimes, the facilitator also suggested issues that he had identified during the discussions. However, this only occurred after one of the observers indicated that there were a number of issues arising from the first discussion that had not been recorded. The facilitator then sought general agreement that the issues had been satisfactorily met by the discussions. He often asked 'Is this [issue] okay?' rather than more probing questions. The facilitator did not seek

individual responses and in many cases accepted a lack of response as an indication that the issues had been satisfactorily addressed. Furthermore, the facilitator did not endeavour to determine what further action – if any – was required. Consequently, actions were identified for only a few of the issues during the meeting, and many of the resolutions and actions recorded in the minutes were not discussed during the review meeting (see also Section 6). Therefore, the meeting participants might not agree with, or be satisfied with, the resolutions or actions. It is suggested that the facilitator should have spent more time encouraging the participants to try to determine how to address each issue and ensuring that all of the participants agreed with the chosen course of action.

Despite this, the general perception (mean greater than 0 in Table 6) of the review participants was that the use of a facilitator offered benefits over the traditional review process. One of the respondents to the post-review questionnaire indicated that there was very little conflict or domination in the review, not because of the facilitator, but because of the nature of the participants. However, as previously discussed, there was evidence from the observations as well as the post-review questionnaire (F_2 and F_3) that the facilitator took steps to ensure that all of the participants had an opportunity to contribute to the review – particularly during the generation of scenarios. There was also evidence that the facilitator helped prevent conflict from escalating, as well as preventing one person from dominating the discussion during the early stages of the review.

Table 6: Participants' perceptions of the impact of the facilitator.

Variable	Question	Mean	Number of Responses	Standard Deviation
SAAM reviews are better than traditional software reviews because[6]:				
F_1	The facilitator was able to prevent conflict from escalating.	0.5	6	0.5
F_2	The use of the facilitator prevented one person from dominating the discussion.	0.7	6	0.8
F_3	The use of the facilitator ensured that everyone's goals were considered when each issue was discussed.	0.8	6	0.7
F	$F_1 + F_2 + F_3 / 3$	0.7	6	0.6

In interpreting these results, it should be remembered that the facilitator had only general experience as a facilitator and that he did not have any prior exposure to the SAAM review process. The facilitator performed best at the SAAM review activities that were similar to more general negotiation or meeting activities – such as brainstorming. The facilitator did not perform as well at SAAM specific activities –

[6] While this might be considered a leading question when considered in isolation, the question was asked in conjunction with question E_1 of Table 4. Question E_1 states that "SAAM reviews offer no benefit over traditional software reviews". The participants were allowed to respond to either of the questions. (See Appendix G for the layout of the Questionnaire.)

such as scenario elaboration and issue resolution. This suggests that facilitators should receive specific training in the SAAM review process. It is believed that the development of guidelines that indicate what is expected from the facilitator during the various phases of the review process would also be beneficial.

5.3 Goals

The pre-review questionnaire and the post-review questionnaire were used to collect information about the participants' goals. This information provides part of the characterisation of the EXC3ITE OCP/IMAD review. The ability of the SAAM review process to align the goals of the participants is discussed in Section 5.5.

The questionnaires listed standard goals, to enable consistent coding of the participants' goals, and provided space for the participants to identify up to six other goals, in case the participants had additional goals that were not in the list. The standard goals were derived from one of two sources. First, some of the goals were derived from the goals identified during the Project Llama study. Second, some of the goals were derived from the goals used in the laboratory studies discussed in Section 4. These goals were adapted to reflect the nature of the EXC3ITE OCP/IMAD system rather than the design used in the laboratory studies. The review organisers were asked to check the list to determine whether or not any additional goals should be included. No major changes were made to the list of goals during this final step.

The pre-review questionnaire was used to determine whether the participants had the same goals or whether their goals were different or conflicting. The participants were asked to indicate and rank their goals. Numbers were not to be repeated during the ranking. Six participants did not follow these instructions. The goals were provided in three columns, and three participants ranked the goals in each column separately, one made an error and repeated a number, and two provided only a single tick. Therefore, the ranking was not used during the analysis. Instead, the analysis was based on the number of participants who indicated that they had a particular goal. As shown in Table 7, most of the goals of the participants were covered in the list provided, eliminating the need to verify the coding of the goals. One participant stated that he had most of (15 out of 16) of the listed goals, while 7 of the 10 participants had only around six to eight goals of particular interest, and two of the participants were interested in less than 3 goals. Six of the goals were held by at least half of the review participants. In contrast, three of the goals were held by less than three participants. Thus, the participants had some common goals, but they also had different goals.

Table 7: Participants' stated goals for the review. Shading indicates the goals that were selected by participants. Numbers indicate that multiple (other) goals were recorded.

Goal	Participants with goal												
	Total Times Selected	Facil-itator	Archi-tect	Recor-der	Reviewers						Observers		
		1	2	3	4	5	6	7	8	9	10	11	12
Listed goals													
Performance	4					■		■		■			■
Requirements Met	6	■	■		■	■		■				■	
Support for Experiment	8	■			■	■	■	■	■	■			■
Minimise IMAD changes	2					■	■						
Minimise OCP changes	1						■						
Flexibility	8	■	■		■	■	■	■	■			■	
Documentation Quality	1					■							
Future Supportability	6		■		■	■	■	■	■				
Proof of OCP Quality	6		■		■	■		■	■			■	
Risk Assessment	3					■			■	■			
Cost/Benefit alternatives	3					■	■			■			
Verify architecture	9	■	■		■	■	■	■	■	■		■	
Secure agreement	5	■	■		■	■		■					
Present information	1					■							
Address scenarios and issues	4	■			■	■							■
Other goals	6			■			■				■		■

The post-review questionnaire was designed to determine the degree to which the participants' goals were met. The participants were asked to state whether each of their goals was met Completely, Partially or Not at all. Figure 8 shows the total number of goals in each category as calculated from the post-review questionnaire, summing the responses of all the participants over all the listed goals. The participants in the SAAM review appear to have been less successful in meeting their goals than the participants in the study of the Project Llama review (Figure 9). There are two alternative explanations for these differences. First, the questionnaire used in the study of Project Llama only had a 33% response rate. It is possible that the goals of the other review participants were not met. Second, the two studies used different methods to capture whether or not the participants' goals were met. The participants in the SAAM review had to indicate whether or not *their* goals were met, using a list of possible review goals. The participants in the Project Llama review were not provided with a list of goals, because of the exploratory nature of the study. Instead, they had to identify their own goals as well as indicating whether or not they were met. Thus, it is possible that the participants only remembered the goals that had been at least partially met. Additional studies are needed to determine which, if any, of the two interpretations is correct.

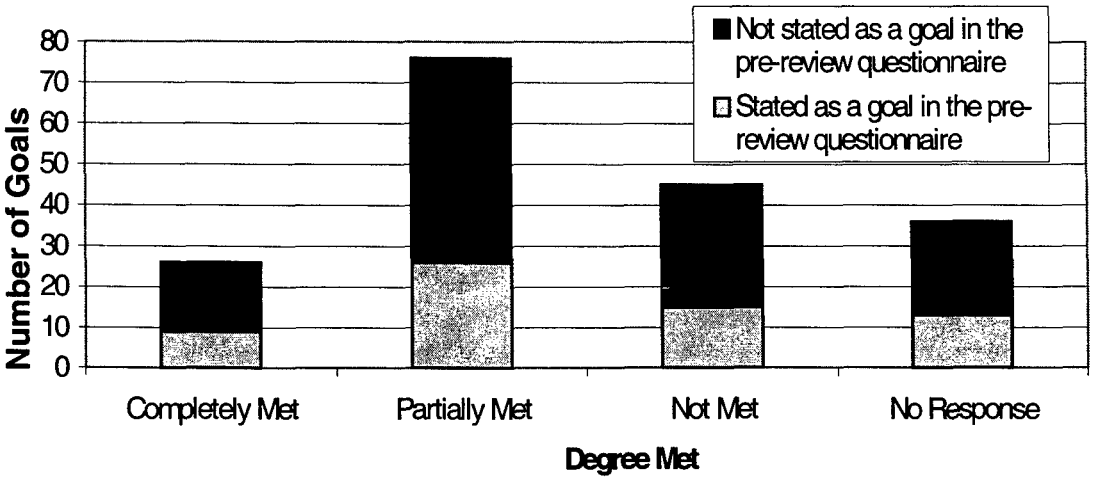


Figure 8: Degree to which the participants' goals were met by the review as stated in the post-review questionnaire for the EXC3ITE IMAD/OCP Review.

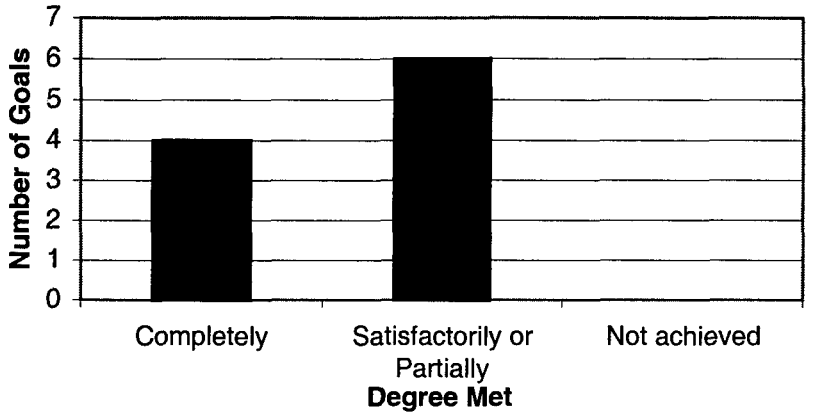


Figure 9: Degree to which the participants' goals were met during the Project Llama review.

Figure 8 not only indicates the degree to which the participants' goals were met, but also whether or not (see the index to the figure) the participants had indicated that they had the goals in the pre-review questionnaire. The participants were not explicitly asked if their goals had changed during the review. However, from the results in Figure 6 it appears that the participants' goals had changed. The participants had additional goals – the first three black bars in Figure 6 show that goals rated by the participants included goals that they had not held at the time of the pre-review questionnaire. However, the participants did not rank all of the goals, as indicated by the right-most bars. Therefore, it appears that these goals were additional goals. Furthermore, the participants failed to rate some of the goals that they held during the pre-review questionnaire (the fourth light bar). It appears that the participants abandoned these goals. Changes to the participants' goals were also discussed in Section 5.2.

5.4 Scenario generation

The scenario generation process used by the EXC3ITE OCP/IMAD review team was a two-step process. The first step was a brainstorming session where scenarios were identified and recorded on the whiteboard. The second step was an elaboration step where scenarios were clarified, and similar scenarios were combined or discarded.

Two sources of information were used to characterise the scenario generation process. First, information was obtained from observing the scenario generation process. The research observers recorded each time a scenario was stated or modified, and each time two or more scenarios were combined. Second, information was obtained from the review documentation and artefacts. Copies of the information on the whiteboard were obtained at the end of each phase. These were used to determine the number of scenarios that were removed, and as a second source of information on the number of scenarios that were combined. The number of scenarios on the final whiteboard was compared with the number of scenarios in recorder's database and with the order of the scenarios in the minutes of the review to check for consistency.

The results of the study are given in Table 8. The brainstorming session appears to have been reasonably efficient; 35 scenarios were identified in less than 25 minutes. Each of the reviewers identified at least 3 scenarios. As expected, the review observers, the architect, the facilitator and the recorder did not identify any of the scenarios. There was no process loss[7] during this stage of the meeting as all 35 scenarios were recorded.

Table 8: Observed characteristics of the scenario generation phase.

Category	**Number of scenarios**
Brainstorming Step (25 min)	
Scenarios Stated	35
Scenarios Recorded	35
Scenarios Modified	3
Elaboration Step (5 min)	
Scenarios Combined	0
Final number of scenarios	35

The elaboration step was less successful. The facilitator asked the participants to identify scenarios that could be grouped or combined. When there was no constructive input from the review participants after five minutes, the facilitator commenced the next stage in the review. It is believed that the elaboration process could be improved. One of the participants remarked that two of the selected scenarios were almost identical during the scenario evaluation phase. However, this similarity was not noticed until one of the scenarios was clarified to provide the architect with sufficient

[7] Process Loss occurs when the review process does not ensure that all the scenarios or issues are recorded. It includes scenarios or issues that are "lost" because the participants do not have time to raise them during the meeting.

information for him to discuss it. Further analysis by two of the review organisers revealed other groups of similar scenarios.

Two mechanisms for grouping scenarios were identified. First, existing scenarios could be combined to form a more complex scenario. For example, the three scenarios 'A server starts up', 'A client starts up' and 'A server shuts down' could be combined into one scenario where all three events occur in the order given. Second, a representative scenario could be selected from a set of similar scenarios. This is relevant when the scenarios are all specific methods of evaluating the same underlying concern. For example, the two scenarios 'We want to update to Java 1.2' and 'We want to update to the latest version of OCP' are both specialisations of the same concern – 'How do we deal with updates to technology?' If the two scenarios are related in this manner, it might be possible to remove one from further consideration. This will prevent votes being split between similar scenarios. (See also Section 5.5).

It is believed that it would be difficult to use these grouping mechanisms without modifying the elaboration phase. First, the participants need more instruction in the grouping mechanisms. They were not given any instruction in the EXC3ITE OCP/IMAD review, and they were not told why it was important to group scenarios. Second, it is believed that more detailed descriptions of the scenarios are probably required to do effective comparisons of scenarios. Elaborating scenarios is time consuming, thus mechanisms for elaborating and grouping scenarios need to be investigated. One possible mechanism is to randomly select a scenario and have the person who suggested it describe it in more detail. The other participants can then decide whether or not the scenario is similar to other scenarios that have been identified. After similar scenarios have been identified and grouped, another scenario can be selected.

5.5 Scenario selection

After generating scenarios, the reviewers selected 5 scenarios[8] for further analysis using the voting process described in Section 5.5. The research observers recorded the allocation of votes by the participants. The participants' perceptions on the benefits and limitations of the combination of scenarios and the voting scheme were also analysed using the post-review questionnaire and the white-board summary of the strengths and weaknesses of the review. The results of this analysis are summarised in Table 9 and in Figure 10.

[8] The review organiser chose the number of scenarios to be selected based on the available time for the review. The organisers who had attended the training session on the SAAM review process believed that approximately 15 minutes would be required to discuss each scenario.

Table 9: Voting patterns over two rounds. Shading indicates participants who raised or modified scenarios.

Scenario	Votes by Reviewer and Round						Total Votes		
	A	B	C	D	E	F	Research observers		White-board
	1st/ 10th	2nd/9th	4th/8th	3rd/7th	5th/6th	11th	First Round	Total	
1				2/0			2	2	2
2					2/5		2	7	7
3	**2/8**						**2**	**10**	**10**
4		6/0		0/2			6	8	8
5			3/6				3	9	9
6				3/4			3	7	7
9	**2/0**		**3/0**			**0/5**	**5**	**10**	**10**
11	**2/0**				**2/2**	**0/5**	**4**	**11**	**10**
14				2/0			2	2	2
15	**0/2**	**2/2**	**0/4**				**2**	**10**	**10**
18					6/3		6	9	8
19				3/4			3	7	7
21	2/0						2	2	2
24	2/0		3/0				5	5	5
26			1/0				1	1	1
29		**2/8**					**2**	**10**	**10**
TOTAL	**10/10**	**10/10**	**10/10**	**10/10**	**10/10**	**0/10**	**50**	**110**	**108**

The first column of Table 9 lists the scenarios; only those scenarios that received at least one vote are listed. The next six columns represent the reviewers, A, B, C, D, E and F. The figures in the columns represent the number of votes they allocated to each scenario in each of the two rounds of voting. Thus, the figure of 2/8 in the second column indicates that reviewer A allocated 2 votes to scenario 3 in the first round and 8 votes to scenario 3 in the second round. The final row indicates that reviewer A allocated 10 votes in the first round and 10 votes in the second round. The figure 1st/ 10th in the third row indicates that reviewer A had the first vote (round 1) and the tenth vote (round 2). The last three columns summarise the results of the voting process. The first two of these columns indicate the results as recorded by the research observers and correspond to the total of the figures in the previous 6 columns. The final column is provided for comparison; it contains the figures recorded by the facilitator, and used to select the scenarios.

Figure 10 contains an alternative representation of the information in Table 9. The blocks indicate votes. The shading of a block indicates the participant who voted, the vertical position of a block indicates the scenario that they voted for. The horizontal position indicates the relative time at which the votes were given. Thus, participant A distributed the first ten votes evenly between scenarios 3, 9, 11, 21 and 24. The vertical bar separates the first round from the second round.

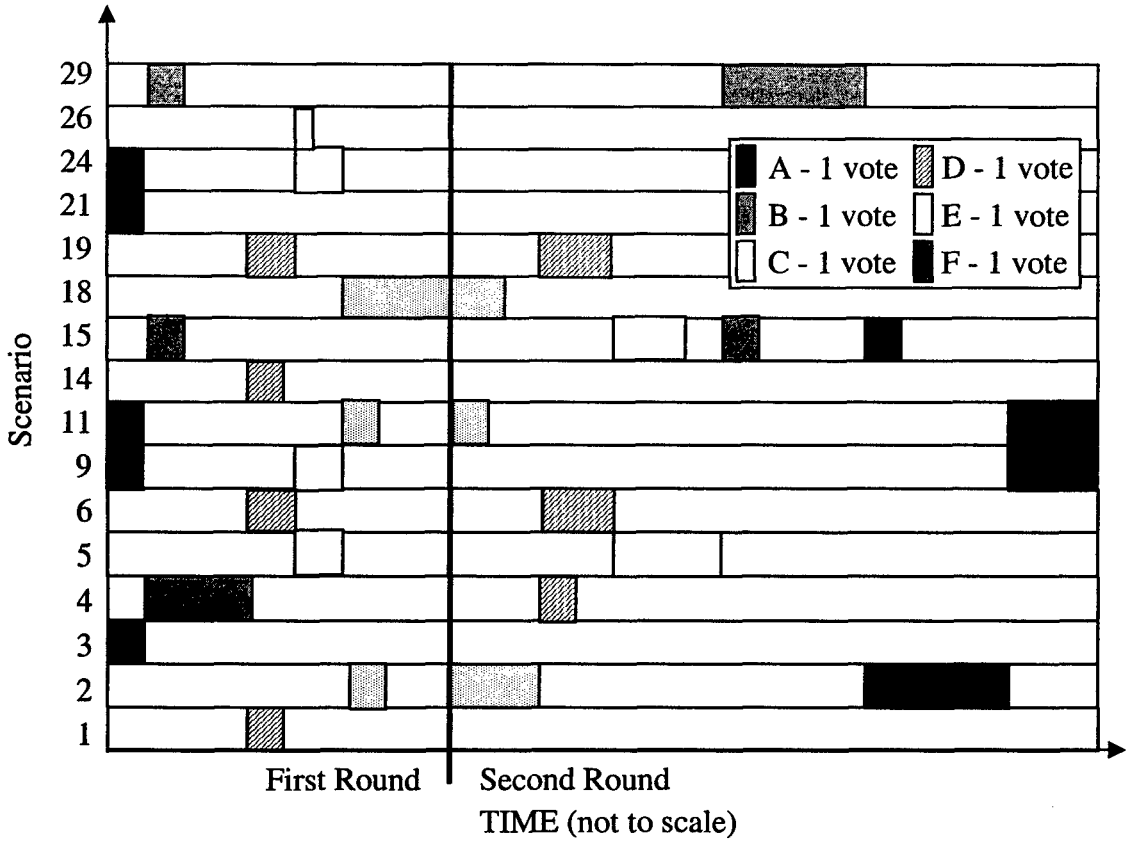


Figure 10: An alternative representation of voting patterns.

Three problems with the voting process were identified. First, one participant complained that the voting process was not fully explained before the start of the voting process. The facilitator only informed the participants that there would be two rounds of voting after the first round of voting had occurred. The second problem is that one of the reviewers was only given a vote in the second round (Table 9). These problems could both be due to the demands made on the facilitator, who had no prior experienced with SAAM. However, the third potential problem with the voting scheme was not due to the facilitator. The voting process used in the EXC3ITE OCP/IMAD review was a public voting scheme proposed by two of the meeting organisers after they attended a seminar on SAAM. The facilitator implemented the voting scheme by asking each of the participants, in turn, to allocate all of their points for the round. While not strictly followed, the intent appeared to be to work around the table, first clockwise and then anticlockwise. The order in which the participants voted is indicated in the third row of Table 9 and in Figure 10. Towards the end of the second round, the conversation revealed that participant B was trying to determine the number of votes that a scenario required to be selected for evaluation. Having determined that scenarios would require 10 votes to be selected he proceeded to add the required number of votes to his preferred scenarios. It appeared that participant A took a similar approach. Of the five scenarios selected for evaluation, one was *only* voted for by participant A and one was *only* voted for by participant B. Of the three other scenarios selected, two were influenced by the vote of participant F, who had the

final vote but no other vote. It appears that the participants with the last votes had a greater influence on the outcome of the voting stage than the other participants had.

This problem with the voting scheme is reflected in the results of the post-review questionnaire (Table 10). The participants were asked to compare SAAM reviews with other joint software reviews. The questions aimed to determine whether or not the SAAM review process resulted in better goal alignment than in traditional joint software reviews, and why. The first two questions, relating to the variables GC_1 and GC_2 concerned the resulting alignment of the participants' goals. The first question, relating to variable GC_1, asked whether or not the participants' goals were fully aligned, or the same. The second question, relating to GC_2, was included in case the goals were not fully aligned, but were still similar enough to enable them to all be addressed. The remaining variables concerned indirect measures of goal alignment. Goal alignment should be reflected in the focus of the review (GC_3) and is hampered when a single person dominates the review (GC_4).

The review participants who had attended other joint software reviews believed that the voting scheme did not give everyone the opportunity to ensure that their goals were met (mean less than 0). The standard deviation for this question was high – possibly because the goals of some of the participants were met while the goals of other participants were not met. The use of alternative voting schemes – such as an anonymous voting scheme – should be investigated to determine if this problem can be reduced. The voting scheme could still be conducted over two stages, but no participant would gain an advantage from the order in which the votes were elicited.

Table 10: The perceived impact of scenarios and the voting scheme on the software review process.

Variable	Related Question	Mean	Number of Responses	Standard Deviation
SAAM reviews are better than traditional joint software reviews because:				
GC_1	After voting on the scenarios everyone had the same goals.	0.2	6	1.2
GC_2	The voting scheme gave everyone the opportunity to ensure that their goals were addressed.	-0.3	6	1.2
GC_3	The scenarios provided a focus for the reviews.	1.7	6	0.5
GC_4	The use of scenarios and the voting scheme prevented one person from dominating the discussion.	0.3	6	1.3
GC	$(GC_1 + GC_2 + GC_3 + GC_4) / 4$	0.4	6	0.9

The participants who had attended other joint software reviews still believed that the SAAM review approach was beneficial (GC, mean greater than 0) because it provided a

focus for the review (GC_3). Note that there is strong support for this statement, as the mean is close to the value 2, which corresponds to the response option, Strongly Agree.

All of the participants were also asked a series of questions designed to determine whether or not the review process successfully aligned the goals of the participants. The questions (Table 11) were designed to look for indicators of poorly aligned goals (GA_1 and GA_4) as well as indicators of well-aligned goals (GA_2 and GA_3). The goals covered two areas. The first area concerned the identification of a common purpose, amongst factions (GA_1), or the group (GA_2). The second area, which was of particular interest, was whether or not the participants felt that they were able to contribute to scenarios that had been raised by other participants (GA_3 and GA_4).

Table 11: Participants' perceptions of goal alignment

Variable	Related Question	Mean	Number of Responses	Standard Deviation
$-GA_1$	During the meeting, it was easy to classify the participants into factions or groups with similar interests.	0.9	12	0.9
$+GA_2$	The meeting created a greater sense of shared purpose amongst the stakeholders.	1.1	12	0.8
$+GA_3$	During the meeting I tried to contribute towards **all** of the scenarios.	0.7	12	1.0
$-GA_4$	During the meeting it was difficult to identify issues in scenarios proposed by other meeting participants.	0.6	12	0.5
GA	$(GA_2 - GA_1 + GA_3 - GA_4)/4$	0.8	12	0.3

It appears that the goals of the participants were reasonably well aligned after the voting stage. The results from the questionnaire (Table 11) show no obvious signs that the participants were not working towards to common goals (GA_1 and GA_2) and the participants generally tried to contribute towards all of the scenarios (GA_3 and GA_4).

Table 12: Relationships between reviewers and scenarios.

Scenario	Reviewers		
	Raised or Modified	Voted On	Discussed
3	A	A	A, C, D
9	C	A, C, E	A, B, C
11	A, E	A, E, F	A, B, E, F
15	B	A, B, C	B, C, D, E
29	B	B	B, C, D, E

There was also some support for goal alignment from the observations. Table 12 shows the relationship between the scenarios, the reviewers and the activities that use the scenarios. The first lists the selected scenarios. The other columns indicate the activities undertaken by the reviewers. The letters in the cells refer to the reviewers who

contributed to the selection of, and the discussion of, the scenario through each activity. For example, scenario 3 was raised, and voted for, by participant A. However, during the scenario evaluation activity, it was discussed by participants A, C and D.

Of the six reviewers, at least three participated in the discussion of each scenario, and these reviewers were not necessarily those who voted on the scenarios. For example, participant D, who did not vote for any of the selected scenarios, still participated in the discussion of three of the scenarios. However, on average, only 60% of the reviewers participated in any of the discussions. When all the meeting participants are considered, the participation rate was slightly lower at 55%. The architect, the facilitator and one of the observers participated in all of the discussions, while the other two observers and the recorder did not participate in any of the discussions. Additional evidence from other reviews is required to determine whether or not these participation rates are higher than those of traditional joint software reviews. (See also the discussion on Freeloading in Section 7.6.)

6. Performance

Joint software reviews have two main objectives: 1) to identify and resolve issues that could affect the system being developed, and 2) the transfer of knowledge to develop a shared understanding between the acquirer and the developer. Section 6.1 discusses the performance of the EXC3ITE OCP/IMAD review according to the first objective and Section 6.2 discusses the performance of the review according to the second objective.

6.1 Issues stated, raised and resolved

Previous studies on software reviews where the participants can have conflicting goals have measured performance by comparing the number of issues raised and the number of issues resolved with the number of issues identified during the individual preparation phase (Kingston 1999a; Kingston et al. 1999c). This study determined the first two measures – the number of issues raised and the number of issues resolved – from analysis of the meeting minutes.

The third of these measures – the number of issues identified during the individual preparation phase – cannot be used for a SAAM review, because the individual preparation phase is used primarily for familiarisation and the identification of scenarios rather than issues. Instead, an alternative measure must be used. The measure chosen was the number of issues implied or stated in the discussions. That is, the number of issues either indirectly (*implied*) or directly stated (*stated*). This measure is the closest available measure. However, it only provides a lower bound for the number of issues *identified*. The participants in software reviews might identify issues, but not state them because of negotiation stumbling blocks. It is therefore likely that the number of issues stated in the earlier studies would be less than the number of

issues identified during the preparation phase. The number of issues *stated* is a valid measure for characterising the performance of SAAM reviews. However, the differences between this measure and the number of issues *identified* should be taken into account when interpreting the comparison of the results of the SAAM review with the results of other reviews.

The number of issues stated was obtained from the observations and checked by analysing the video tape. This analysis was done by three of the meeting organisers. Additional issues were recorded when two of the three organisers agreed with the statement of an issue, and the third organiser did not object to the statement being recorded as an issue. This new measure could be used in other observational studies of software reviews.

The results from the analysis of the video tape and the review minutes are given in Table 13 and Figure 11. Figure 12 provides a more detailed depiction of when issues were stated and recorded during the discussion of one of the scenarios. Interpretation of the table and the figures depends to some extent upon the discussion of the results. Therefore, the discussion of the features of the table is intertwined with the discussion of the results, and the figures are explained later.

Table 13 identifies four major categories associated with issue progress. These categories are shown in Italics, and are issues stated, issues raised, issues agreed and issues resolved. From this information, it appears that the SAAM review of EXC3ITE had reasonably large process losses ($1/3^{rd}$ of the issues that were stated were *not* recorded). However, large process losses were also observed in the other studies. In the study of Project Llama, all issues in a sample of issues identified during preparation were *not* recorded in the meeting minutes. In the laboratory studies reported in (Kingston 1999a; Kingston et al. 1999c), approximately half of the identified issues were *not* recorded – regardless of whether the participants had the same goals, different goals or conflicting goals. As previously discussed, the number of issues stated in the early studies would probably have been smaller than the number of issues identified. Therefore, only an upper bound for the level of process loss (as measured in the Case Study) can be determined for the earlier studies. A process loss for the SAAM review of EXC3ITE that was greater than 50% would indicate that the SAAM review process had greater process losses than the previous studies. However, it cannot be assumed that the lower value (33%) obtained for the EXC3ITE OCP/IMAD review reflects an improvement over the other reviews studied. Additional studies on other review processes, using the same measures of process loss are required to determine whether or not SAAM reviews have less process loss than do traditional software reviews.

Interpretation of the number of issues resolved also requires greater consideration than in the laboratory studies. In the laboratory studies, an issue was said to be resolved if the review participants had decided how the issue was to be addressed. However, in practice, the review participants also need to determine who is responsible for

addressing the issue. For example, the documentation may need to be reworked to the satisfaction of the architect and a particular reviewer or an independent expert.

Table 13: Issue progression through the EXC3ITE OCP/IMAD review.

Category	Number of Issues				
	EXC3ITE	Llama [9]	Laboratory Study		
			Same	Differ -ent	Conflict- ing
Identified	N/A	59	10.65	10.70	9.40
Implied or stated	33	Unknown			
Raised and recorded	21	0	5.10	4.35	4.20
Agreed	19				
Addressed by Discussions	(12)				
Action Required	(7)				
Resolved		0	4.15	3.55	3.60
Fully resolved	3				
No action required	(2)				
Action identified and assigned	(1)				
Partially Resolved	8				
Action identified but not assigned	(6)				
Could be included if required. No clarification of whether or not it was required.	(2)				
Not Clear	8				
Insufficient information available	(5)				
Minuted that resolution not clear	(3)				

Therefore, three minor categories and 8 sub-categories are used in Table 13. The minor categories are all associated with the resolved issues, and appear in the standard formatting used in the report. The subcategories were used to provide detailed information on how issues were resolved. In Table 13, the subcategories are indented, and the numbers of issues related to the sub-categories are given in brackets. Three minor categories were identified. Issues were said to be fully resolved if action was identified and assigned or if no action was required. Issues were said to be 'Partially Resolved' if an action has been determined, but no one was assigned responsibility for

[9] The numbers in this column reflect the issues identified before the review by a multi-disciplinary team that consisted of selected experts from the Information Technology Division. The issues were taken to the review by a single member of the team.

the action. A course of action was implied for two additional issues. During the discussion of these issues, the architect indicated that the features could be implemented if required. However, no one was assigned responsibility for determining whether or not the features were required. These issues were also classified as Partially Resolved. For the purposes of comparison with the laboratory studies, Partially Resolved issues are treated as Resolved issues. This is because they contain all of the information required for an issue to be classified as Resolved in the laboratory studies.

Analysis of the number of issues resolved in the EXC3ITE OCP/IMAD review involved additional difficulties. The meeting minutes indicated that there was some difficulty in determining whether or not action was required on three of the issues. The reviewers had claimed that these issues were satisfactorily addressed by the discussions, but they were raised again later in the meeting. The recorder therefore believed that action to further address these issues was probably required. These issues were classified as 'Not Clear'. There were five other issues where the minutes contained no record of how the issue was addressed by the discussions. Furthermore, because the issues were related to specialist areas, analysis of the videotape offered no insights into how these issues were addressed. Therefore, these five issues were also classified as Not Clear.

This is not just a problem with coding the issues. It also indicates that important issues might not be satisfactorily addressed after the review. It is believed that this could have been avoided if the facilitator had encouraged the participants to specify a course of action for each issue and to assign responsibility for the course of action. Cases where no action was required could then be clearly indicated in the minutes of the review. Instead, the facilitator sought general agreement that the issues had been satisfactorily addressed by the discussion.

Only two of the twenty-one issues raised were not resolved with explicit disagreement between the review participants. This is comparable with the resolution rates in the laboratory studies where groups resolved almost all of the issues raised. However, as in the laboratory studies, there is some doubt that the issues were resolved to the satisfaction of all the participants. The resolutions of seven of the twenty-one issues were classified as Not Clear.

Figure 11 shows an alternative representation of the progress of issues through the review. The horizontal axis shows the activities where issues were identified, in time order. The activities include the Architecture Description phase of the review, the discussion of the six scenarios, and the final discussion of the issues raised. Each line in the graph represents the cumulative total of the number of issues stated, recorded, agreed or resolved by the end of the activity.

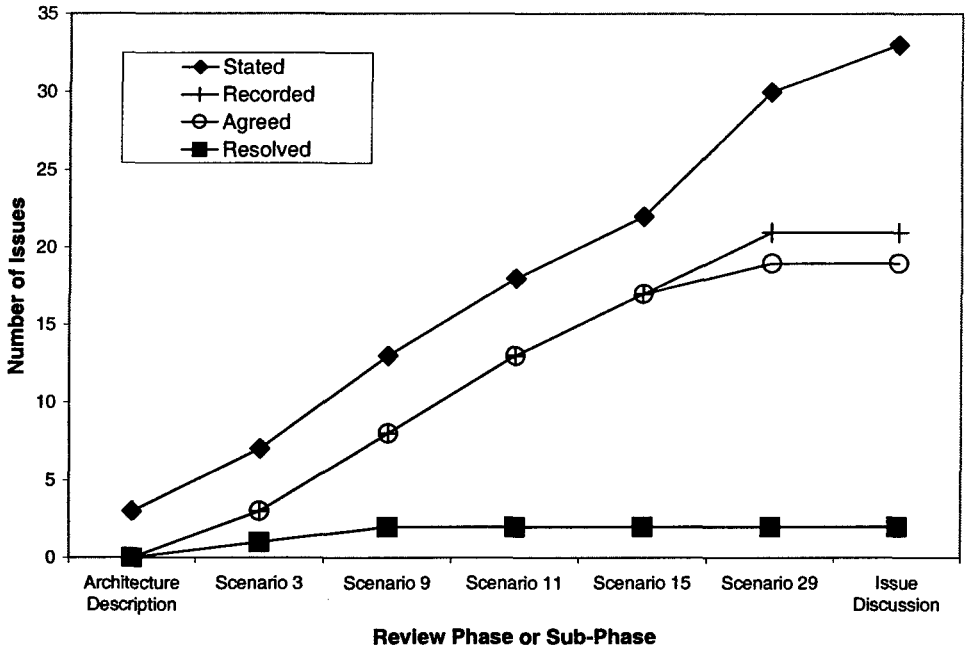


Figure 11: Progress of issues through the SAAM review (cumulative totals)

Figure 12 provides a more detailed depiction of the process of issues through the discussion of a single scenario. It shows the contribution of each reviewer during the evaluation of scenario 3, which took approximate 11½ minutes. The reviewers are shown on the vertical axis and time on the horizontal axis. The colour of the blocks indicates the type of contribution – presentation, discussion, raising an issue or recording an issue. Interruptions are indicated by a dashed line between the person who was interrupted, and the person who did the interrupting.

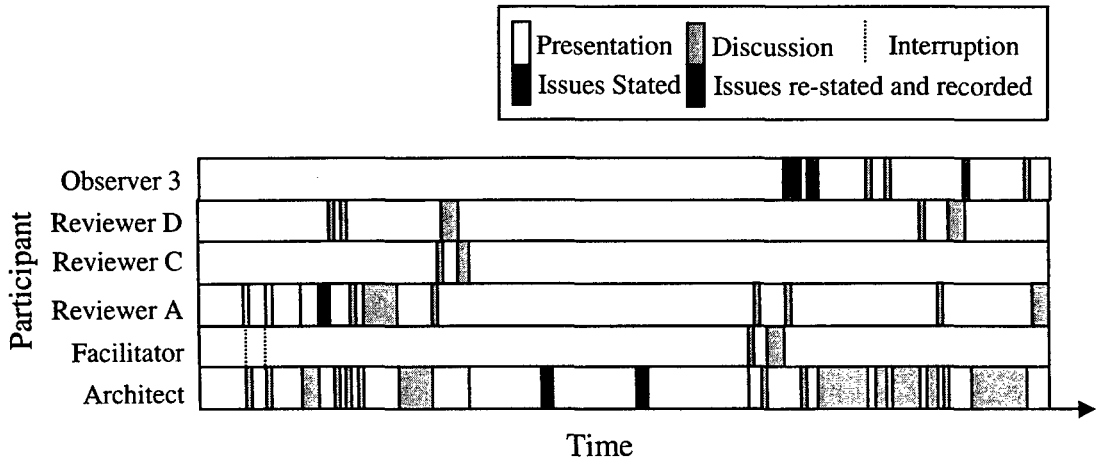


Figure 12: Discussion of Scenario 3. Issues were stated or implied during discussions and collected at the end of discussions. No attempt was made to agree on or resolve issues.

6.2 Knowledge transfer

The second main objective of a joint software review is to transfer knowledge to develop a shared understanding between the acquirer and the developer. One of the review organisers, the DSTO manager of the EXC3ITE OCP/IMAD development, identified five outcomes related to an increase in shared understanding that he hoped would result from the SAAM review. The success of the SAAM review in achieving these outcomes was assessed through 5 variables associated with the post-review questionnaire. These variables concerned: the identification of conflicts and trade-offs (P_1), the identification of new requirements (P_2), the clarification of existing requirements (P_3), improvements to the architecture, or architecture description (P_4) and improved individual understanding of the architecture (P_5). The results from the analysis of these variables (Table 14) show that the review participants believed that the EXC3ITE OCP/IMAD review was beneficial (mean greater than 0) in four of the five areas. It helped the participants achieve a better understanding of the architecture (P_5). The participants also believed that the review helped uncover requirements (P_2 and P_3) and that the review would lead to a better software architecture, or description of the architecture (P_4). However, as a whole, the participants did not believe (mean less than 0) that the review was good at uncovering conflicts and trade-offs (P_1). This is despite the architect explicitly discussing some trade-offs – such as the difficulty in determining when to incorporate new technologies such as JavaBeans. This suggests that the identification, and statement, of trade-offs by one party is not sufficient for a shared understanding of the trade-offs to be achieved. Additional work is probably required to understand how acquirers and developers achieve a shared understanding of the trade-offs that were made in developing the architecture.

Table 14: Shared understanding and the perceived benefits of the review.

Variable	Question	Mean	Number of Responses	Standard Deviation
P1	The SAAM review uncovered conflicts and trade-offs.	-0.3	12	0.8
P2	The SAAM review revealed new requirements.	1	12	0.9
P3	The SAAM review clarified existing requirements.	0.6	12	0.9
P4	The SAAM review should result in an improved architecture, or description of the architecture.	1	12	0.6
P5	I have a better understanding of the architecture since the SAAM review.	0.8	12	0.8
P	P1 + P2 + P3 + P4 + P5	0.6	12	0.5

The Case Study gave little indication as to why the review did not identify conflicts and trade-offs. Three possibilities are suggested for investigation in other studies. First, SAAM reviews, as implemented, might offer limited opportunity for uncovering

conflicts and trade-offs. It is possible that the Scenario Interaction stage, which was not conducted for the EXC3ITE OCP/IMAD review, would be beneficial in uncovering trade-offs even when a single architecture is being evaluated. Second, the quality and quantity of review documentation might have precluded the discovery of conflicts and trade-offs. The documentation consisted of 210 pages, some of which was derived from templates. However, the documentation contained neither an overview of the components in the architecture nor a discussion of how the requirements were reflected by the architecture. Third, the architect believed that the architecture could support almost any issue identified and often stated that 'anything was possible'. However, it was not clear how much the architecture needed to be changed or extended to support these features. The review was conducted as part of a research program and not as part of a fixed price contract. Thus extensions to the architecture were to be funded by the purchasers. This could have affected the attitude of the architect and the number of conflicts and tradeoffs that were identified. All of these alternative explanations require further investigation.

7. Negotiation stumbling blocks

It is believed that the performance of software reviews is adversely affected by negotiation stumbling blocks. This section discusses the level and impact of several negotiation stumbling blocks on the EXC3ITE OCP/IMAD SAAM review. The design of the study enabled a wide range of negotiation stumbling blocks to be investigated: Conformance Pressure, Cognitive Inertia, Production Blocking, Domination, Communication Breakdown and Freeloading. Some of these negotiation stumbling blocks had previously been investigated in the studies of software reviews that were introduced in Section 4. Where possible, the results of the EXC3ITE case study are compared with the results of these other studies. The additional information presented in this report could be used as a baseline for comparison in further studies on similar projects.

7.1 Conformance Pressure

In negotiations, Conformance Pressure is said to occur when participants are reluctant to criticise the comments made by other participants (Nunamaker et al. 1991). Conformance Pressure was studied because there was anecdotal evidence of Conformance Pressure in the Project Llama study (Kingston 1999b). Laboratory studies also indicated that perceived Conformance Pressure increases when review participants have different or conflicting goals, rather than the same goals (Kingston 1999a).

The EXC3ITE case study used two sources of information on Conformance Pressure. The first was the post-review questionnaire, which contained four questions pertaining to the level of Conformance Pressure experienced by the group. The first three questions were associated with the three occasions during a software review when

Conformance Pressure can arise; when issues are stated (CP_1), when they are raised (CP_2), and when they are resolved (CP_3). Participants may not state issues if they expect that the other group members will disagree with them. However, there are many other reasons why issues may not be stated. For example, an issue may not be raised because similar issues have already been considered. Therefore, rather than asking the participants about the issues which they did not raise, the questionnaire asked the participants about the issues they did raise. The question pertaining to CP_1 asked participants if they focused on the issues that they believed would be easy to resolve.

If another person disagrees with an issue that a participant has stated, then the participant may drop it due to Conformance Pressure. The question corresponding to the variable CP_2 addresses this component of Conformance Pressure.

Participants may also be discouraged from discussing issues where the review participants disagree on the issue or its resolution. The question corresponding to the variable CP_3 addresses this component of Conformance Pressure. These three questions were indirect questions that had been used in laboratory studies of software reviews (Kingston 1999a). The fourth question (CP_4) was a more direct question about the level of Conformance Pressure experienced by the group.

Table 15: Perceived Conformance Pressure in the EXC3ITE OCP/IMAD review, and in laboratory studies of software reviews.

Variable	Question	EXC3ITE OCP/IMAD			Laboratory Studies	
		Mean	Number of Responses	Standard Deviation	Mean Same	Mean Different or Conflicting
CP_1	During the meeting, the group tried to focus on issues that would be easy to resolve.	-0.6	11	0.8	0.15	-0.30
$-CP_2$	During the meeting, the group recorded all issues raised – even issues we did not agree on.	-1.1	12	1.0	0.55	0.00
CP_3	During the meeting, the group did not discuss the issues before we tried to resolve them.	-0.3	12	0.8	-0.95	-0.55
CP_4	During the meeting, I felt pressured to agree to issues that I either disagreed with, or did not fully understand.	-1.4	12	0.8	N/A	
CP	$(CP_1 - CP_2 + CP_3 + CP_4)/4$	-0.8	11	0.4	-0.10	-0.30

The results (Table 15) indicated that Conformance Pressure in the EXC3ITE review was generally low. Indeed, for the components CP_1 and CP_2, the level of Conformance Pressure was lower than that experienced by all groups in the laboratory studies – regardless of whether the members of the laboratory groups had the same goals, or different or conflicting goals. One possible explanation is that Conformance Pressure is greater between students than between researchers. However, this explanation is unlikely because the Conformance Pressure associated with resolving issues (CP_3) was higher in the EXC3ITE review, although this is possibly because very few issues were explicitly resolved in the EXC3ITE review. Indeed, the facilitator provided little opportunity for the reviewers to comment and discuss issues after they were identified. This has already been discussed in Section 5.2 and should improve as the facilitator gains more experience with SAAM reviews, or if facilitators were given additional training on the SAAM review process.

The second source of information about Conformance Pressure was provided by the analysis of the observations of the review. Interruptions and disagreements were recorded. During the meeting, there were thirty-five cases where a person was interrupted and six cases where disagreement with a statement was expressed. However, these do not appear to have affected the performance of the review as all the reviewers continued to raise issues after the interruptions and disagreements. This is probably because of the nature of the interruptions and disagreements. Most of the interruptions were to provide clarification of information during the presentations. For example, the two interruptions shown in Figure 12 were for clarification of information. Most of the disagreements were resolved when a second participant, with appropriate technical knowledge, agreed with one of the two participants between whom the disagreement had arisen. Resolution of disagreements might not have been so easy if the participants did not have respect for the other participants' technical knowledge, or if the disagreements had concerned the system's requirements rather than technical issues.

In summary, the main source of Conformance Pressure in the EXC3ITE OCP/IMAD review appears to have occurred during the resolution of issues. This is in contrast to the results of the laboratory studies where Conformance Pressure appeared to affect the stating and recording of issues, as well as the resolution of issues. Thus, SAAM reviews might have less Conformance Pressure than other joint software reviews. However, additional studies in other environments are required to confirm these results.

7.2 Cognitive Inertia

Cognitive Inertia occurs in negotiations when participants refrain from making comments that are not related to the current discussion (Nunamaker et al. 1991). That is, it occurs when comments remain focused on a single topic. However, SAAM reviews focus on one scenario at a time. If a scenario focuses the reviewers on a particular class of issue, it would be acceptable for similar issues to be raised during

the discussion of the scenario. Therefore, this study focused only on Cognitive Inertia during the scenario identification phase. The main limitation with this definition of Cognitive Inertia is that it does not allow reviews that use scenarios to be compared with other software review techniques. However, the author knows of no accepted measure of Cognitive Inertia that could be applied to all software reviews. The advantage of this measure of Cognitive Inertia is that it can be both subjectively and objectively evaluated, as in this study. The subjective evaluation was conducted using the post-review questionnaire and the objective evaluation was conducted by analysing the review documentation.

The questionnaire contained one question on Cognitive Inertia that explicitly asked the participants if they had experienced Cognitive Inertia (Table 16). The results were somewhat mixed (standard deviation of 0.8). Five of the twelve participants had neutral responses (response equivalent to 0), five disagreed with the statement (response equivalent to –1 or –2) and two of the participants agreed with the statement (response of 1).

Table 16: Perceived Cognitive Inertia in the EXC3ITE OCP/IMAD review.

Variable	Question	Mean	Number of Responses	Standard Deviation
CI_Q	After a scenario had been identified it was easier to identify similar scenarios than to identify a completely different type of scenario.	-0.3	12	0.8

Table 17: Demonstrated Cognitive Inertia in the EXC3ITE OCP/IMAD review.

Variable	Definition	Value
N	The number of scenarios identified.	35
S	The number of scenarios similar[10] to the previous scenario.	4
CI_O	S / (N - 1) [11]	0.12

However, the minutes of the review were more objective and showed that there was very little Cognitive Inertia (Table 18). The minutes contained a list of scenarios, and

[10] Similar scenarios were determined by asking two of the review organisers to organise the scenarios into groups of related scenarios. They classified 23 of the scenarios in an identical manner and a further 9 scenarios received the same high-level classification. Thus, while the classification was subjective, the inter-rater reliability was reasonably high (32/35) and it is reasonable to use the groupings as a classification scheme. Scenarios that were placed into the same group by either organiser were classed as similar scenarios.

[11] The level of cognitive inertia is normalised. The normalisation was designed so that a value of 1 means that all the scenarios were similar to the previously identified scenario and a value of 0 means that none of the scenarios were similar to the previously identified scenario. The normalisation divides by N-1 rather than N because the first scenario is not compared to any of the other scenarios.

the order in which they were identified. The level of Cognitive Inertia was calculated from this list of scenarios as described in Table 18. That is, the number of similar scenarios that were identified sequentially was compared with total number of scenarios identified. While the low level of Cognitive Inertia is promising, the level of Cognitive Inertia in other software reviews could be influenced by the amount of time the participants were given to prepare for the review, and the method used to elicit scenarios. In the EXC3ITE review, the participants were given time to identify scenarios. The facilitator then asked each participant to state one scenario before allowing the participants to state the other scenarios they had identified. Thus, future studies should be conducted to identify the impact of preparation time and facilitation strategies on Cognitive Inertia.

7.3 Production Blocking

Production Blocking is the term used for negotiation stumbling blocks that arise because the participants in a meeting or a negotiation cannot all contribute to the meeting at the same time (Diehl and Stroebe 1991; Lebie et al. 1996). Production Blocking was studied because there was anecdotal evidence of Production Blocking in the Project Llama study (Kingston 1999a). The three components of Production Blocking are Attenuation Blocking, Concentration Blocking and Attention Blocking. Attenuation Blocking occurs when "members who are prevented from contributing comments (issues) as they occur, forget or suppress them later in the meeting because they seem less relevant, original or important" (Nunamaker et al. 1991). Concentration Blocking occurs when new issues are not identified because the reviewers are trying to remember the previous discussions (Nunamaker et al. 1991). Attention Blocking occurs when the reviewers are concentrating on the discussions and do not have sufficient time to identify new issues (Nunamaker et al. 1991).

The level of Production Blocking in the SAAM review was characterised using two questions in the post-review questionnaire (Table 18). The variable PB_1 captures the perceived degree of Attenuation Blocking, while PB_2 captures the perceived degree of Concentration Blocking and Attention Blocking. The last two factors were not addressed separately, because it was believed that the participants would find it difficult to distinguish between the factors. Objective measures of Production Blocking could not be used to confirm the results from the questionnaire, because Production Blocking specifically refers to issues that were not stated. The results (Table 18) showed a wide variation in the level of Production Blocking. The participants generally believed that they were able to state all their concerns – concerns could either require clarifications of information or require new issues to be raised. However, some participants believed that they were not able to contribute fully because of the volume of information and the need to concentrate on the presentations. A large amount of documentation (over 200 pages) was supplied for the EXC3ITE OCP/IMAD review. This could have affected the level of Production Blocking in the review.

Table 18: Perceived Production Blocking in the EXC3ITE OCP/IMAD review.

Variable	Question	Mean	Number of Responses	Standard Deviation
PB_1	During the meeting, I had several concerns that I did not state because I believed that they would not be of interest to the rest of the group or because I believed they were not relevant to the review.	-1.0	12	0.6
PB_2	During the meeting, it was difficult to identify scenarios and issues because of the amount of information that had to be presented and the need to concentrate on what other people were saying.	-0.1	12	0.9
PB	(PB1 + PB2) / 2	-0.5	12	0.7

Additional studies should be conducted to determine the effect of documentation quantity and quality on Production Blocking. There was also anecdotal evidence of Production Blocking in the Project Llama study. However, this had a specific cause – a difference in how the reviewer had collated issues and the organisation of the meeting. Further studies are also needed to determine how the level of Production Blocking compares with the level in traditional software reviews, and to see if the level of Production Blocking can be further reduced.

7.4 Domination

Domination occurs when 'some group member(s) exercise undue influence or monopolize the group's time in an unproductive manner' (Nunamaker et al. 1991). Domination was studied using both the post-review questionnaire and observation of the EXC3ITE OCP/IMAD review.

The observations were analysed for large differences in the time that the participants were given to state their opinions. A participant was said to be dominating discussions if they spent more than twice the average time stating their opinions. This includes the time spent stating scenarios and issues, as well as time spent in discussion. It does not include the time spent presenting information. For example, the time spent by two of the reviewers presenting information on the SAAM review process was not included. The time spent presenting scenarios during the evaluation phase was also excluded.

The analysis uses a very coarse measure of time spent in discussion. The two observers recorded whenever a new participant commenced speaking. The time of the recording was automatically included. In these discussions, a single interval of speech by a single participant will be called a 'conversation block'. Over 400 conversation blocks were recorded in the SAAM review meeting, some of which lasted under 10 seconds. The duration of the conversation blocks was subject to measurement error. There was a

time delay between the start of a conversation block and when it was recorded. The time delay was not systematic as it varied depending on the number of people who spoke in quick succession, and the amount of additional information that had to be recorded. These errors mean that the times were not accurately recorded. There was an average variation of 30% in the total time attributed to the individual reviewers. However, this variation is acceptable given the definition of Domination.

The architect, the facilitator and the observers were excluded from the analysis of Domination. Thus, the results reflected Domination amongst the reviewers. This measure was chosen to avoid a possible bias in the measure due to the ratio of reviewers to the number of observers, and the single architect and facilitator. It was anticipated that the architect and the facilitator would spend more time in discussions than the other participants would. It was believed that the architect would spend more time in discussions because of his knowledge of the work product. It was believed that the facilitator would spend more time in discussions because of his role in coordinating the meeting and summarising the issues as they were raised. This was supported by the data (Table 19). The two observers who did not participate in the meeting were excluded from the analysis. They spent no time in discussion. Including them would lower the mean and skew the results of Domination. The analysis was conducted both with and without the third observer, who did participate in the meeting, because they acted as a reviewer. Their inclusion or exclusion made no difference to the figures reported in Table 19. As shown in the table, there was no Domination between the remaining review participants.

Table 19: Observed Domination in the EXC3ITE OCP/IMAD review.

Participants	**Time Spent in Discussion** Average for both observers (minutes)	**Domination**
Architect	59	Excluded from analysis as discussed above.
Facilitator	25	
Minimum Reviewer	4	False
Average Reviewer	8	
Maximum Reviewer	11	

The post-review questionnaire also asked the participants whether or not they believed that the meeting was dominated by one or two individuals (D). The participants generally disagreed (mean less than 0 in Table 20) that one or two individuals dominated the meeting. However, there was some variation in the responses (standard deviation of 0.8).

Table 20: Perceived Domination in the EXC3ITE OCP/IMAD review.

Variable	Question	Mean	Number of Responses	Standard Deviation
D	The meeting was dominated by one or two individuals.	-1.2	12	0.8

7.5 Communications Breakdown

Communications Breakdown occurs in negotiations when the participants stop discussing existing proposals and fail to propose alternative proposals for consideration (Nunamaker et al. 1991). Communications Breakdown has also been investigated in the laboratory studies of software reviews introduced in Section 4. In the EXC3ITE review, it was investigated using the post-review questionnaire. For consistency, the questions in the post-review questionnaire were identical to questions used in the laboratory studies. However, because the participants in the study did not identify issues before the SAAM meeting, only two of the questions were relevant. Communication Breakdowns occur in software reviews when the participants discuss issues without resolving them (CB_1 and CB_2). Table 21 summarises the questions and variables this study uses to determine the presence of Communications Breakdown in software reviews.

The responses from the questionnaire (Table 21) indicate that there was generally a low level of Communications Breakdown. However, there are some interesting differences between the results of the EXC3ITE OCP/IMAD study and the results of the previous laboratory studies. The EXC3ITE OCP/IMAD review showed a lower level of Communications Breakdown related to CB_2 and a higher level of Communications Breakdown related to CB_1. This was probably because the participants in the EXC3ITE OCP/IMAD review spent very little time trying to resolve issues. (The reasons for the low level of issue resolution were discussed in Section 5.2.) It appears reasonable that this would affect Communications Breakdown in the manner observed. Additional studies are required to determine if the observed changes in Communications Breakdown are characteristic of SAAM reviews, or if the changes in Communications Breakdown resulted from the short time spent discussing issues. In particular, additional studies are required on SAAM reviews using facilitators who are trained to identify conflict, especially latent conflict – that is conflict which is not explicitly expressed.

Table 21: Perceived Communications Breakdown in the EXC3ITE OCP/IMAD review.

Variable	Question	EXC3ITE OCP/IMAD			Laboratory Studies	
		Mean	Number of Responses	Standard Deviation	Mean Same	Mean Different or Conflicting
$-CB_1$	The group worked together to identify and/or resolve issues.	-1.1	12	0.5	-1.45	-1.35
CB_2	The group wasted time trying to resolve issues without coming to an agreement.	-1.3	12	0.6	-1.15	-0.95
CB	$(CB_2 - CB_1)/2$	-1.2	12	0.4	-1.30	-1.15

7.6 Freeloading

Freeloading participants attend the meeting, but do not prepare for the meeting or contribute to the meeting. Freeloading, as calculated in the study of Project Llama, for a review with N participants, is a normalised (/2N) measure of the number of participants that do not prepare for the review (N-P) plus the number that do not contribute to the review (N-C). The same calculation is used in this study. However, where possible, the measures of C and P are based on objective evidence rather than relying on self-reported measures of the participants' contributions to the meeting. In the EXC3ITE OCP/IMAD review, the reviewers and observers either contributed both scenarios and issues, or neither. Therefore, a participant was deemed to have contributed to the meeting if they were a reviewer or an observer and they contributed scenarios (0.5 of a contribution) or issues (0.5 of a contribution) or both (1.0 of a contribution). In contrast, self-reported contributions were used in the Project Llama study.

The facilitator, architect and recorder were not included in the analysis. There are arguments for and against this decision. The facilitator, architect and recorder have different roles than the other participants. Because of these roles, they should always contribute to the review in a specific manner. They should be included if the review is to be compared with other software reviews where the participants have different roles, or multiple participants in these roles. However, inclusion of these participants would make it difficult to compare the results from review meetings with more or less participants. That is, reviews with a small number of participants would always have relatively low levels of Freeloading regardless of the contributions of the reviewers, because of the contributions of the architect, the facilitator, and the recorder. For this reason, both sets of results are reported. The architect, facilitator and the recorder are deemed to have contributed to the review if they presented information or if they contributed issues. They were not allowed to contribute scenarios.

Table 22: Reviewers' and observers' preparation for the EXC3ITE OCP/IMAD review.

Activity conducted before the review meeting	Responses								
	Total	By Respondent							
Preparation or collation of the architecture Documents	3		■		■				■	
Familiarisation with the architecture Documents	9	■	■	■	■	■	■	■	■	■
Familiarisation with the SAAM process	4	■	■		■				■	■
Identification of scenarios	4			■			■	■	■	
Identification of issues	4		■		■		■		■	
Other	2					■			■	

Table 23: Facilitator, architect and recorder's preparation for the EXC3ITE OCP/IMAD review.

Activity conducted before the review meeting	Responses			
	Total	By Respondent		
Preparation or collation of the architecture Documents	1		■	
Familiarisation with the architecture Documents	2	■		■
Familiarisation with the SAAM process	3	■	■	■
Identification of scenarios	1		■	
Identification of issues	0			
Other	2	■		■

To determine whether or not the participants had prepared for the review, they were requested to describe how they had prepared for the review. The participants were given six options as shown in the first columns of Table 22 and Table 23. The responses from the facilitator, the architect and the recorder (Table 23) were separated from those of the reviewers and observers (Table 22). The stated activities of the participants are shown by the shaded entries in the tables. The 'Total' columns indicate the number of participants shown in the table, who undertook each type of activity. All the participants, who indicated that they had undertaken any preparation for the review, were deemed to have prepared for the review.

Table 24 summarises the information used to calculate the level of Freeloading in the review. As shown in the second-last two columns, two values of Freeloading were calculated to enable comparison with the results of the Project Llama review, which are given in the far-right column. The second column from the right contains the value of Freeloading calculated using the same method as in the Project Llama study. This value is independent of the review process. However, in the SAAM review process, the roles of the architect, the facilitator, and the recorder mean that they *have* to contribute to the review. This means that SAAM reviews with a small number of participants could have a higher level of Freeloading, without necessarily having had a larger percentage of participants contribute to the identification of scenarios and issues. Therefore, Freeloading was also calculated for a subset of the participants, for the reviewers and the observers.

Table 24: Freeloading in the EXC3ITE OCP/IMAD review and the Project Llama review.

Variable		Number or Value		
		EXC3ITE OCP/IMAD		Project Llama
		Reviewers and Observers	Total	
N	Number of meeting participants.	9	12	5 Respondents
C	Contributors	7	10	3
P	Participants who prepared for the review	9	12	5
FL	Free loading = (2N – P – C)/2N	0.11	0.08	0.1

The level of Freeloading recorded for the Project Llama review is very similar to the level of Freeloading amongst the reviewers and observers in the EXC3ITE OCP/IMAD review and slightly higher than that observed amongst all the participants in the review (Table 24). However, the questionnaire used to calculate Freeloading in the Project Llama study had a low response rate (33%) and there was some evidence to suggest that the actual level of Freeloading in the Project Llama review could have been significantly higher (Kingston, 1999b). It is also possible that the level of Freeloading in the EXC3ITE OCP/IMAD review is not indicative of SAAM reviews. One of the stakeholders could not attend the review and instead sent two replacements. The role of these replacements was not clear, and the facilitator did not solicit information from them. If the role of these participants had been made clear to the facilitator, it is possible that they would have contributed to the review. Thus, it appears that the level of Freeloading in SAAM reviews is at least as low as that in traditional software reviews. Additional studies are necessary to confirm whether the level of Freeloading in SAAM reviews is equal to or lower than that in other joint design reviews.

8. Implications

The study has several implications for the conduct of, and research into, joint software reviews and SAAM reviews. The implications depend, in part, on the external validity of the study.

8.1 External validity

The external validity of this study is limited by five main factors. First, the study was self-selecting. The systems architect and the EXC3ITE OCP/IMAD project manager both volunteered to trial the SAAM review process during the review of the EXC3ITE OCP/IMAD concept demonstrator. They had both attended training in the SAAM process and both believed that the process would be beneficial. Thus, the systems architect and the project manager might have been more willing to cooperate than might the participants in a typical joint software review. That is, the project being studied could be atypical.

Second, only a single review was studied. The 4-hour review had 12 participants, 210 pages of documentation, and poorly understood and documented requirements, some of which were only delivered on the day before the review (Section 4.4). While the requirements might be poorly understood in many reviews, care should be taken when interpreting the results of the study for reviews where the requirements are well understood. Variations to the environment in which a review is conducted could affect the review characteristics. For example, the preparation time and facilitation strategies might effect the level of Cognitive Inertia and Communications Breakdown. The amount and quality of the documentation might effect the level of Production Blocking. More importantly, the EXC3ITE project is not a typical software project. DSTO, in partnership with DAO, have a role in managing the EXC3ITE project. Consequently, the participants in the EXC3ITE OCP/IMAD review all had a strong technical (software engineering) background rather than a strong military background. It is believed that this resulted in the generation of more issues that were technical, and less issues that were related to the requirements, than would be generated in the review of a typical project. The participants' respect for each other's technical expertise might also mean that the participants had less difficulty resolving issues than might the participants in a typical joint software review.

Third, the experience of the participants with the SAAM review process was limited. None of the participants, including the facilitator, had previously participated in a SAAM review. Several of the limitations identified with the SAAM review process could have been artefacts of this inexperience. However, this might not be a severe limitation on the study. The participants in traditional joint software reviews do not always have previous review experience.

Fourth, not all the steps of the SAAM review process were included in the EXC3ITE OCP/IMAD review. The Scenario Interaction step and the Scenario Weighting step were not included. The Scenario Weighting step was not included because only a single architecture was under review. The facilitator did not include a Scenario Interaction step. This step identifies the components that would need to be modified to support the scenarios. It is believed that classification of scenarios, which contributed little to the EXC3ITE OCP/IMAD review, would be most beneficial during the Scenario Interaction step. It is believed that assessment of the value of the stages in the SAAM review process needs to consider the context of the review.

These first four limitations can be addressed with additional studies of the SAAM review process conducted on a range of projects. Future studies are planned, which will use a more experienced facilitator and a modified review process, in order to address some of these limitations.

The fifth limitation is more difficult to address. The study did not allow for a direct comparison between traditional software reviews (of designs) and SAAM reviews. The

strengths and limitations of the SAAM review process are now discussed in light of this limitation.

8.2 SAAM reviews and traditional joint software reviews

The study provides limited evidence to support the use of SAAM reviews. No direct comparisons were made between SAAM reviews and traditional joint software reviews. While comparisons were made with the results of previous studies, these reviews were conducted in different contexts. Additional studies are required to compare SAAM reviews and traditional joint software reviews for a range of projects. It is unlikely that a study of SAAM reviews and traditional joint software reviews could be conducted on the same stage of a single project, without introducing alternative biases. Therefore, multiple studies of both traditional joint software reviews and SAAM reviews are required to clearly identify the strengths and limitations of the two approaches.

The conclusions drawn from the comparison between SAAM reviews and the previous studies of software reviews can only be tentative, because of the differences between the three studies. The studies use different empirical techniques. For example, this study is the only study to use detailed observations. The reviews being studied also had different characteristics. The laboratory studies used two-person reviews of students reviewing a 6-page *design* document. The subjects were also provided with a 16-page document containing *supporting information* that included explicit system requirements. The subjects' goals for the review concerned additional, implicit requirements. The Project Llama and the EXC3ITE OCP/IMAD review were both case studies conducted in an industrial setting. The traditional architecture review of Project Llama and the SAAM review of the EXC3ITE OPC/IMAD architecture had similar numbers of participants (15 versus 12). The requirements were not fully specified for either system, although Project Llama was designed to replace an existing system. Less documentation was reviewed during the Project Llama review (142 pages) than for the EXC3ITE OPC/IMAD (210 pages). However, the main differences between the reviews were the nature of the projects. Project Llama was being acquired through a project run by the DAO. The project aimed to develop a replacement for a system component with which the system users interact. In contrast, EXC3ITE OCP/IMAD was part of a research project jointly managed by DSTO and the DAO. The project aimed to develop a concept demonstrator for future command and control systems. These differences should be considered when comparing the results of the study.

According to the review participants who had experienced traditional software reviews, the SAAM review of the EXC3ITE OCP/IMAD architecture was better than traditional software reviews (Table 4 and Table 6). However, the objective evidence shows only limited support for this opinion.

The SAAM review had lower levels of process loss (30%) than observed for either the Project Llama review (100% of sample issues) or the laboratory reviews (mean of

approximately 50%). However, the process loss in the SAAM review process would tend to be lower because of variations in how process loss was defined. The participants in the SAAM review process did not try to identify issues before the review, so the process loss is based on the number of issues stated rather than the number of issues identified before the review. Additional studies on traditional software reviews using this definition of process loss are needed before the process losses can be accurately compared.

Other characteristics of the SAAM review process were similar to the characteristics of the traditional software reviews. The SAAM review had similar issue resolution rates to the reviews in the laboratory studies, and in both studies it appears that issues might have been resolved prematurely (Section 6.1). A comparison between the negotiation stumbling blocks encountered by the groups in the laboratory studies and the group in the Project Llama review showed mixed results (Section 7). The level of some components of Conformance Pressure and Communication Blocking increased, while others decreased. Further, a comparison between the Project Llama review and the EXC3ITE OCP/IMAD review shows that the participants' goals were partially met using both techniques (Section 5.2). The level of Freeloading was also similar in the two studies (Section 7.6). However, the true level of Freeloading in the Project Llama review was probably considerably higher than that reported. It is believed that the respondents who did not return the questionnaire were observers, who made limited contributions to the review. Thus, early identification of stakeholders in the SAAM review process might reduce the level of Freeloading and therefore the cost of joint software reviews.

SAAM reviews may still offer benefits over traditional joint software reviews. The process had considerable support from the participants and exhibited a number of desirable properties. A number of potential improvements to the SAAM review process were also identified. The strengths and limitations of SAAM reviews are now discussed in more detail.

8.3 SAAM Reviews

The SAAM review process has two main differences from traditional joint software reviews: the use of a facilitator and the use of scenarios including voting on the scenarios. Both were generally believed to be beneficial. However, the study identified several areas where improvements should be possible.

The facilitator used in the study was inexperienced with the SAAM review process. While he received some training in the process before the review, it is believed that additional training and guidance would have been beneficial. Additional studies could be conducted to determine the best methods of training experienced facilitators in the SAAM review process. A trained facilitator should be able to ensure that the participants' roles are well understood (Section 5.1) and that issues are clearly identified and resolved to the satisfaction of all participants (Section 5.2). They should

also facilitate the elaboration and grouping of scenarios before the selection of scenarios.

Methods of elaborating and grouping scenarios, and methods of selecting scenarios would also benefit from additional studies. The review group were given little guidance, and failed to identify and group similar scenarios (Section 5.4). Alternative processes for elaborating and grouping scenarios could be developed, investigated, and used to produce guidelines for reviewers and for facilitators.

The participants in the EXC3ITE OCP/IMAD review also believed that the voting process used to select scenarios for evaluation was biased. Most of the scenarios investigated were raised and voted on by the participants who had the final three sets of votes. Alternative mechanisms should be investigated. The current voting process has two design features. First, the voting patterns are visible to all the participants. Second, the voting has two rounds so that voting patterns can be modified based on the results from previous rounds. There is no evidence that these design guidelines have been empirically or theoretically validated, although other SAAM reviews have been conducted. Therefore, the guidelines, and alternative guidelines or voting schemes should be investigated further.

Both the voting schemes and the guidelines for elaborating and grouping scenarios could be investigated outside the context of the SAAM review process. However, the study also provides a baseline for testing the impact of these and other changes on the SAAM review process. The SAAM review had a high level of participant satisfaction, but the participants did not identify conflicts and trade-offs. The SAAM review had a very low level of Domination. It also had reasonably low levels of other negotiation stumbling blocks – such as Cognitive Inertia and Production Blocking. Furthermore, the evidence from this study suggests that these levels could be improved, further reducing the loss of issues due to Cognitive Inertia and Production Blocking. The results suggest that improved voting schemes and guidelines for grouping scenarios could improve the degree of Goal Alignment and possibly increase the percentage of goals completely met by the review, thus improving the performance of the review.

9. Recommendations

The study highlighted the characteristics of the SAAM review process and provided a preliminary, but unconclusive, comparision between SAAM reviews and traditional software reviews. The use of SAAM reviews could result in improved review performance (lower loss of issues or problems) and a decreased level of Freeloading (or increased contibution from all participants). However, the SAAM review process has some limitations and the participants did not believe that the SAAM review process was useful for identifying conflicts and tradeoffs.

Consequently, this report recommends the continued study of the SAAM review process. Three areas deserve particular attention. First, the required skills and training of the facilitator need to be identified. For example, in the study, the facilitator was observed to exert Conformance Pressure resulting in the premature resolution of issues. It is believed that this problem could be addressed through training. Second, the process by which scenarios are elaborated and combined should be studied. Participants in the study found this activity difficult because of lack of guidance. Third, the voting process used to select scenarios should be studied. The voting process was incorrectly implemented, and the participants believed that, even if implemented correctly, the voting process would be biased. If SAAM reviews are to be used before further study, it is recommended that the participants have to record their votes before the start of each round of voting. This would ensure that the last participants to state their votes are not able to ensure that their scenarios are selected by adjusting their votes.

Until these areas are addressed, the SAAM review process cannot be recommended as a replacement for traditional joint architecture reviews. However, the SAAM review process did appear to be beneficial at identifying and clarifying requirements. SAAM reviews might therefore be a useful addition to some projects, such as those using evolutionary acqusition. During an evolutionary acquisition the detailed requirements are refined during each phase in the acqusition (Henderson and Gabb 1997).

References

Adler, P. A. and P. Adler (1994). Observational Techniques. In *Handbook of Qualitative Research.* (N. K. Denzin and Y. S. Lincoln. Eds.), pp. 377-392, SAGE Publications, Thousand Oaks, California.

ADO (1998). Electronic Systems Acquisition Division Compendium, Department of Defence, Canberra.

Basili, V. R., R. W. Selby and D. H. Hutchens (1986). Experimentation in Software Engineering. *IEEE Transactions on Software Engineering*, **SE-12(7)**, 733-743.

Bass, L., P. Clements, R. Kazman and G. Abowd (1998). Analyzing Development Qualities at the Architectural Level: The Software Architecture Analysis Method. In *Software Architecture in Practice.* (L. Bass, P. Clements and R. Kazman. Eds.), Addison Wesley.

Chung, W. Y. and P. J. Guinan (1994). Effects of Participative Management on the Performance of Software Development Teams. In *ACM SIGCPR Conference* (J. W. Ross. Ed., pp. 252-260, ACM Press.

Diehl, M. and W. Stroebe (1991). Productivity Loss in Brainstorming Groups: Tracking Down the Blocking Effects. *Journal of Personality and Social Psychology*, **61**, 392-403.

Ellis, L. (1994). *Research methods in the Social Sciences.* WCB Brown & Benchmark.

EXC3ITE (1999). OOPL Architecture.
http://exc3ite.dsto.defence.gov.au/pubs/architectureDesign/OOPLarch/OOPLarch.htm.

Fenton, N. E. and S. L. Pfleeger (1996). *Software Metrics: A Rigorous and Practical Approach*. 2nd Edition. Chapman & Hall.

Foroughi, A, W. C. Perkins and M. T. Jelassi (1995). An Empirical Study of an Interactive, Session-Oriented Computerized Negotiation Support System (NSS). *Group Decision and Negotiation*, **4**, 485-512.

Hay, D. (1997). Software Development Plan for the Java Situation Monitoring (Project Llama) No. ADI-Perth-65, ADI, Perth.

Henderson, D. and A. Gabb (1997). Using Evolutionary Acquisition for the Procurement of Complex Systems. Technical Report No. DSTO-TR-0481, DSTO, Adelaide.

ISO/IEC 12207 (1995). Information Technology Software Life Cycle Processes. Standard No. ISO/IEC 12207, ISO/IEC.

Kaplan, B. and D. Duchon (1988). Combining qualitative and quantitative methods in information systems research: a case study. *MIS Quarterly*, **12(4)**, 571-86.

Kazman, R. (1999). Using Scenarios in Architecture Evaluations. *SEIinteractive*, **2(2)**.

Kazman, R., G. Abowd, L. Bass and P. Clements (1996). Scenario-Based Analysis of Software Architecture. *IEEE Software*, **13(6)**, 47-55.

Kingston, G. (1999a). Goal Conflict in Software Reviews. *School of Information Systems*. PhD Thesis submitted to University of New South Wales, Sydney.

Kingston, G. (1999b). Joint Software Reviews: A Case Study from JP2030. Research Report No. DSTO-RR-0156, Defence Science and Technology Organisation, Adelaide.

Kingston, G., R. Webby and R. Jeffery (1999c). Different and Conflicting Goals in Software Reviews. In *European Conference on Information Systems*, Copenhagen, Denmark.

Kitchenham, B., L. Pickard and S. L. Pfleeger (1995). Case Studies for Method and Tool Evaluation. *IEEE Software*, **Jul**, 52-62.

Lebie, L., J. A. Rhoades and J. E. McGrath (1996). Interaction Process in Computer-Mediated and Face-to-Face Groups. *Computer Supported Cooperative Work*, **4(2-3)**, 127-152.

Lee, A. S. (1989). A scientific methodology for MIS case studies. *MIS Quarterly*, **13(1)**, 33-50.

Lethbridge, T. C., S. E. Sim and J. Singer (1998). Software Anthropology: Performing Field Studies in Software Companies. *http://www.site.uottawa.ca/~tcl/papers/TSE/TSE-swanthrop.doc*.

Lewicki, R. J., S. E. Weiss and D. Lewin (1992). Models of conflict, negotiation and third party intervention: A review and synthesis. *Journal of Organizational Behaviour*, **13**, 209-252.

Likert, R. (1932). A technique for the measurement of attitudes. *Archives of Psychology*, **21(40)**, 1-55.

Luthans, F. (1985). *Organizational Behaviour*. McGraw-Hill.

Macdonald, F., J. Miller, A. Brooks, M. Roper and M. Wood (1996). Automating the Software Inspection Process. *Automated Software Engineering*, **3**, 193-218.

MIL-STD-498 (1994). Software Development and Documentation. AMSC No. N7069, Department of Defense (USA).

Nunamaker, J. F., A. R. Dennis, J. S. Valacich, D. R. Vogel and J. F. George (1991). Electronic Meeting Systems to Support Group Work. *Communications of the ACM*, **34(7)**, 42-61.

Porter, A. A., L. G. Votta and V. R. Basili (1995). Comparing Detection Methods for Software Requirements Inspections: A Replicated Experiment. *IEEE Transactions on Software Engineering*, **21(6)**, 563-575.

Quin-Conroy, D. (1999). Distributed Command and Control Applictions Using Java. In *JavAus'99* (E. Crameri. Ed., DSTC Pty Ltd, Sydney, Australia.

Sauer, C., R. Jeffery, L. Lau and P. Yetton (1999). A Behaviourally Motivated Programme for Empirical Research into Software Development Technical Reviews. *Accepted for publication in IEEE Transactions on Software Engineering.*

Shea, P and R. A. Guzzo (1987). Groups as Human Resources. In *Research in Personnel and Human Resources Management*, pp. 323-356, JAI Press Inc.

Stake, R. E. (1994). Case Studies. In *Handbook of Qualitative Research.* (N. K. Denzin and Y. S. Lincoln. Eds.), pp. 236-247, Thousand Oaks, California.

Weidenhaupt, K., K. Pohl, M. Jarke and P. Haumer (1998). Scenarios in Systems Development: Current Practice. *IEEE Software*, **15(2)**, 34-45.

Yin, R. K. (1984). *Case Study Research Design and Methods*. Sage Publications, Newberry Park, CA.

Acknowledgements

This study would not have been possible without the contributions and approval of several people. In particular, the author thanks Brett McDowall (OOPL) and Phil Collier (IAG) for allowing the study of the EXC3ITE OCP/IMAD review. The quality of the study was greatly enhanced by the willingness of the review participants to be observed, recorded, and to complete the questionnaires. In addition to review organisers the participants included Kevin Nuttall; Derek Henderson, Dr Michael Pilling, Matthew Phillips and Dr Rudi Vernik (SSE); Greg Chase (IMF); John Hildebrandt and Mark Grigg (C3ISC); and Mark Daniel and Tony Gray (SSD).

Thanks go to those who reviewed the questionnaires: Phil Collier (IAG); Prof Ross Jeffery and Dr Richard Webby (UNSW); and Derek Henderson and Dr Rudi Vernik (SSE) for their comments. The author also thanks Stephen Fry for recording his

observations of the review, which greatly enhanced the confidence that can be placed in the results of the study.

Appendix A: Pre-Review Questionnaire

The pre-review questionnaire was used to collect three types of information from the participants in the EXC3ITE OCP/IMAD review. First, information was obtained about participants' experience with joint software reviews and with the SAAM review process. This provided information about the context in which the review was conducted. According to models of software reviews (Kingston 1999a; Sauer et al. 1999), the experience of participants can affect the performance of a software review. Second, information was obtained about the participants' roles in the review. This also provided information about the context in which the review was conducted. Third, information was collected about the participants' goals. This information was used in the interpretation of the information collected in the post-review survey about the participants' goals that met by the review.

Id:

SAAM Review: Pre-Review Questionnaire

Part A: Background

Project: *EXC3ITE*

Review: *OCP/IMAD ARCHITECTURE*

Date(s): *2 March 1999*

Name (Optional): _____

Phone Number (Optional): _____

Organisation: _____

What were your expectations for the review, and what did you hope to achieve at the review?

_____

_____

_____

Experience

Previous Joint Software Review Experience*:

- ❑ 0 Reviews
- ❑ 1-5 Reviews
- ❑ > 5 Reviews

Last Joint Software Review* (other than a SAAM review)

- ❑ Never
- ❑ < 12 months
- ❑ > 12 months

Previous SAAM Review Experience:

- ❑ No experience
- ❑ Training or 1 Review
- ❑ > 1 Review

* A joint software review is a review attended by both acquirers and the developers.

Part B: Review Information

What is your role in the review? (Tick all that apply).

- ❑ Facilitator
- ❑ Recorder
- ❑ Reviewer
- ❑ Observer
- ❑ EXC3ITE Developer
- ❑ EXC3ITE Project Manager
- ❑ EXC3ITE System Manager
- ❑ CTD Developer
- ❑ Architect/Presenter
- ❑ Research User / Technical Expert
- ❑ Other _____

What are your goals for the review? Using all three columns and the space provided for listing additional goals, rank your goals for the review. Use rank 1 for the most important goal, 2 for the second most important goal etc. Do not repeat numbers. You do **not** have to rank every goal.

Consider Technical Design:

- ❑ Performance Aspects
- ❑ Requirements Met
- ❑ Support for Experiments
- ❑ Minimise IMAD changes
- ❑ Minimise OCP changes
- ❑ Flexibility (for other CTDs)

Consider other issues:

- ❑ Documentation Quality
- ❑ Future supportability
- ❑ Proof of OCP Quality
- ❑ Risk Assessment
- ❑ Cost/Benefit of alternatives

- ❑ Verify the architecture
- ❑ Secure agreement on approach
- ❑ Present information
- ❑ Address issues and scenarios

Other:

❑ _____ ❑ _____ ❑ _____

❑ ❑ ❑

Appendix B: Post-Review Questionnaire

The post-review questionnaire was used to collect three types of information from the participants in the EXC3ITE OCP/IMAD review. First, information was obtained about whether the participants' goals were met completely, partially, or not at all. Second, information was obtained about the negotiation stumbling blocks encountered by the participants. Third, information was obtained about the perceived strengths and weakness of the SAAM review process. All three sources of information were used in the characterisation and evaluation of the SAAM review process.

Id:

SAAM Review: Post-Review Questionnaire

Part A: Goals

Indicate which of your goals were met completely (C), partially (P) or not met (N).

Consider Technical Design:

- ❑ Performance Aspects
- ❑ Requirements Met
- ❑ Support for Experiments
- ❑ Minimise IMAD changes
- ❑ Minimise OCP changes
- ❑ Flexibility (for other CTDs)

Consider:

- ❑ Documentation Quality
- ❑ Future supportability
- ❑ Proof of OCP Quality
- ❑ Risk Assessment
- ❑ Cost/Benefit of alternatives

- ❑ Verify the architecture
- ❑ Secure agreement on approach
- ❑ Present information
- ❑ Address issues and scenarios

Other_____

_____

_____

Part B: SAAM Evaluation

1. During the meeting, the group recorded all issues raised - even issues we did not agree on.

1	2	3	4	5
Strongly Agree				**Strongly Disagree**

2. During the meeting, the group did not discuss the issues before we tried to resolve them.

1	2	3	4	5
Strongly Agree				**Strongly Disagree**

3. During the meeting, the group tried to focus on issues that would be easy to resolve.

1	2	3	4	5
Strongly Agree				**Strongly Disagree**

4. The meeting was dominated by one or two individuals.

1	2	3	4	5
Strongly Agree				**Strongly Disagree**

5. After a scenario had been identified it was easier to identify similar scenarios than to identify a completely different type of scenario.

1	2	3	4	5
Strongly Agree				**Strongly Disagree**

6. The group worked together to identify and/or resolve issues.

1	2	3	4	5
Strongly Agree				**Strongly Disagree**

7. The group wasted time trying to resolve issues without coming to an agreement.

1	2	3	4	5
Strongly Agree				**Strongly Disagree**

8. During the meeting, I felt pressured to agree to issues that I either disagreed with, or did not fully understand.

1	2	3	4	5
Strongly Agree				**Strongly Disagree**

9. During the meeting, I had several concerns that I did not state because I believed that they would not be of interest to the rest of the group or because I believed they were not relevant to the review.

1	2	3	4	5
Strongly Agree				**Strongly Disagree**

10. During the meeting, it was difficult to identify scenarios and issues because of the amount of information that had to being presented and the need to concentrate on what other people were saying.

1	2	3	4	5
Strongly Agree				**Strongly Disagree**

11. The SAAM review uncovered conflicts and tradeoffs.

1	2	3	4	5
Strongly Agree				**Strongly Disagree**

12. During the meeting, it was easy to classify the participants into factions or groups with similar interests.

1	2	3	4	5
Strongly Agree				**Strongly Disagree**

13. The meeting created a greater sense of shared purpose amongst the stakeholders.

1	2	3	4	5
Strongly Agree				**Strongly Disagree**

14. Requirements were well documented and I understood them prior to the review.

1	2	3	4	5
Strongly Agree				**Strongly Disagree**

15. The SAAM review revealed new requirements.

1	2	3	4	5
Strongly Agree				**Strongly Disagree**

16. The SAAM review clarified existing requirements.

1	2	3	4	5
Strongly Agree				**Strongly Disagree**

17. During the review meeting I tried to contribute towards **all** of the scenarios.

1	2	3	4	5
Strongly Agree				**Strongly Disagree**

18. During the review meeting it was difficult to identify issues in scenarios proposed by the other meeting participants.

1	2	3	4	5
Strongly Agree				**Strongly Disagree**

19. The SAAM review should result in an improved architecture, or description of the architecture.

1	2	3	4	5
Strongly Agree				**Strongly Disagree**

20. I have a better understanding of the architecture since the SAAM review.

1	2	3	4	5
Strongly Agree				**Strongly Disagree**

Part C: SAAM Comparison

If you have participated in *one or more* joint software reviews that did *not* use the SAAM review process, please tick all of the following statements with which you agree.

21. SAAM reviews offered no benefit over traditional joint software reviews.

1	2	3	4	5
Strongly Agree				**Strongly Disagree**

SAAM reviews are better than other software reviews because:

22. The facilitator was able to prevent conflict from escalating.

1	2	3	4	5
Strongly Agree				**Strongly Disagree**

23. After voting on the scenarios everyone had the same goals.

1	2	3	4	5
Strongly Agree				**Strongly Disagree**

24. The voting scheme gave everyone the opportunity to ensure that their goals were addressed.

1	2	3	4	5
Strongly Agree				**Strongly Disagree**

25. SAAM reviews are more structured than ordinary reviews.

1	2	3	4	5
Strongly Agree				**Strongly Disagree**

26. The scenarios provided a focus for the reviews.

1	2	3	4	5
Strongly Agree				**Strongly Disagree**

27. The use of scenarios and the voting scheme prevented one person from dominating the discussion.

1	2	3	4	5
Strongly Agree				**Strongly Disagree**

28. The use of the facilitator prevented one person from dominating the discussion.

1	2	3	4	5
Strongly Agree				**Strongly Disagree**

29. The use of the facilitator ensured that everyone's goals were considered when each issues was discussed.

1	2	3	4	5
Strongly Agree				**Strongly Disagree**

Part D: Expectations and Improvements

Describe how your expectations for the SAAM review compare with the actual SAAM review.

How do you believe that the SAAM review could have been improved?

Appendix C: Coding Observations

The two research observers used a systematic process to record the activity during the EXC3ITE OCP/IMAD review. The observers recorded conversation blocks. A conversation block is a period when a single person spoke on a single general topic, such as a particular issue, in a single manner (conversation type). At the start of a conversation block, the observers recorded the person speaking using a database that automatically recorded the time of the entry. The observers then tried to further classify the conversation as in Table 25. Additional information was collected to support the evaluation of the SAAM review process.

Table 25: Basic information recorded by the research observers.

Information	**Recording mechanism**
Start of conversation block	Time automatically recorded by the database. The completion time was identified by the start of the next recording as breaks in verbal communication
Person speaking	Selected from a list of the meeting participants.
Conversation type	Selected from the following list: • *Presentation.* This included material presented at the whiteboard, on the overhead projector and opening presentations by the facilitator. It included the presenation of scenarios by the person who identified the scenario, except when questions were asked. • *Discussion.* This included asking and answering questions, identifying scenarios. • *Social.* This included introductions. It was also intended to include asides about the weather, sport, news or the participants themselves. Although none of these types of conversation blocks were observed.

The research observers recorded information about the scenarios and issues discussed during the review (Table 25). These observations were combined with information obtained from the review minutes to assess the efficency and effectiveness of the review.

Table 26: Information about the review performance recorded by the research observers.

Information	Recording mechanism
Scenario Information	Information about scenarios was recorded during the scenario brainstorming session, the scenario elaboration phase. the scenario selection phase, and the classification of scenarios. In addition to recording a unique scenario identifier for each scenarios, information about the type of conversation block was recorded from the following list. • *Stated.*This was used when an issue was stated or restated. • *Modified.* This was used when the description of a scenario was deliberately modified. • *Combined.* This was intended for use during the scenario elaboration phase. • *Voted on.* This was used during the scenario voting phase. Extra information recorded included the number of votes given to each scenario. • *Direct or Indirect.* Scenarios were classified as either direct or indirect. (See Section 2.1 for more detail.)
Issue Information	Information about issues was recorded during the description of the candidate architecture, during the evaluations of scenarios and during the review of issues. In addition to recorded a unique issue identifier, information about the type of conversation block was recorded from the following list. • *Stated.* This was used when an issue was stated or restated. • *Modified.* This was used when an issue was deliberately modified. • *Agreed.* This was used when all the participants stated that they agreed with the description of an issue. • *Resolved.* This was used when all the participants stated that they agreed with the action needed to address the issue, and with the person responsible for ensuring that the issue was addressed. • *Combined.* This was intended to capture when two or more issues had the same resolution and were therefore combined. • *Priority.* This was intended for use when the issues were reviewed. However, priorities were not assigned to issues during this stage.

The research observers recorded information about the facilitator's activities and indicators of negotiation stumbling blocks (Table 26) to assess impact of the facilitator on the review process.

Table 27: Information about the facilitator recorded by the research observers.

Information	Recording mechanism
Indicators of Negotiation Stumbling Blocks	Selected from the following list: • *Off Topic.* Discussions that were not relevant to the current topic, or more generally not relevant to the review. • *Topic Change.* Used when a change in the topic occurred, such as changing form discussion of one scenario to another. • *Interruption.* Used when one person interrupted another person's presentation or discussion. There had to be a definite interruption, not just a change in the person speaking. • *Disagreement.* Used when one person stated that they disagreed with a statement, or contradicted a recent statement.
Facilitator	The facilitator's activities were selected from the following list: • *Prompts for input from a person* • *Suggests alternatives* • *Encourages topic change* • *Stops domination by one person* • *Stops people from socialising* • *Encourages more discussion*

DISTRIBUTION LIST

Software Design Reviews Using the Software Architecture Analysis Method: A Case Study

Gina Kingston

(DSTO-RR-0170)

AUSTRALIA

DEFENCE ORGANISATION

S&T Program

Chief Defence Scientist FAS Science Policy AS Science Corporate Management	1 shared copy
Director General Science Policy Development	1
Counsellor Defence Science, London	Doc Control Sheet
Counsellor Defence Science, Washington	Doc Control Sheet
Scientific Adviser to MRDC Thailand	Doc Control Sheet
Scientific Adviser Policy and Command	1
Navy Scientific Adviser	Doc Control Sheet and distribution list
Scientific Adviser - Army	Doc Control Sheet and distribution list
Air Force Scientific Adviser	1
Director Trials	1

Aeronautical and Maritime Research Laboratory

Director	1

Electronics and Surveillance Research Laboratory

Director	Doc Control Sheet and distribution list
Chief of Information Technology Division	1
Research Leader, Military Computer Systems	1
Research Leader, Command & Control and Intelligence Systems	1
Research Leader, Command, Control and Communications	1
Research Leader, Advanced Computer Capabilities	Doc Control Sheet
Research Leader, Joint Systems	1
Head, Software Systems Engineering Group	1
Head, Information Architectures Group	1
Head, Information Warefare Studies Group	Doc Control Sheet
Head, Trusted Computer Systems Group	Doc Control Sheet
Head, Advanced Computer Capabilities Group	Doc Control Sheet
Head, Systems Simulation and Assessment Group	Doc Control Sheet
Head, C3I Operational Analysis Group	Doc Control Sheet
Head, Information Management and Fusion Group	Doc Control Sheet
Head, Human Systems Integration Group	Doc Control Sheet

Head, Distributed Systems Group	Doc Control Sheet
Head, C3I Systems Concepts Group	1
Head, Military Systems Synthesis Group	Doc Control Sheet
Head, Systems of Systems Group	Doc Control Sheet
Head, Advanced Network Integrity Group	Doc Control Sheet
Derek Henderson (SSE Group ITD Lab 205)	1
Gina Kingston (MSS Group ITD Fernhill)	1
Philip Collier (MSS Group ITD Fernhill)	1
Mark Grigg C3ISC Group ITD Fernhill)	1
Matthew Phillips (SSE Group ITD Lab 205)	1
Michael Pilling (SSE Group ITD Lab 205)	1
John Hildebrandt (C3ISC Group ITD Fernhill)	1
Mark Daniel (SSD Lab 200)	1
Tony Gray (SSD Lab 200)	1
Stephen Fry (SSE Group ITD Lab 205)	1
Publications & Publicity Officer, ITD / Executive Officer, ITD1	shared copy

DSTO Library and Archives

Library Fishermens Bend	1
Library Maribyrnong	1
Library Salisbury	2
Australian Archives	1
Library, MOD, Pyrmont	Doc Control Sheet
US Defence Technical Information Centre	2
UK Defence Research Information Center	2
Canada Defence Scientific Information Service	1
NZ Defence Information Centre	1
National Library of Australia	1

Capability Development Division

Director General Maritime Development	Doc Control Sheet
Director General C3I Development	Doc Control Sheet
Director General Aerospace Development	Doc Control Sheet

Navy

SO (Science), Director of Naval Warfare, Maritime Headquarters Annex, Garden Island, NSW 2000.	Doc Control Sheet

Army

ABCA Strandardisation Officer, Puckapunyal	4
SO (Science), DJFHQ(L), MILPO, Enoggera, Qld 4051	Doc Control Sheet
NAPOC QWG Engineer NBCD c/- DENGRS-A, HQ Engineer Centre Liverpool Military Area, NSW 2174	Doc Control Sheet

Intelligence Program

DGSTA Defence Intelligence Organisation	1
Manager, Information Centre, Defence Intelligence Organisation	1

Acquisition Program

Director General, Command and Support Systems	1
Director General, Communications Systems	1

Manager, Information Centre, Defence Intelligence Organisation 1

Corporate Support Program
OIC TRS, Defence Regional Library, Canberra 1

UNIVERSITIES AND COLLEGES
Australian Defence Force Academy
Library 1
Head of Aerospace and Mechanical Engineering 1
Deakin University, Serials Section (M list), Deakin University Library, Geelong, 3217 1
Senior Librarian, Hargrave Library, Monash University 1
Librarian, Flinders University 1

OTHER ORGANISATIONS
NASA (Canberra) 1
Info Australia 1
State Library of South Australia 1
Parliamentary Library, South Australia 1
Brett McDowall, OOPL 1
Kevin Nuttall, 1

OUTSIDE AUSTRALIA

ABSTRACTING AND INFORMATION ORGANISATIONS
Library, Chemical Abstracts Reference Service 1
Engineering Societies Library, US 1
Materials Information, Cambridge Scientific Abstracts, US 1
Documents Librarian, The Center for Research Libraries, US 1

INFORMATION EXCHANGE AGREEMENT PARTNERS
Acquisitions Unit, Science Reference and Information Service, UK 1
Library - Exchange Desk,
National Institute of Standards and Technology, US 1

SPARES 5

Total number of copies: **71**

DEFENCE SCIENCE AND TECHNOLOGY ORGANISATION
DOCUMENT CONTROL DATA

1. PRIVACY MARKING/CAVEAT (OF DOCUMENT)

2. TITLE

Software Design Reviews Using the Software Architecture Analysis Method: A Case Study.

3. SECURITY CLASSIFICATION (FOR UNCLASSIFIED REPORTS THAT ARE LIMITED RELEASE USE (L) NEXT TO DOCUMENT CLASSIFICATION)

Document	(U)
Title	(U)
Abstract	(U)

4. AUTHOR(S)

Gina Kingston

5. CORPORATE AUTHOR

Electronics and Surveillance Research Laboratory
PO Box 1500
Salisbury SA 5108 Australia

6a. DSTO NUMBER	6b. AR NUMBER	6c. TYPE OF REPORT	7. DOCUMENT DATE
DSTO-RR-0170	AR-011-217	Research Report	February 2000

8. FILE NUMBER	9. TASK NUMBER	10. TASK SPONSOR	11. NO. OF PAGES	12. NO. OF REFERENCES
N8316-7-34	DST 97/243	CITD	80	36

13. URL of Electronic Copy on the World Wide Web

http://www.dsto.defence.gov.au/corporate/reports/DSTO-RR-0170.pdf

14. RELEASE AUTHORITY

Chief, Information Technology Division

15. SECONDARY RELEASE STATEMENT OF THIS DOCUMENT

Approved for public release

OVERSEAS ENQUIRIES OUTSIDE STATED LIMITATIONS SHOULD BE REFERRED THROUGH DOCUMENT EXCHANGE, PO BOX 1500, SALISBURY, SA 5108

16. DELIBERATE ANNOUNCEMENT

No Limitations

17. CASUAL ANNOUNCEMENT Yes

18. DEFTEST DESCRIPTORS

Software architecture
Software engineering
Reviews
Scenarios

19. ABSTRACT

Software reviews are conducted on most software-intensive Defence projects and are an important component of the software acquisition process. However, software reviews are often conducted in an *ad hoc* manner, and may be inefficient. This report investigates an alternative review process that is based on the Software Architecture Analysis Method (SAAM).

The SAAM review process is driven by the identification of scenarios that capture how the system might be used or modified. It offers potential benefits over the traditional review process in the identification and clarification of requirements, but was less effective at identifying conflicts and trade-offs. Consequently, it is recommended that projects continue to use traditional review processes, and where appropriate, supplement these reviews with SAAM reviews to clarify and identify requirements.