Universidade de Lisboa
Faculdade de Ciências
Departamento de Informática

Evaluating Mobile Collaborative Applications Support of Teamwork in Critical Incidents Response Management

Cláudio Miguel Garcia Loureiro dos Santos Sapateiro

Doutoramento em Informática
Especialidade de Engenharia Informática

2013

Universidade de Lisboa
Faculdade de Ciências
Departamento de Informática

Evaluating Mobile Collaborative Applications Support of Teamwork in Critical Incidents Response Management

Cláudio Miguel Garcia Loureiro dos Santos Sapateiro

Tese Orientada pelo Prof. Luis Manuel Pinto da Rocha Afonso Carriço
Co-orientada pelo Prof. Pedro Alexandre Mourão Antunes e
Prof. Joaquim Belo Lopes Filipe,
especialmente elaborada para a obtenção do grau de Doutor
em Informática na especialidade de Engenharia Informática

2013

Abstract

This thesis contributes to the understanding of the use of Information and Communication Technology, more specifically on the introduction of Mobile Collaborative Applications to assist teamwork in Critical Incidents Response Management. Due the increasing complexity of organizations' socio-technical system, existing work structures and processes may be challenged when they are required to cope with the particular demands posited by unanticipated events. Such events, as for instance, the failure of key organizational resources, may be classified as critical when they entail disruptive consequences for the regular organizational activity. Although, from risks and vulnerabilities assessments organizations may devise business continuity and contingency plans, training programmes and set up teams to address such situations, like service maintenance or help desk teams, an inherent characteristic of the critical incidents considered on this thesis is that they posit novel situations that lead teams to depart from pre-established work arrangements toward an emergent and adaptive behaviour.

Under such work contexts teams often rely on their experience to develop improvised and creative solutions to mitigate the effects of a disruptive event. The development of Team Situation Awareness had been put forward by the related literature as a fundamental asset under this settings.

The inherent affordances brought by mobile devices, namely, situated use and real time information sharing and persistence, lead to the consideration of the use of Mobile Collaborative Applications to assist operational teamwork in cases that the Critical Incidents Response Management endeavour move teams to operate distributed through different locations.

Evaluating the use of Mobile Collaborative Applications on assisting the operational level of teamwork, and their role on Team Situation Awareness development, reveals a challenging research effort, since the more established collaborative technology evaluation methods may reveal short for the considered work contexts. The associated difficulties of conducting field research or achieving a trade-off between the control of the evaluation process without constraining the inherent openness of human behaviour within a collaborative setting, restrict the adoption of more typical evaluation approaches. Moreover, by considering Team Situation Awareness one of the main evaluation dimensions, it should be noticed that it is required to account for the different levels that compose the construct. As it had been debated on the literature Team Situation Awareness should be accounted at both individual and team level and therefore requires the consideration of numerous interwoven factors, that range in the realm of individual cognition to the team processes that bound teamwork.

This thesis puts forward the adoption of a Microworld environment to support *quasi-naturalistic* oriented experiments toward a fine-grain understanding of the use of Mobile Collaborative Applications in operational settings.

Although, the use of Microworld environments as an experimental paradigm is not new, its adoption on (collaborative) software applications evaluation is still emerging. Moreover, it had been noticed from most of the related research works, that Microworlds are typically bounded by specific research aims and lack a frame of reference to inform their development in a more phenomena and domain independent manner, so that they provide a well-grounded experimental instrument.

One contribution of this thesis is the comprehensive specification of the set of foundational building blocks that should guide a Microworld environment development in order to constitute a test-bed for the experimental evaluation of collaborative applications. The demonstration of such specification on informing Microworld environments development is accomplished by its implementation on the selected target application domain that supported the conducted experiments. Although, bounded by the domain characteristics the implementation of the Microworld constitutes also a contribution since it holds a set of software components that are reusable in other contexts of software applications evaluation (particularly those addressing collaborative work support).

The selected target application domain had been the Help Desk Teams operational work enacted on the context of organizational network infrastructures Critical Incidents Response Management, since has it is discussed in this thesis it constitutes a representative domain for the present research aims.

The combination and extension of existing Team Situation Awareness measures and measurement techniques that supported the definition of the devised experiment's dependent variables constitute the third contribution of this thesis.

The fourth contribution of this thesis draws from the results of the conducted experimental trials that yield that Team Situation Awareness had not been enhanced or impaired by the introduction of a Mobile Collaborative Application, but the unveiled usage points that the design of Mobile Collaborative Applications to assist Critical Incidents Response Management teams, should be focused mainly on functional features that support operational information management in disregard to those that assist team management. Such insights could only be achieved by the capability of the Microworld environment to trace team operational work accounting for the context of the team collective task at different levels of granularity. The experimental results show that team management will still be carried out through speech based communications, mainly as team task complexity increases. This result is consistent with several research works, which seems provide some evidence for the validity of the Microworld as an experimental paradigm.

Keywords: Evaluation; Computer Collaborative Work Support; Emergency Response Management; Mobile Computing; Synthetic Environments; Microworld Environments; Team Situation Awareness.

Resumo (Portuguese Abstract)

A presente tese contribui para a compreensão da utilização das Tecnologias de Informação e Comunicação por equipas em contextos operacionais que desafiam as estruturas e procedimentos estabelecidos, definidos para orientar os fluxos de trabalho e partilha de informação. Em particular, o foco é colocado na introdução de aplicações para dispositivos moveis, no suporte ao trabalho colaborativo desenvolvido no âmbito da gestão da resposta à ocorrência de incidentes críticos.

Ocorrências como a falha de recursos organizacionais fundamentais para a regular actividade da organização podem comprometer a sua qualidade de serviço e, nalguns casos comprometer os limites de segurança operacional. Estas ocorrências são classificadas como criticas em situações que acarretam elevadas consequências negativas e que assim sendo a sua contenção e mitigação reveste-se de particular urgência.

Organizações, particularmente aquelas cujo a actividade pode ser mais vulnerável em relação a determinadas ocorrências, realizam análises de vulnerabilidades e riscos, no sentido de desenvolver planos de contingência e programas de formação dirigidos à promoção de uma gestão da resposta a eventos disruptivos mais eficaz.

Nesse esforço são comumente definidas unidades organizacionais, constituídas por equipas, como é o caso das equipas de manutenção e suporte, especialmente vocacionadas para desenvolver as diligencias necessárias no âmbito da gestão da resposta a incidentes críticos.

Neste trabalho são considerados os incidentes críticos que decorrem de situações mais extremas, que incluem ocorrências que em larga medida ultrapassam os planos de contingência estabelecidos e habituais procedimentos, dada a sua natureza inesperada e em certa medida sem precedentes.

Estes contextos requerem às equipas de gestão da resposta a incidentes críticos uma colaboração estreita entre os seus elementos de forma a se adaptar aos imperativos emergentes que decorrem da dinâmica associada à progressão dos incidentes. Esse imperativos levam muitas vezes ao desenvolvimento de acções temporárias e improvisadas, que têm por base a experiência dos elementos da equipa, e que objectivam a contenção e mitigação dos efeitos disruptivos que sucedem da ocorrência do incidente, até que soluções mais definitivas possam ser desenvolvidas.

Contudo a natureza oportunista dessas acções assim como a crescente complexidade da realidade sociotécnica das organizações, restringem a percepção/entendimento colectivo da equipa sobre o alcance das consequências do incidente, bem como, das acções realizadas no âmbito da resposta ao mesmo.

A literatura relacionada aponta que, nestes exigentes contextos operacionais, a promoção da percepção/entendimento da progressão do incidente e das suas consequências, bem como do efeito das acções levadas acabo na gestão da respectiva resposta ao incidente, constitui um activo fundamental que é desenvolvido quer ao nível individual (membros da equipa) quer ao nível colectivo (equipa), de forma a empreender um esforço integrado.

As características inerentes aos dispositivos moveis, nomeadamente, a possibilidade de operar em qualquer local e assim partilhar em tempo real informação operacional, leva à sua consideração para o suporte ao trabalho de equipa, em particular quando estas têm de operar com os seus elementos distribuídos por vários locais afectados pela ocorrência.

Contudo a avaliação de propostas de aplicações moveis para assistir o trabalho de equipa enquadrado neste contexto operacional, mostra-se um desafiante esforço de investigação.

Os mais estabelecidos métodos de avaliação de aplicações de suporte ao trabalho colaborativo revelam-se pouco apropriados dadas as restrições características destes contextos. O controlo subjacente às tradicionais experiências de laboratório dita que estas se revelam mais adequadas para o estudo de fenómenos mais contidos (por exemplo, testes de usabilidade) o que para o presente trabalho se mostra insuficiente dada à riqueza de comportamentos associada aos processos colaborativos e à determinante influência do contexto operacional de trabalho, na utilização efectiva das funcionalidades fornecidas pelas aplicações moveis. Por outro lado, métodos de avaliação baseados em simulações reais ou trabalho de campo (por exemplo, estudos etnográficos) são dispendiosos em termos de custo, tempo e recursos; adicionalmente a avaliação de propostas de aplicações em estados mais iniciais do seu desenvolvimento dificilmente poderá ser suportada nestas abordagens. Embora os métodos de inspecção realizados com recurso a peritos possam em certa medida endereçar este ultimo constrangimento, pela sua natureza não acomodam a questão relacionada com a avaliação do impacto das soluções propostas atendendo ao seu contexto de utilização. Adicionalmente, esses métodos ainda não foram sistematizados para avaliação de aplicações de suporte ao trabalho colaborativo para dispositivos moveis.

Esta tese enquadra o processo de avaliação das aplicações para dispositivos moveis na utilização de um ambiente sintético (Micromundo) que permite, por um lado preservar algum realismo do contexto operacional de trabalho das equipas, e por outro providenciar um ambiente seguro, dotado de algum controlo experimental na aquisição de dados para suportar o rigor da avaliação (embora não restringindo o comportamento dos participantes) e cujo a utilização em termos de recursos e tempo permite o seu uso frequente no ciclo de desenho-desenvolvimento-avaliação que caracteriza o iterativo processo de desenvolvimento de aplicações.

A capacidade de um ambiente sintético capturar a utilização das aplicações moveis de forma contextualizada com o trabalho operacional desenvolvido pela equipa no âmbito de um conjunto de tarefas que promovem a manifestação dos comportamentos individuais e colectivo similares aos exibidos em contextos reais, permite a avaliação de diferentes dimensões do impacto da introdução das aplicações moveis no suporte ao trabalho de equipa na gestão da resposta a incidentes críticos.

No trabalho reportado nesta tese a principal dimensão de analise e avaliação recai sobre como é conduzido o processo de desenvolvimento da percepção/entendimento do estado da situação e da estratégia da resposta à ocorrência do incidente critico; quer ao nível individual quer ao nível colectivo (equipa). Como tem sido debatido na literatura correspondente, a análise e avaliação desse processo requer que sejam tidos em conta vários factores que vão desde dos relacionados com questões do foro cognitivo, aos associados aos processos colectivos relacionados com a colaboração e comunicação.

A utilização de um Micromundo revela-se apropriada para a recolha de dados com diferentes níveis de granularidade, relacionados com actividade, quer individual, quer colectiva, para informar a análise pretendida.

Apesar da utilização de ambientes como o Micromundo, como paradigma experimental não ser completamente nova, na área da avaliação de aplicações colaborativas a sua adopção é ainda uma pratica emergente. A literatura revela que a utilização de Micromundos tem sido guiada por iniciativas especificas em termos de domínio de aplicação ou de investigação de determinados fenómenos de interesse. Como tal, é apontado a falta de um modelo de referência que oriente o seu desenvolvimento de forma a que este constitua um fundamentado instrumento experimental na avaliação de aplicações.

Uma contribuição desta tese é a especificação detalhada do conjunto de constituintes fundamentais, independentes do domínio de aplicação, que devem orientar o desenvolvimento de um Micromundo, de forma a que este possa constituir uma bancada de ensaio que suporte a avaliação experimental de aplicações colaborativas.

A demonstração da concretização da especificação proposta é feita através da implementação de um Micromundo num domínio de aplicação representativo para os objectivos de investigação. O domínio seleccionado foi o da gestão de infra-estruturas tecnológicas de suporte ao sistemas de informação e comunicação organizacionais, em particular abordando o trabalho realizado pelas equipas de suporte e manutenção da infra-estrutura (*Help Desk Teams*), no âmbito da gestão da resposta a incidentes críticos que podem ocorrer sobre a mesma e que são comprometedores da continuidade e qualidade de serviço das organizações cuja a actividade é altamente dependente desta.

Apesar do Micromundo implementado ter tido o referido domínio de aplicação como alvo, a implementação do mesmo, seguindo as especificações do modelo de referência, constitui uma contribuição adicional uma vez que na grande maioria as suas componentes podem ser reutilizáveis no desenvolvimento de novos Micromundos direccionados a outros domínios de aplicação e dimensões de avaliação de aplicações, em particular (mas não só) aquelas destinadas ao suporte de trabalho colaborativo.

A combinação e extensão das actuais medidas (e métodos de medida) que informam a análise sobre os processos de desenvolvimento da percepção/entendimento do ambiente operacional, quer ao nível individual (elementos da equipa), quer colectivo (equipa), apoiaram a definição das variáveis dependentes das experiências realizadas sobre o Micromundo desenvolvido. Entre estas, estão definidas novas medidas que apesar de serem recorrentemente apontadas pela literatura como necessárias, a sua operacionalização tem sido escassa. Pelo que a formulação subjacente sua definição operacional constituí a terceira contribuição desta tese.

A quarta contribuição da tese resulta dos resultados das experiências realizadas. Estes apontam para que a introdução de aplicações moveis no suporte ao trabalho de equipa no âmbito da gestão da resposta a incidentes críticos não favorece a percepção/entendimento, nem individual, nem colectivo (equipa), sobre o estado do contexto operacional de trabalho, contudo, também não se revela intrusivo neste aspecto. Os resultados apontam que a contribuição das aplicações moveis nestes contextos de trabalho, assenta no seu valor acrescido na gestão da informação operacional, uma vez que as equipas participantes nas experiências utilizaram consideravelmente as funcionalidades que suportam o reportar da actividade operacional desenvolvida, em detrimento da utilização de um maior numero comunicações verbais (telefonemas). Contudo as comunicações verbais (telefonemas) continuam a ser o meio de comunicação preferencial relativamente à gestão/coordenação da equipa. Este ultimo resultado, está alinhado com outros que têm vindo a ser consistentemente reportados na literatura relativa à gestão de emergências, e que aponta para a validade do Micromundo como ferramenta experimental. Estes resultados proporcionam uma orientação para futuras iniciativas no desenvolvimento de aplicações colaborativas moveis, para assistir contextos críticos de trabalho operacional; estas deverão atender maioritariamente às questões relacionadas com a gestão de informação operacional.

Palavras-Chave: Avaliação; Trabalho Colaborativo Assistido por Computador, Gestão da Resposta a Emergências; Computação Móvel; Ambientes Sintéticos; Micromundo; Percepção Colectiva da Situação

Acknowledgements

My first and foremost thanks is directed to Professor Pedro Antunes which was the main supervisor of my work until his recent departure for the Victoria University of Wellington. I thank him for his interest, availability, support and initiative and collaboration regarding the publications submission and projects proposals even after his leaving.

This work would not had been possible without the collaboration of Professor Joaquim Filipe and Professor Luis Carriço, who has assumed the role of principal advisor since Professor Pedro Antunes departure. To them may most sincere thanks.

I would also like to express my gratitude to the co-authors of accomplished publications for their precious collaboration and contributions. Particularly, I give thanks to António Ferreira for his continuous interest and support on my research efforts.

A word of appreciation is more than worthy to Eng. Nuno Costa and Eng. Fábio Carreto which had help me in the software development endeavour and had followed unquestioningly the specifications that I provided to them.

I would like to thank also Professor Pedro Cunha and Eng. Pedro Ferreira from the CENI centre, which had provide me the space for conducting the experiments. I extend this more institutional thanks to LaSIGE, which had provided me the materials for such experiments, FCT (Fundação para a Ciência e Tecnologia) and Instituto Politiectnico de Setúbal for the funding that made possible the participation on multiple conferences.

To the consulted help desk teams and all the volunteers that participated in the experiments my deep thanks.

It is unavoidable to thank my colleagues on Escola Superior de Tecnologia from Instituto Politécnico de Setubal for their fellowship in helping me to comply with my professional agenda. A very emphasized thanks to Professor Patricia Macedo and Professor Nuno Pina.

My very hearty thanks is toward my close family and friends for their unconditional support, patient and comprehension of my frequent unavailability either physical or mental. To them, and particularly to Ana, my eternal gratefulness.

At last, thank you, my son for without even knowing, provide me the strength to carry on.

Dedicated to Miguel, Ana,
my wonderful Parents and Sister ,Family, Friends
and to the memory of my dear grandmother Adelaide

Contents

List of Figures

List of Tables

1 Introduction

This dissertation addresses the evaluation of Mobile Collaborative Applications (MCA) in Critical Incident Response Management (CIRM). The fundamental dimension of such evaluation relies on the Team Situation Awareness (TSA) construct, which has been put forward in the research literature as a core team asset in complex work contexts. This chapter presents the motivation for this research and frames the scope of the research effort. It is further introduced the main research question, adopted research approach, research goals and underlying research hypothesis. The chapter concludes with an outline of the remaining chapters of the thesis.

1.1 Motivation and Scope

Organizations are subject to several exceptional events that have the potential to disrupt their regular activity as, for instance, the failure of key operational resources. Whenever an exceptional event occurs, the existing organizational structure and work procedures are stressed and may reveal short to cope with the demands posited by the required response actions, therefore compromising organizational efficiency and efficacy (Rosen, Fiore et al. 2008).

The gap between established work procedures and operational work requirements emerging from exceptional events is not strictly a result of poor or faulty organizational and work structures design but instead an inherent property of the nature of **Critical Incidents** (CI) (Dynes and Quarantelli 1997; Hardeman, Pauwels et al. 1998).

Critical incidents have been characterized as time critical, unwanted, unexpected, with uncertain dynamics, and to some extent unprecedented events that disrupt the regular organizational work's course and bear high consequences (Rosenthal, Boin et al. 2001; Wybo and Latiers 2006). Critical incidents, if not properly contained, may scale to more acute situations as one may find in the literature reporting organizational accidents, disasters and crisis (Shrivastava 1992; Vaughan 1996; Hopkins 2000; Hollnagel 2005).

Organizations typically frame the CI life cycle management through the consideration of four main stages: (1) Mitigation, (2) Preparedness, (3) Response, and (4) Recovery (Kelly 1999; Nasghar, Alahakoon et al. 2005). The first and second stages mainly comprehend vulnerability and risk assessment and, accordingly, focus on devising proper contingency plans and training programs. The third stage consists in the enactment of the operational response, which comprises activities like diagnosis, mitigation, containment, and overcoming the CI disruptive effects. Although these operational activities can borrow from training and contingency plans, a number of additional factors contribute to the enacted operational response, including the available technology support. Finally, the fourth stage of CI life cycle management addresses long-term needs concerning preparedness for future events.

Regarding the present research, the focus has been placed on **providing the means for evaluating the role of Mobile Collaborative Applications (MCA) in the Response stage**.

More specifically, this research **evaluates the role of MCA support of the operational activities of teams** involved in Critical Incidents Response Management (CIRM).

CIRM teams like for instance fire fighters, service maintenance or help desk teams, are constituted by highly knowledgeable experts and are furnished with team processes through training programs intended to leverage team members knowledge, experience and expertise toward an integrated and effective CIRM enactment.

Although CIRM may be to some extent grounded on pre-established strategies defined through contingency plans and training, due to the very emergent and uncertain nature of CI it is not unusual that these teams are required to move beyond established plans, procedures and roles (Turoff, Chumer et al. 2004). Furthermore, it has been pointed that in more extreme contexts, strict reliance on anticipated plans may reveal quite difficult or even impossible, since involved actors continuously renegotiate goals, tasks and resources, supported on informal relationships and their knowledge and experience (Tuomisto 1999; Bruinsma and Hoog 2006).

Accordingly, one may bring up Rasmussen framework regarding human performance (Rasmussen 1983), which distinguishes three Human Performance Modes: skill-based, rule-based and knowledge-based. Under novel and unexpected situations, human performance moves from skill- and rule-based modes to the knowledge-based mode, on which human performance is not limited to following pre-defined procedures but rather entails a pro-active attitude leading to emergent work processes typically characterized by having no best structure or sequence and dynamically evolving (Markus, Majchrzak et al. 2002), which contrasts with the efficiency purported by the other performance modes, that are informed by more stable work requirements (Markus, Majchrzak et al. 2002). This research addresses is directed to the knowledge-based mode of human performance which is the one that better frames the operational teamwork under CIRM settings.

1.2 Problem Statement and Research Question

The development of an effective evaluation paradigm for studying the impact of MCA in CIRM contexts is constrained by a number of challenging factors. By their very definition CIs are unexpected and thus difficult for the researcher to witness. Moreover, safety and security policies highly restrict the researchers access to CIRM teams, especially in the context of their operational work. Such factors highly constrain the ability of the researcher to obtain direct information regarding the phenomena of interest. Even though post-mortem analysis of past CIRM instances is possible, such analysis hardly possesses the sensitivity to capture the operational context where it took place: This is a fundamental concern since, as previously discussed, the operational response is highly contingently bounded.

Also due to safety and security concerns, technology use by CIRM teams has been primarily evaluated through simulation. However, full scope simulations are costly regarding both time and resources and therefore may not be conducted very often.

This work pursues the development of an evaluation approach that, at the same time, preserves as much as possible the operational conditions framing CIRM, and is cost effective in terms of effort and costs associated with the evaluation procedure.

One emphasise in particular, that the evaluation of technological support to teamwork has proved to be a very complex endeavour (Neale, Carroll et al. 2004; Herskovic, Pino et al. 2007; Antunes, Herskovic et al. 2008). Teamwork is highly dependent on various factors such as the team members' cognitive abilities, task characteristics, environmental stressors and teams' collective past experiences. For instance, if team members become information overloaded, caused by an excessive amount of communication or inadequate collaboration support systems (e.g. crowdedly populated user interface screens), they may experience a number of information processing bias (e.g. attention tunnelling, misplaced saliencies) that constrain their higher level cognitive functions (e.g. decision making, situation monitoring) and impair their operational efficiency and efficacy (Endsley, Bolté et al. 2003). These cognitive phenomena are difficult to examine directly in the broader scope of CIRM, since teamwork comprises multiple entangled types of interactions and concurrent activities (Endsley and Garland 2000). **Therefore, a comprehensive evaluation paradigm with applicability in CIRM must accommodate measures with different granularity levels to reflect the individual and collective dimensions of teamwork.** This arguments leads to the definition of the main research question of this thesis:

How to achieve a fine-grain understanding of the role of Mobile Collaborative Applications in supporting teamwork under Critical Incidents Response Management contexts?

1.3 Research Approach, Objectives and Hypothesis

A research approach may be rooted on different philosophies regarding the means for conducting the inquiry and constructing scientific knowledge based on a set of ontological and epistemological assumptions about the phenomena of interest (Cecez-Kecmanovic 2001). Ontological and epistemological assumptions, respectively, refer to the nature of the phenomena of interest (Nandhakumar and Jones 1997; Becker and Niehaves 2007) and how knowledge about the phenomena of interest may be achieved (Galliers 1992).

A number of factors frame the selection of a particular research approach (Trauth 2001). First, one should consider the extent of the overlap between conceptual, descriptive and prescriptive research activities. Second, the feasibility of some research methods in the target application domain. And third, the researcher's skills, background and community often also influence what is thought to constitute an eligible research approach.

Research on the evaluation of Information Systems (IS) has not been rooted on one single overreaching framework. Orlikowski and Baroudi (Orlikowski and Baroudi 1991) outline that research approaches in the IS community have been framed in two dominant philosophical trends, the first grounded on *positivist* and the second on *interpretativism* philosophies. This methodological pluralism is consistent with the results of a more recent literature survey which outlines that 89% of US journals publications and 66% of European journals publications have a positivist undercurrent (Chen and Hirschheim 2004). This study further acknowledges that the European IS research community tends to be more receptive to methods rooted in *interpretativism* than the IS research community in USA.

Reflecting its roots in natural sciences, the positivist research philosophy assumes the hypothetic-deductive logic of the scientific method, framing the research on empirically testable theories, and typically pursuing law-like generalizations based on cause-effect relationships (Chua 1986). *Interpretativism* , which had been widely adopted on behavioural sciences, relies mainly on inductive logic to construct interpretations and explanations about the phenomena of interest (Putman 1983) . Through the interpretative research lens, causality is seen as reciprocal, acknowledging a circular relation between reality and human activity (Creswell 2002).

Applied research initiatives, particularly those in the Human Factors (HF) field had focused on problem-solving interventions through the design of new artefacts toward for instance, work performance or user experience improvements, whether more explicitly or not, much of this research had been oriented according **Design Science** (DS) tenets, which are particularly accounted in this research (Iivari 2007; Purao, Baldwin et al. 2008).

The roots of DS lies in the work of Herbert Simon's The Sciences of the Artificial (first published in 1969) (Simon 1996) where it is argued that **design is an inherent part of the problem solving strategy**. DS has not been specific in terms of its commitment to a particular research philosophy, since it offers a broader perspective on the linkage between theory and practice grounded on the practical relevance of the developed promising solutions.

According with the foundations of DS, knowledge and understanding of a problem domain are achieved through building, deploying and understanding the use of a designed artefact (Purao, Baldwin et al. 2008). Such design involves creativity and innovation (Jay F. Nunamaker, Chen et al. 1990; March and Smith 1995; Hevner, March et al. 2004) and therefore its validation is not straightforward, since a designed artefact may very well change the current status of the problem domain, as well as, its understanding. As so, **theory, design and evaluation should be considered an iterative process**.

Hevner (Hevner 2007) defines a DS framework that comprehensively integrates the main principles of Design Science Research in Information Systems (DSRIS), which concern 1) relevance of designed artefact to the target application domain; 2) rigor underlying the design and evaluation of promising solutions; and 3) iteration of design and evaluation. According with the author, relevance requires practical convergence between the designed artefact and application domain problem requirements, while rigor is achieved by grounding the design on well-established theories. Furthermore, rigor is accomplished by aligning the evaluation method to the research purposes, phenomena of interest and target application domain distinctive characteristics.

This openness to different evaluation methods, ranging from those grounded on *interpretativism* (e.g. observation techniques and structured interviews) to more *positivist* methods (e.g. controlled laboratory experiments and surveys) lead Hevner to frame his framework on the notion of **pragmatism, which is followed in this research**.

Pragmatism posits that the adopted research method for conducting inquiry (e.g. regarding data collection and analysis) is chosen as that which is most likely to provide insights about the phenomena of interest. Thus, pragmatism takes ontological and epistemological assumptions as instrumental to establish a comprehensive understanding on the linkage between theory and practice for a given problem formulation (Creswell 2002; Johnson and Onwuegbuzie 2004). Therefore, different instances of DSRIS may differ in their ontological and epistemological assumptions (Iivari 2007), and may be found in the literature references to *positivist* (Burrell and Morgan 1979), *critical realism* (Carlsson 2006) or *interpretativist* approaches (Niehaves 2007).

Introducing MCA to assist teamwork in CIRM inherently holds a pragmatic dimension, since theory development must be accompanied with artefact design and evaluation. In order to define a baseline for such iterative process, the research objective number one is:

Research Objective 1: ***Define a characteristic set of functional features for Mobile Collaborative Applications to support team processes in Critical Incident Response Management.***

As purported by DSRIS, rigorous knowledge creation requires an iterative design-evaluation cycle supported on proper evaluation methods. The selection of a suitable evaluation method is a chief concern of the present research.

A recent study points out that only 10% of collaborative systems evaluations relied on laboratory experiments (Antunes and Pino 2010). Laboratory experiments are mostly suited for closed and repeatable phenomena (e.g. usability studies), since they provide a controlled and cost effective environment for performing frequent evaluations at several stages of maturity of promising solutions. Nevertheless, collaboration, particularly in work contexts such as CIRM, is associated with the openness of human behaviour in realistic scenarios. The dependence on multiple factors (e.g. individual cognitive skills, task complexity, environmental stressors) makes it more difficult to define an experimental setting for conducting laboratory experiments. It had been noticed however that more systematic and cost effective (both in time and resources) methods rather than the traditional field ethnographic studies should be used to assess teamwork, at least in early stages of team interventions (Paris, Salas et al. 2000; Kiekel, Cooke et al. 2001).

The use of synthetic task environments to support experimental simulations in laboratory with humans in the loop has been put forward as a suitable evaluation paradigm for CIRM work contexts. Variants of synthetic task environments appear in related literature under several designations (e.g. scaled worlds, virtual environments). In the present work it is adopted the **Microworld** designation, since the term has been widely adopted in research addressing concerns similar to those discussed by this thesis.

A **Microworld is a synthetic task-oriented environment that allows studying human behaviour under simulated conditions within a laboratory setting**. Microworlds hold basic real-world characteristics in order to diminish the artificiality of human behaviours, while omitting other aspects deemed secondary for the purposes of the research interests, in order to keep some degree of control regarding the experimental apparatus (Brehmer and Dörner 1993).

Experimental studies supported by Microworlds are not usually extremely long, expensive or user-demanding, and since they afford to collect large sets of experimental data they support both formative and confirmatory purposes, serving at the same time to explore design ideas and to validate promising solutions. The adoption of Microworlds as the fundamental experimental paradigm for this research grounded research objective number two:

Research Objective 2: ***Specify the fundamental building blocks for developing a Microworld environment that reproduces the operational context of Critical Incident Response Management teams, in order to promote that teams behave in a quasi-naturalistic way under the experimental setting.***

In order to address the rigor imperative of DSRIS regarding the evaluation method, the Microworld environment must provide the means to collect representative measures of team behaviour. Toward this goal, **this research had focused mainly on the Team Situation Awareness (TSA) construct.**

Team Situation Awareness (TSA) have been pointed as a critical asset for CIRM teams when their operational activity departs form pre-established procedures (McManus, Seville et al. 2007). Situation Awareness (SA) has been studied in several domains that posit complex, highly dynamic and critical work environments. One may found in the literature various definitions of the construct, as deeper discussed in chapter 3. Based on a synthesis of the most prominent SA definitions (Vidulich, Dominguez et al. 1994) provides a summary definition where, the development of SA is thought as individuals continuous extraction of environmental information and integration of such information with previous knowledge (that is developed through training programs and past experiences) to form a coherent representation of the situation that guide action and support the projection of its evolution. Lower levels of TSA have been reported to lead to an unbalanced CI response e.g. overloading some team members, prioritizing less urgent actions or fail to consider mutually exclusive tasks (Horseley and Barker 2002).

TSA development in spatially distributed team arrangements is highly dependent on the available communication channels that support operational information and activity management (Bowers, Braun et al. 1994; Kanno and Futura 2006; Milis and Walle 2007). Such acknowledgement had grounded the pertinence of the consideration of the use of MCA on the operational level of CIRM. MCA provide a communication channel with some particular affordances: (1) they may be operated *in situ*, which would be of most valuable to CIRM teams operating distributed through several physical locations, (2) they support real time information sharing among several users, which may meet the demands posited by the time criticality and the highly dynamic pace of CI and (3) they support information persistency, which the typical speech communication channels (e.g. phone or radio) used by distributed CIRM teams does not, requiring thus, an additional explicit communication effort for monitoring/accounting the enacted operational activities.

Nevertheless, the introduction of MCA to support teamwork, as in the case of any intervention based on technology introduction, inherently imply a mutual influence between technology design and the team work processes intended to be supported (Bygstad 2005). Team communications and operational activity patterns underlying its collaboration and coordination processes will potentially change often in an unanticipated or even in an unintended and undesirable way. MCA real usage and impact on TSA may depart from those envisioned at design time, due for instance their intrusiveness in operational work practices, the suitability and required operative effort of provided functional features regarding the context in which they are used, such factors are hardly comprehensively anticipated at design time due the exceptional nature of the CIRM contexts.

Research on TSA provides different theoretical formulations of the construct, which in turn, led to a myriad of measures and measurement techniques. A deep analysis of the related literature will root what constitutes suitable measures and measurement techniques that should be accommodated by the Microworld in order to assess the impact of MCA on teamwork. Toward this endeavour the research objective number three is formulated as:

Research Objective 3: ***Establish a set of representative measures and measurement techniques of Team Situation Awareness in Critical Incidents Response Management, which must be supported by the Microworld environment.***

By targeting the research objectives mentioned above on a specific application domain one addresses the relevance imperative posited by the DSRIS. Several factors bound the selection of target application domain. First and of paramount importance, the domain should be representative, in the sense that it offers manifestations of the phenomena of interest. Second, the researcher's acquaintance with the domain drives the extent to which design is more or less informed and thus also leads to practical relevance. And third, the availability of domain knowledge (namely the possibility to witness the operational conditions in which work is accomplished) and accessibility to domain experts also limits the target domain selection.

The selected application domain is Information and Communication Technologies (ICT) infrastructure management. **More specifically, the domain operational teams, commonly referred as Held Desk Teams (HDT), that constitute the organizational units which ensure regular infrastructure service levels.**

The representativeness and relevance of HDT to this research aims are rooted on the acknowledgment that HDT occasionally have to address CI, typically classified as low probability - high impact, since although they do not occur very often, may nevertheless imply significant consequences to business activity. Disruptive events such as major server failures, critical software services break downs (e.g. mail, web proxy, domain naming) or loosing network connectivity that occur from unprecedented and unanticipated factors are perceived as critical to organizations that heavily rely on their ICT infrastructures.

According to (Barret, Kandogan et al. 2004), nearly one third of the operational work of HDT concerns maintenance and troubleshooting. Thus, although a large amount of work accomplished by HDT rely on highly standardized activities (e.g. reconfiguring routers, updating virus-scanning utilities, monitoring service levels), CI may lead HDT to perform beyond established service continuity plans and procedures, to collectively develop creative and temporary workarounds to contain and mitigate the effects of disruptive events.

The development of a set of Microworld based experiments to evaluate MCA usage in helpdesk operations under CIRM contexts constitute the research objective number four:

Research Objective 4: ***Conduct a set of Microworld based experiments to assess the role of a Mobile Collaborative Applications in Critical Incident Response Management performed by helpdesk teams.***

These four research objectives inherently hold the two hypotheses investigated by this thesis, which are stated bellow:

Research Hypothesis 1: ***The introduction of Mobile Collaborative Applications to support Critical Incidents Response Management will drive new ways on how teams develop Situation Awareness.***

Research Hypothesis 2: ***Microworlds provide a valuable experimental paradigm to develop a fine-grain understanding on how teams use Mobile Collaborative Applications in Critical Incidents Response Management.***

Research hypothesis number one draws from the prescriptive nature of introducing a MCA on CIRM teams work context. The underlying assumption is that MCA constitute an additional support for team communications and information sharing. Nevertheless, the extent that MCA are actually integrated in CIRM is an inherent incertitude of any technology design. Therefore, the evaluation process must contrast the use of MCA against the current Team Situation Awareness development support mechanisms.

This understanding is accounted in research hypothesis number two, where it is considered the use of Microworlds as an experimental paradigm. Given the already discussed constrains surrounding the evaluation of interventions in CIRM settings, Microworlds constitute an appealing medium for conducting experiments in a safe and controlled way. The underlying assumption of hypothesis number two is that Microworlds provide an acceptable trade-off between experimental control and ecological validity (Brehmer 2005).

The arguments that support the two stated research hypothesis reflect the claim of DSRIS that applied research inherently encompasses both a prescriptive and descriptive endeavour (Hevner, March et al. 2004).

1.4 Thesis Outline

This thesis is organized in seven chapters. The first, present, chapter frames the motivation and scope for the research work. It was presented the main research question, goals, hypothesis and adopted approach to the formulated research problem.

Chapter number two holds the presentation of the conducted background work, which had consisted on a literature reviewed regarding Critical Incidents aetiology and the nature of human and teams performance in demanding work contexts such as those purported by Critical Incidents Response Management settings.

The third chapter presents research's related work divided on two main sub-sections: the first addresses the existing techniques for measuring Team Situation Awareness, since it had been considered a key dimension of evaluation of technology support regarding teamwork, particularly pertinent, in Critical Incidents Response Management contexts; while the second, overviews existing collaborative systems evaluation approaches and underlying methods.

Chapter number four presents a Microworld Reference Model, and provides a comprehensive specification of the requirements for developing Microworld environments to perform collaborative work support applications evaluation in the context of experimental research.

The preformed immersion on the selected target application domain, Help Desk Teams, is described on chapter number five, where the domain representativeness is discussed regarding the research aims. The chapter further presents, an instance of the Microworld environment bounded by the specifications put forward in chapter four, for supporting experimental trials with help desk teams.

In the sixth chapter the experimental design that guided the conducted experimental trials, as well as, the presentation and discussion of the results of those experiments are provided.

Chapter seven concludes this thesis, by discussing the accomplished work in terms of the devised research goals and hypothesis and pointing out its major contributions and future work directions.

2 Background Work

This chapter is constituted by three main sections. It starts by presenting the main theories and models that had framed the Critical Indents aetiology research; then offers an overview of the research addressing the human factor in critical work contexts and, the third section contrast existing theories on how the research on teamwork in critical work contexts has been approached. The chapter concludes with an integrated discussion of the contents of this three sections.

2.1 Theoretical Underpinnings on Critical Incidents Aetiology

In this section the discussion of Critical Incidents (CI) aetiology is grounded on the existing body of knowledge, that had provided theories and models for explaining the occurrence and escalation of incidents. A number of research communities such as engineering, cognitive psychology and organizational sociology have developed several complementary approaches towards organizational safety, organizational resilience, organizational reliability and organizational accidents models, all providing valuable contributions for CI aetiology and development of CIRM strategies (Perrow 1984; Ferry 1988; Rasmussen 1997; Weick, Sutcliffe et al. 1999; Leveson, Dulac et al. 2006; Woods 2006).

Earlier approaches seek to explain how CI evolve over time. These approaches tend to adopt a sequential description of a chain of discrete events occurring in a particular temporal order. Since the earlier Domino model (Heinrich 1931), several initiatives have adopted a sequential and event-based perspective, namely Failure Modes and Effects Analysis, Fault Tree Analysis, Event Tree Analysis, and Cause-Consequence Analysis (Leveson 1995). These models are grounded on the theory that the CI's progression may be traced back through a temporal sequence of failures like a resource failure or a human operator error. However, this oversimplified view reveals short when CI scale up, mostly because large-scale CI involve multiple intertwined factors, as observed in complex organizational failures (Qureshi 2007).

The Normal Accident Theory (NAT) (Perrow 1984) establishes that organizations, as complex socio-technical systems, inherently entail two main susceptibilities which should be considered when grasping CI aetiology, interactive complexity and coupling. Interactive complexity refers to the presence of a multitude of unfamiliar, unplanned and unexpected interactions within the socio-technical system that bound the work structure and processes, which may not be anticipated in design time and thus are neither visible nor readily comprehensible. Therefore, no one on its own has complete understanding of the overall system's dynamics, particularly when operating in exceptional conditions.

The concept of coupling refers to the interdependency of system’s constituents. Coupling can be either tight or loose. A tightly coupled system is one that its constituents are highly interdependent. Tightly coupled systems tend to respond more quickly to changes, but this will also stand when propagating failures. Conversely, in loosely coupled systems, the system’s constituents work relatively independently, which may refrain a catastrophic failure.

Therefore, the basic argument of NAT is that more interactive complexity and more tight coupling lead systems to unpredictable interactions and insufficient understanding of how to control accidents. In that sense accidents are therefore inevitable, or “normal”.

Several means for reducing failures has been proposed in the literature: identify and reduce as much as possible the unnecessary complexity and the coupling; and accommodate in the socio-technical system several design concerns regarding security, safety, controllability and monitoring (Leveson 1995).

These principles are aligned with the insights brought from studies conducted with High Reliability Organizations (HRO) such as nuclear power plants, air traffic control and fire-fighters (Weick 1987; Roberts 1990; Weick, Sutcliffe et al. 1999; LaPorte 2006). These organizations are characterized as highly reliable because of their track record of consistent and successful contention of CI over long periods of time. The (Weick, Sutcliffe et al. 1999) analysis of HRO has pointed out several common key qualities exhibited by HROs. Human resources are characterized by a strong technical expertise and emphasis is made on continuous training and learning from incidents and their escalation. The design of operational work processes accounts for alternative courses of action, necessary to deal with exceptional demands. Moreover, work echelons exhibit a more collegial structure, empowering the operational levels to cope with unexpected events (Hayes 2006).

HRO principles are usually witnessed in organizations that have prioritized exceptional courses of action in the design of their work structures and processes. However, this attitude may not smoothly generalize to organizations designed around different priorities (e.g. productivity) (Roberts 1990; Weick, Sutcliffe et al. 1999; LaPorte 2006).

Epidemiological theory of organizational accidents had been emphasizing the importance of risk assessment and accidents analysis. Epidemiological theory confronts the oversimplification of the sequential theory by positing that CI progression results from a combination of multiple factors. Reason’s Swiss Cheese model (Reason 1990; Reason 1997) is a major contribution to this theory. It acknowledges that although organizations may instantiate multiple defence barriers and safeguards to mitigate hazards, often CI find a combination of active and latent conditions that open a pathway for CI progression towards catastrophic accidents. Active conditions are, for instance, resource failures, human error, and procedures violations. Latent conditions are those that reside dormant in the organization, often as a result of design errors, which may be triggered by a particular combination of events.

Through this notion of latent conditions, the Swiss cheese model suggests a broader understanding of CI beyond the more obvious sequential paths of events-causes. This broader understanding has been proved to be particularly advantageous in many domains, such as the oil and gas industry (Wagenaar, Groeneweg et al. 1994), commercial aviation (Maurino, Reasonson et al. 1995), and medicine (Reason, Carthey et al. 2001).

An alternative to sequential and epidemiological theories, coming from the fields of engineering and human factors, has fuelled the development of systemic theory on CI aetiology (Rasmussen 1983; Hollnagel and Woods 2005; Leveson, Dulac et al. 2009; Leveson, Dulac et al. 2009).

The premises underlying systemic theory have already fuelled various domains, like cybernetics (Wiener 1961), engineering (Hall 1962), sociology (Buckley 1967), and complexity theory (Kauffman 1995). The central concept of the so called "systems thinking" is that a system embodies a set of elements that are connected together and mutually influence each other to form a whole holding specific properties, beyond the combination of properties exhibit by its parts.

Systemic theory conceives organizations as complex systems formed by human and non-human components that mutually constrain each functional role through a complex web of dynamic relationships and transactions. Accordingly, to fully grasp how CI come to be and scale, one has to depart from simplistic cause-effect explanations, which hardly accommodate the role of dysfunctional interactions between the systems' constituents, towards the consideration of non-linear and complex interactions. In this view, CI scale through the interactions of multiple system's constituents, which may be perfectly functional on their own). This perspective moves the discussion from component failures (in the individual system's constituents) towards the recognition that CI aetiology can be regarded as an emergent phenomena comprising the whole socio-technical through flawed complex interactions between people, processes and technology (Hollnagel 2004; Leveson 2004).

Systemic models such as the Hierarchical Model of Socio-Technical Systems (HMSTS) (Rasmussen 1997), Systems-Theoretic Accident Model and Processes (STAMP) (Leveson, Daouk et al. 2004), and Functional Resonance Accident Model (FRAM) (Hollnagel 2004), have been applied in *post-hoc* analysis of accidents (Johnson and Holloway 2003; Woo and Vicente 2003; Hollnagel, Pruchnicki et al. 2008; Leveson 2008). STAMP had also rooted some initiatives on conducting an accident prone factors assessment in the early stages of systems design (Dulac and Leveson 2004; Dulac and Leveson 2009).

Both STAMP and HMSTS provide holistic frameworks for organizational accidents addressing the interwoven constrains imposed by multiple hierarchical levels of responsibility and activities. These levels comprise, at a higher level, directives emanating from regulatory agencies and, at a lower level, operational procedures. STAMP and FRAM however, provide a more comprehensive formulation regarding the dynamic of the operational level of work.

FRAM drills down operational processes into lower-level functions. These functions are considered in the realm of human, automated and group activities. FRAM posits the analysis of dependencies between functions considering 1) how volatile they are, 2) how much they depend on the context, and 3) their overall rate of change.

STAMP takes a different approach grounded on the tenets of control theory, assuming that at the operational level work processes should have a set of controllers enforcing safety boundaries, and also that such control is exerted by both human operators and technological components. In the case of technological systems, control is embedded by design, whereas in the case of human operators, they are typically developed through training and experience. According with STAMP, CI arise from inconsistent, incomplete, and incorrect process models held by the controllers. Models hold by technological components are inherently limited by design, while in the case of humans they are limited by knowledge and cognitive abilities.

The recognition of the increased cognitive demands experienced by human operators brought forward another theory: Cognitive Systems Engineering (CSE). CSE emerged in the early 1980s specifically concerned with identifying the constrains that shape human operators complex work in critical industries such as nuclear power, aviation and healthcare (Rasmussen, Pejtersen et al. 1994; Vicente 1999) . The CSE precursors were Hollnagel and Woods (Hollnagel and Woods 1999), who introduces the concept of Joint Cognitive System (JCS), where operators and technological components are thought to constitute an integrated unit of analysis.

CSE research have been exploring the dynamic interactions among workers, tools, tasks and structures that make up the working environment, identifying several systemic factors bounding cognitive ergonomics and performance e.g. (Hoffman and Militello 2008; Xiao, Broxham et al. 2010). An exemplary initiative is the Cognitive Reliability and Error Analysis Method (CREAM) (Hollnagel 1998). CREAM provides guidance for identifying conditions (e.g. work procedures, work environment, collaboration structures) framing the operators' reliability. Reliability is categorised through four performance control modes: Strategic, Tactical, Opportunistic, and Scrambled defined according the situation demands imposed to the operator. In the Strategic mode, actions are well informed, since the work conditions are not demanding; it is thus the most reliable mode. At the opposite extreme, in the scrambled control mode, actions are considered to have little or no reflection, because of the demands emanated by the critical work context; it is therefore the most prominent operational mode for failure.

The next section reviews the main research fields addressing human performance under critical work contexts.

2.2 The Human Factor in Critical Work Contexts

The study of processes in which human beings interpret and make sense of required work activities under demanding work conditions had constituted on of the core tenets of cognitive and social psychology. Human behaviour results from a complex interplay between external demands, internal psychological processes and social context (Gasson 1999). The external demands imposed by unexpected situations challenges the actors' internal cognitive structures that represent the external reality (Gasson 2004). These structures are often referred as schemas (Bartlet 1932; Neisser 1976) or Mental Models (MM) (Johnson-Laird 1983). The mental codification of experience constitutes a particular way of perceiving and responding to the stimulus purported by the work context. Therefore MM establish the knowledge structures, cognitive representations and mechanisms used by humans to predict, explain, describe, recognize, and instantiate their behaviour (Mathieu, Heffner et al. 2000; Paris, Salas et al. 2000).

When handling new information, it may be simply accommodated on existing MM or it may dictate a structural change of the existing MM. Humans inherently seek to avoid ambiguity between what is experienced and their MM. This so called “cognitive dissonance” has been evident in strategies like “ambiguity aversion” (Heath and Tversky 1991; Camerer and Weber 1992), “maintaining the felling of competence” and “coping with cognitive overload” (Dorner 1997). These strategies, mostly unconscious, often lead to information processing and decisions bias. Typical decision bias consistently reported by the related literature includes: “tunnel vision”, leading to fail to notice information that is inconsistent with their MM; focus on conforming/reinforcing information and distorting information to fit the current MM (Kahneman, Slovic et al. 1982; DeKeyser and Woods 1990); oversimplifying a complex situation (Sterman 1994); and fitting current situations to past knowledge (cognitive conservatives (Reason 1990)). (Tversky and Kahneman 1974) also point two heuristic principles that explain why humans seek “cognitive economy”: 1) representativeness, where the situation is assessed against similar experiences; and 2) availability, where the situation is assessed relying on the easiness that similar experiences can be retrieved from memory. The level of adherence to these strategies is closely related with the demands posited by the situation (and their criticality) and the operator’s ability to cope with the associated stress (McGrath 1976).

In CIRM scenarios, the operators divide their attention between the task execution and the continuous scan of the environment to update Situation Awareness (SA). Background control (Dorner and Pfeifer 1993) is the mechanism by which the cognitive system alternates between concentration (attention focus on the primary task, suppressing distractions) and situation assessment (scanning the environment on a regular basis). Background control happens mostly without conscious planning. If a task is very demanding, or if the stress level rises, background control may be reduced or completely abolished, which may compromise current SA.

To minimize information processing and decision bias arising from flawed SA, several frameworks specifically developed to guide information retrieval and assessment in critical work contexts had pursued to furnish operators with a conscious process to lead the information processing and decision-making processes. These frameworks provide a set of orientations ranging from more general ones e.g. the Observe-Orient-Decide-Act (OODA) framework (Boyd 1996; Osinga 2007) to more comprehensive e.g. Detect-Estimate-Choose-Identify-Do-Evaluate (D.E.C.I.D.E.) (Benner 1975) and Facts-Options-Risks-Decision-Execution-Check (FOR-DEC) (Hoermann 1995).

However, research on naturalistic decision-making (NDM) has shown that experienced people under critical work contexts rarely process information and make decisions grounded on such normative orientations (Klein 1998; Grant and Kooter 2005). According with NDM, people do not decompose information to derive a set of options and compare those to a criteria set, because a criteria set cannot be defined given the complexity and uncertainty of the exceptional demands that bound the unexpected demands of critical and complex work context; plus the salience of information frequently change according with the specific situation pace and emerging contingent requirements.

Opposing the normative models, the most popular NDM model is the Recognition-Primed Decision model (RPD), a descriptive decision-making model proposed by (Eisenberger 1993) based on field observations and interviews with fire fighters, intensive care nurses, surgeons, military field commanders, and pilots. In scenarios characterized by time constrains, lack of information, and dynamically changing goals, the information processing and decision-making process is not a once-through process searching for the best option but rather a cyclic process where the aim is to choose an acceptable option for the course of action and then improve upon the observed consequences.

Accordingly, much of the research in NDM concerns assessing the situation as it evolves over time (Zsambok, Beach et al. 1992). The typical principles underlying NDM models are: 1) focus on situation assessment; 2) single option construction, evaluation and modification (instead of alternative generation and selection); and 3) deeply relying on previous experience. NDM models emphasize that previous experience delivers a referent which is triggered by current contextual cues. These cues will recall previous experience and will dictate courses of action by suggesting critical things to look for, typical actions to perform, and their feasibility. This process is typically iterative because expectancies may be violated and readjustments in the course of action have to be made. Considering the main NDM tenets, SA development plays a fundamental role in bridging the gap between referent experience and the current experienced situation.

The study on how individuals overcome the gap between their referent and current situation status has also rooted the development of the sensemaking theory (Weick 1988; Grant and Kooter 2005; Landgren 2007; Muhren and Walle 2009). The sensemaking theory claims that environmental cues are extracted, interpreted and revised based on the enacted actions and their consequences, since from a sensemaking perspective action often precedes understanding. According to this analytical lens, people experiencing ambiguity and uncertainty actively interact with the environment in order to create meaning and make retrospective sense of what occurs (Weick and Meader 1992). These interactions give insights about the situation and support the development of an acceptable explanation of the situation. Therefore, it is the interplay between interpretation and action that updates individuals SA. This theory also posits that acting and interpreting are a collective effort, since teams can see more ways and devise more interventions to cope with critical and unprecedented incidents. Therefore the capacity to extract cues, interpret events, communicate, act and provide feedback are fundamental requirements in critical work contexts.

The following section reviews and discusses how team interactions had been studied regarding teamwork requirements to perform in critical work contexts.

2.3 Teamwork in Critical Work Contexts

Critical work contexts such as those that frame CIRM often assume a collaborative dimension because teams can leverage expertise, knowledge, experience, and information processing abilities. The teams' ability to better perform in critical work contexts has been reported in terms of productivity, improved decision making, efficiency under stress and reduced number of operator errors (Orasanu and Fisher 1997; Paris, Salas et al. 2000).

In teamwork research literature, a widely adopted definition of what is a team is "a distinguishable set of two or more people who interact dynamically, interdependently and adaptively toward a common and valued goal/object/mission, who have each been assigned specific roles or functions to perform" (Dyer 1984; Salas, Dickinson et al. 1992). Teamwork, as defined in (Wilson, Salas et al. 2007), is "a multidimensional, dynamic construct that refers to a set of interrelated cognitions, behaviours and attitudes that occur as team members perform a task that results in a coordinated and synchronized collective action".

From the above definitions, one can envision multiple intertwining levels of teamwork, which ground different dimensions for analysis. Research on teamwork in complex environments has mainly focused on either the cognitive demands posited to team members at individual level (e.g. information overload, decision making and cognitive bias), or the team processes (e.g. the underlying activities by which collaboration and coordination are accomplished) and their relationships with team performance (e.g. through overall efficiency and efficacy measures).

Conceiving teams as an integrated unit has fuelled a team level perspective of functions that traditionally lay at the individual domain. Teams can be considered as information-processing units (Hinsz, Tindale et al. 1997) in a manner analogous to early views of human cognition e.g. (Newell and Simon 1972). For instance, (Cooke, Salas et al. 2000) argues that teams perceive environmental cues, assimilate and combine information, solve problems, make decisions, plan courses of action, and develop performance abilities as an integrated unit. This conceptualization leads to the emergence of the T*eam Cognition* construct on teamwork studies addressing particularly complex work environments (Cooke, Gorman et al. 2008; Salas, Fiore et al. 2010). The underlying theory of team cognition reflects the intent of extending individual level cognitive concepts like mental models and associated cognitive processes such as interpretation to team level, in order to develop a more informed understanding of its performance. A paradigmatic example of such reasoning can be found in the conception that teams hold a (collective) memory structure, which is accessed through team processes (e.g. collaboration) supported by team communications (Wegner 1986).

Team cognition has been adopted by a large body of research and has supported research on teamwork in critical environments over the past 20 years (Cooke, Salas et al. 2004; Cuevas, Fiore et al. 2007). As so, the review of its main theoretical positions is unavoidable.

Two competing theoretical perspectives on team cognition can be found in the related literature. The first approaches team cognition by looking to teams as information processing units and adopted the Input-Process-Output (IPO) framework. This stream of research highlights the role of Shared Mental Models (SMM) on performance (an extension of the role of MM in individual performance). SMM refer to the extent that team members have the same understanding of team goals, the roles and functions of each team member, nature of the task, use of resources, environment status, work demands, and required action strategies (Cannon-Bowers and Salas 1997; Klein 2000).

It has been argued that SMM provides a valuable coordination mechanism for effective team functioning, by allowing team members to forecast the behaviour of other team members and facilitate team communication (Cannon-Bowers, Salas et al. 1993; Stout, Cannon-Bowers et al. 1999; Langan-Fox, Code et al. 2000). For instance, when communications channels are limited, SMM allows team members to operate from a common frame of reference and anticipate other team member information requirements without much communication, as demonstrated in (Cannon-Bowers, Salas et al. 1993; Endsley 1995).

Despite numerous works reporting positive impacts of SMM, for instance increased team performance (Entin and Serfaty 1999; Mathieu, Heffner et al. 2000), others, recurring to MM manipulation (e.g. through cross-training and task information availability to each TM), have concluded that SMM does not have a direct impact on team performance, particularly in teams having highly specialized role structures (Levesque, Wilson et al. 2001; Cooke, Kiekel et al. 2003).

These mixed results reflect the variability of the task and team structures. For instance, work settings requiring more heterogeneous teams with very specific expertise and roles, the full overlap of knowledge may not be required or even desirable (Salas, Stagl et al. 2007). In fact, leveraging diverse knowledge and skills may well be the primary impetus for setting up teams in critical work contexts. Also, SMM is more likely to be achieved in highly structured tasks than in unanticipated work scenarios (Kraiger and Wenzel 1997).

The acknowledgement that, in critical work contexts, MM are continuously updated through interactions between the team members and between the team members and the working environment, lead to another approaches on team cognition, which highlights the need to integrate a more dynamic consideration of the interplay between team members (Burke, Stagl et al. 2006).

This more dynamic view of team cognition borrows from developments in alternative perspectives of cognition. The traditional cognitive views regard human action as the outcome of mental processes (Haugeland 1985). Arguing that this conception lacks an understanding of the context or situation, other paradigms like Situated Cognition (SC) (Lave 1988) (or situated action (Suchman 1987)) and Distributed Cognition (DC) (Hutchins 1994) have emerged. Under these paradigms, human action is no longer seen as an exclusive outcome of mental activity but rather as a closely intertwined relationship between mental activity, supporting artefacts and work context.

SC emphasizes the emergent, contingent and improvisational nature of human activity. Mental activity grows directly out of the particularities of a given situation on a moment-by-moment basis, driven by the interactions between actors, and between actors and the environment. Both (Suchman 1987) and (Lave 1988) point out several shortcomings of the more traditional trend under which problem solving is seen as a series of objective, rational and pre-specified means to ends, and argues that the unit of analysis should not be the isolated individual, but the relationships between the individual and the environment.

Although this individual-environment dyad as unit of analysis is aligned with the perspective of DC, DC additionally incorporates a strong emphasis on the role of artefacts that mediate human action. It is a central tenet of DC that knowledge is distributed at the system level, which encompasses individuals and the artefacts they use (Flor and Hutchins 1991). It is therefore fundamental to understand the interactions between the individuals and the artefacts that support their work (Hutchins 1991). Under this lens, the particular characteristics of the artefacts are important to understand human behaviour and interaction.

Both SC and DC have not been absent of criticisms. SC models have a slightly behaviouristic undercurrent in that the subject's reactions to environmental stimulus that determine action. On the other hand, DC views people and artefacts as conceptually equivalent in the sense that they jointly hold and propagate information. This symmetry between humans and artefacts has been criticised by who considers that motive and consciousness belong only to humans. Humans make use of their knowledge in self-initiated ways, according to socially or personally defined motives, while artefacts can only programmatically process information. Artefacts are at most aids for human cognition (Bodker 1989).

Despite these criticisms, the SC and DC perspectives have significantly contributed to the study of cognition in critical work contexts by highlighting the situated nature of human activity in such settings, and stressing the role of artefacts and task/situation contingencies as determinants of the course of action. Such ecological framing is in line with the tenets of ecological psychology. Ecological psychology introduced the concept of affordances as properties of the artefacts and the environment that unveil opportunities/possibilities for action (Gibson 1966). It is also through the perception and interpretation of affordances that we can understand team behaviour in the working environment (Turvey and Shawn 1995).

More recently, in an effort to integrate existing theories in a comprehensive and coherent framework, a number of researchers have recovered the Macrocognition (MC) concept (Letsky 2008; Patterson and Miller 2010). The study of cognitive work in critical work contexts, such as, aviation and nuclear power plants (McNeese 1986; Woods and Roth 1986), had set the foundations of MC (Hoffman and McNeese 2009).

The term macro contrasts the traditional "micro" view of individual cognition (e.g. whether attention is parallel or sequential) studied through carefully and highly directed laboratorial methods (Crandall, Klein et al. 2006). MC suggests moving the study of cognition from overly artificial experiments that evaluate cognitive phenomena in isolation, towards understanding cognition manifestations in the context of real tasks. As an illustrative example, one may consider that the decision-making process is enacted differently according to the context under which it is performed. It is quite different to make a onetime decision in a laboratory experiment than to perform the same decision under real world constraints, involving for instance time pressure without a stable evaluation criteria immersed in a highly dynamic, complex and stressful environment, as discussed in the previous section.

One must however remark that micro and macro cognition are not opposed or divorced but rather complementary. Both, an in-depth understanding of individual cognitive components and processes as well as their manifestations under naturalistic contexts may well provide a richer understanding of cognition.

MC theory was built upon the work of (Cacciabue and Hollnagel 2005), which emphasizes that in order to unveil the actual response of individuals and teams in a critical work context one must consider the affordances of the environment in which work is enacted. The process in which MC may be used to study teamwork is described in (Fiore, Smith-Jentsch et al. 2010). The authors depict teamwork by considering that it involves two high-level constructs, *internalized team knowledge* and *externalized team knowledge*, and two high level processes, *individual knowledge building* and *team knowledge building*. Knowledge is considered emergent in the sense that it is created through the team members' interaction, and tightly coupled with the task and work environment.

According with this formulation, internalized knowledge refers to knowledge held in the individual minds of team members, encompassing both knowledge which is overlapped among team members and non-overlapping knowledge, i.e., specialized knowledge possessed at individual level. Externalised team knowledge encompasses the integrated information that has been made actionable and explicitly agreed upon at team level from individual and team knowledge-building processes. Knowledge-building processes comprehend both team level knowledge building, accomplished through the actions taken by team members to process and disseminate information, and to transform that information into actionable knowledge. On the other hand, individual knowledge building comprehends the actions taken by individuals in order to process and organise their own knowledge. The theory purports that these processes unfold at a multi-level, individual and team, in a parallel, interdependent, and iterative way. Table 2.1 summarizes the main aspects of the theory.

Table 2.1: Macrocognition constructs and processes (adapted from (Fiore, Smith-Jentsch et al. 2010)).

		Description
Process Level	Individual Knowledge Building	Includes actions taken by individuals to build their own knowledge (inside the head or overt actions)
	Team Knowledge Building	Includes actions taken by team mates to disseminate the information and transform it to actionable knowledge for team members
Knowledge Level	Internalized team knowledge	Refers to the collective knowledge held by each team member.
	Externalized team knowledge	Refers to facts, concepts and their relationships which have been agreed upon by the team

MC research has put forward two main approaches for basing the inquiry methods. One, due to its naturalistic claims, regards field research as the richer approach to study MC. On the other hand, some authors, recognizing the difficulties and constrains associated with field research, and that a deeper understanding of team performance requires the elicitation of operational data at multi-level (individual and team) with different granularities, had acknowledge that the degree of control afforded by laboratory settings should constitute a valid alternative provided that they reproduce some representative properties of the work environment that promote the manifestation of teams natural behaviours (Fiore, Smith et al. 2008).

2.4 Discussion

The discussed models and theories explaining CI aetiology indicate that organizations, as complex socio-technical systems, inherently hold latent conditions that, if stressed by particular demands, may provide the sources for systemic failures. Systemic models of CI aetiology offer a broader consideration for the multiple interwoven factors that bound CI occurrence and escalation. They depart from the perspective that CI may be rooted in failure of organizational resources — and the more traditional emphasis is put on human error — by acknowledging the role of dysfunctional interactions among the different constituents of a complex socio-technical system (humans, technology and work processes).

Flawed organizational processes and dysfunctional interactions often manifest at the operational level of organizational work. From the discussed theories and models, FRAM and STAMP are those that more systematically investigated this aspect.

FRAM appears to be more suitable for tightly coupled and relatively stable work processes on which both the functions, the sources of variability, and the relationships between functions are fairly known. Given the dynamics of CIRM, this view misses the intrinsic emergent and contingent nature of the operational requirements of teamwork in such contexts.

STAMP offers a more comprehensive conceptualization about how operational work is carried out in dynamic contexts. STAMP posits that operational activities are kept within safety boundaries as long as the controllers hold a mental model of the process. STAMP claims that control is exerted either by technological components or by humans. In the case of humans, STAMP suggests that situation awareness constitutes a fundamental asset to keep accurate control.

The extent that human operators develop accurate situation awareness is inherently constrained by their cognitive limitations. As it was pointed out by NDM research, the individuals' past experience constitutes a key referent for understanding their activities. Although, this hypothesis holds that individuals take advantage of previously acquired knowledge and this way reduce their cognitive load and enhance their performance, it also constitutes a source of various cognitive bias, which may impair accurate situation awareness.

Given the claims brought by SC and DC, mobile collaborative applications may be hypothesised to provide valuable resources for maintaining accurate situation awareness. Moreover, it has been consistently emphasized that team communication has a paramount role supporting teams making collective sense of unprecedented situations.

The need to understand cognition beyond the internally held mental referents, towards the consideration for the context under which individuals perform, has been emphasized by the tenets of SC, DC, MC and CSE approaches to CI aetiology, such as CREAM. Therefore, to develop a fine grain understanding of the role of mobile collaborative applications in CIRM settings, one must depart from the more traditional usability, user experience and satisfaction studies towards the consideration of technology usage when immersed in a representative work context. According to sensemaking theory, individuals interact with the environment through their actions to make sense of on-going situations.

The adoption of Microworlds to evaluate mobile collaborative applications in CIRM appears to be a suitable alternative when field studies are inaccessible and more traditional laboratory experiments reveals inadequate due their single focused nature and excessive control. If cautiously designed, Microworlds may promote manifestations of teamwork close to those found in real-world settings.

Rooted on the principles of macro-cognition, studying teams' performance in critical work contexts requires the consideration of individual and collective behaviours framed, as much as possible, in naturalistic work settings. Therefore, in order to provide a quasi-naturalistic environment, a Microworld must accommodate the following high-level requirements: 1) Provide a complex and dynamic task environment over which teams enact their activities; 2) Account for the interplay between individual and team levels; 3) Reproduce some of the artefacts used by teams in the real-world working environment; and 4) Promote some of the stress conditions found in the real-world working environment.

3 Related Work

Two major sections constitute this chapter. The first, presents the most prominent theoretical proposals on the Situation Awareness construct formulation that had guided the Situation Awareness measurement techniques also herein presented and discussed. The second part of the chapter, provides an overview and discussion of existing collaborative applications evaluation dimensions and methods.

3.1 Measuring Situation Awareness

Since the late 1980s, situation Awareness (SA) became very popular within the Human Factors (HF) community and many researchers began investigating the topic in a whole host of different domains as, for instance, military aviation (Endsley 1993), civil aviation and air traffic control (Kaber, Perry et al. 2006), health care (Hazlehurst, McMullen et al. 2007), and emergency services (Blandford and Wong 2004).

SA measures and measurement techniques have been devised from different formulations of what SA is about. Early definitions considered SA as a cognitive product, in the sense that SA corresponds to a set of knowledge elements hold in the mind regarding the task environment. This lead to measuring SA through knowledge elicitation techniques. Another line of thought considers that SA also entails a process dimension since its development is grounded on the interactions of individuals with each other and with the task environment, both most often mediated by artefacts that support individual and team work. Thus SA measurement should further be achieved by tracing the work processes.

The debate in the related research literature acknowledges that these are complementarity perspectives and one should account for both product and process dimensions of the construct. The following section discusses the most prominent SA models from which existing SA measurement techniques were devised.

3.1.1 Situation Awareness Models

The mature research on SA comes from individual-oriented models. These models are reviewed first and then those that extended the concept to team level which had put forward different dimensions of Team Situation Awareness (TSA).

3.1.1.1 Individual Situation Awareness

To date, the most widely cited model on SA is Mica Endsley's three-level model (Endsley 1995), presented on Figure 3.1. Endsley posits that SA is achieved through three cognitive levels: Level 1 (Perception) concerns the perception of critical elements in the environment; Level 2 (Comprehension) concerns understanding their meaning in context; and Level 3 (Projection) considers the generation of possible future states and events.

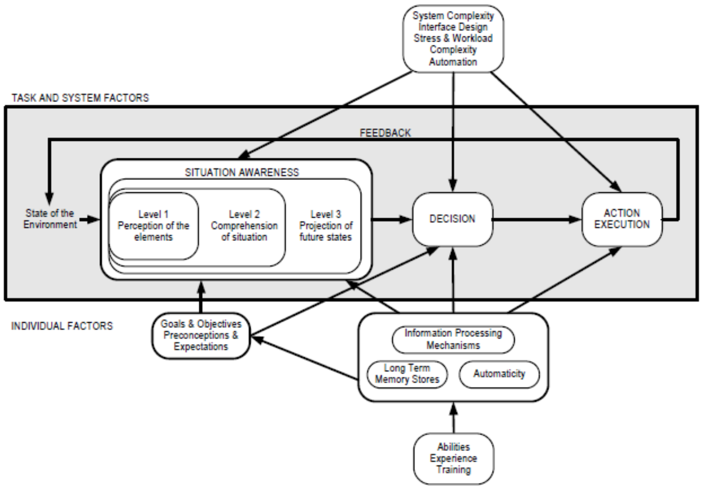


Figure 3.1: Endsley three-level models of Situation Awareness (from (Endsley 1995))

One fundamental characteristic of Endsley's model, which fuelled its wide acceptance, is a clear division of SA in hierarchical levels, which helps disentangling the cognitive phenomena in different levels of analysis that can be used for systems' design and evaluation (Endsley, Bolstad et al. 2003; Endsley, Bolté et al. 2003).

Nevertheless, the complex relationships between the individual and the environment are not comprehensively addressed by the model, since by mainly addressing the cognitive process dimension of SA, Endsley clearly distinguishes SA from the situation assessment process. Although, "perception of elements", "understanding of meaning", and "projection of future states", could be further rooted to external behaviours, the model frame them in internal cognitive activity and lack a more comprehensive definition on how they are accomplished.

This shortcoming is addressed by Bedny and Meister model (Bedny and Meister 1999). The authors rooted their model on activity theory (Leontiev 1974). Activity theory purports that individuals represent an ideal image or desired end state of an activity and perform actions that direct them towards the desired end state. This emphasizes a more dynamic perspective on how SA is developed, by positing that action is motivated by the disparity between the one's goals and the current perceived situation. Overcoming the experienced disparity comprises three stages: 1) Orientational (development of an internal conceptual model of the situation); 2) Executive (proceeding to the desired goal via decision-making and action execution); and 3) Evaluative (assessing the feedback and reframing the Orientational and Executive stages). This model is, to some extent, more comprehensive than Endsley's model by suggesting how such stages are accomplished using the set of functional blocks depicted in Figure 3.2.

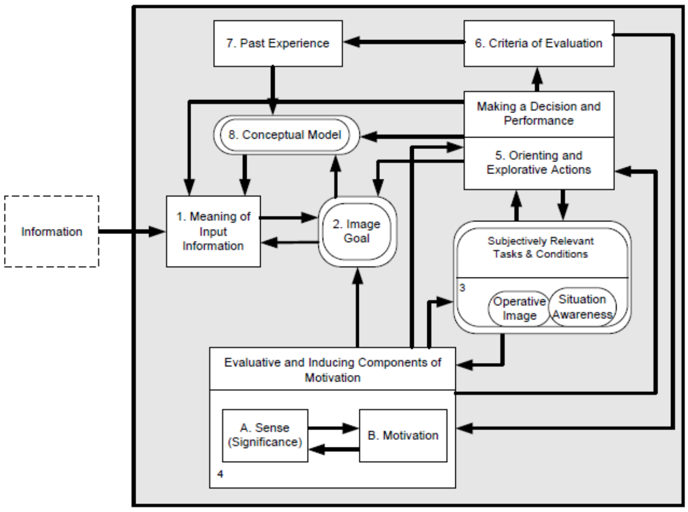


Figure 3.2: Bedny and Meister Activity Theory based Situation Awareness model (from (Bedny and Meister 1999))

According with this model, the interpretation of incoming information (function block 1) is influenced by the individual goals (function block 2), the conceptual model of the current situation (function block 8), and the past experience (function block 7). Interpretation then modifies the individual goals, experience, and conceptual model. Critical environmental features are identified (function block 3) based upon their significance to the task goals and the individual's motivation towards the task goals (function block 4), which on the other hand directs the individual interaction with the world (function block 5).

The extent to which an individual proceeds towards the task goals is determined by the goals (function block 2) and the evaluation of the current situation (function block 6). The resultant experience derived from interaction with the world is stored as experience (function block 7), which in turn informs the conceptual model (function block 8). Still according to this model, the core processes involved in the acquisition of SA are the conceptual model (functional block 8), the image-goal (functional block 2) and the subjectively relevant task conditions (function block 3).

Although, as aforementioned, the model provides a more comprehensive perspective on how SA is developed, it has received less attention than Endsley's model, possibly because its formulation does not smoothly inform the definition of SA measures, and therefore it reveals less practical. Furthermore, the description of the nature of the interactions between the individual and the environment lacks a deeper formulation.

This later aspect has been addressed by Smith and Hancock (Smith and Hancock 1995). Smith and Hancock developed an ecological approach offering a broader formulation of SA characterized as a "generative process of knowledge creation and informed action taking" (1995, p. 138). Their definition is based upon (Neisser 1976) perceptual cycle model, which relates interaction to mental models. According to this model, an individual's interaction with the world (termed exploration) is directed by the internally held schemata. The outcome of an interaction modifies the original schemata, which in turn directs further exploration. This process continues in an iterative cycle.

Smith and Hancock also argue that the schemata accommodates past experiences that facilitate anticipation and provide guidance for future activities. Any unexpected occurrences will trigger exploration and explanation, which in turn update SA. The perceptual cycle model of SA is presented in Figure 3.3.

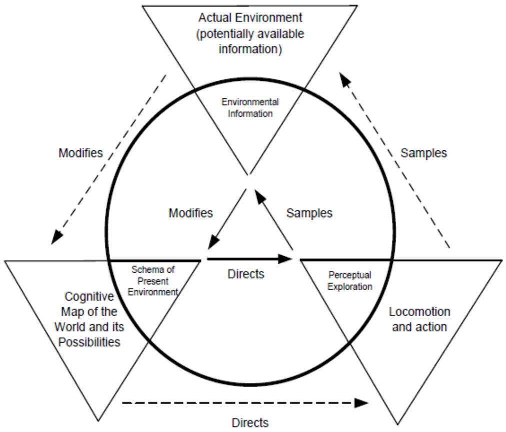


Figure 3.3: The perceptual cycle model of Situation Awareness (from (Smith and Hancock 1995))

Similarly to Bedny and Meister and unlike Endsley, which clearly separate the SA product from the processes underlying its development, SA is formulated through the consideration of its process and product dimensions. In order to fully grasp the nature of the construct, the study of the links between process and the product dimensions is most likely to yield a more comprehensive understanding on how situation awareness is developed (Stanton, Chambers et al. 2001).

However, judging the validity of these models is not straightforward. Aside from Endsley's model, which has been systematically used, the remaining models still lack empirical evidence; and it may be questioned whether or not a testable hypothesis could in fact be unequivocally generated, considering the complexity of the proposed theoretical claims (although this is perhaps a criticism that can be extended to many SA models).

Nevertheless, the complementary SA formulations offered by the reviewed models provide valuable guidance for evaluating SA, as discussed in section 3.1.2.

3.1.1.2 Team Situation Awareness

The formulation of TSA is indubitably more complex than individual SA. As pointed out by (Salas, Prince et al. 1995), TSA goes beyond merely summing individual SA. Ostensibly TSA is multi-dimensional, comprising not only the sum of individual SA but also the combined SA that the whole team possesses as an integrated unit.

One line of research has approached TSA through the notion of Shared Situation Awareness (SSA), which considers TSA as the degree of overlap between individual SA (Nofi 2000; Bolstad, Cuevas et al. 2005).

Acknowledging that efficient teams may not require that all members share the same knowledge, a further distinction between SSA and TSA was made (Endsley and Robertson

2000). While SSA is still defined as the level of overlap between the team members' SA, TSA, on the other hand, is defined as the degree to which every team member possesses the required SA to accomplish the team's task. In this way TSA is understood as partly shared and partly distributed knowledge (Stout, Cannon-Bowers et al. 1999; Shu and Furuta 2005), as depicted in Figure 3.4.

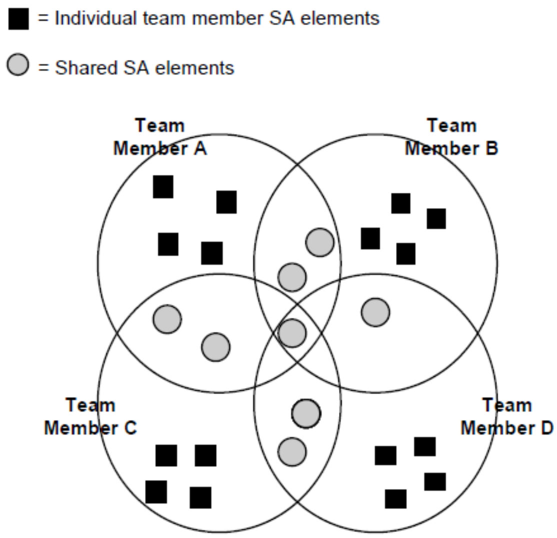


Figure 3.4: Shared Situation Awareness versus Team Situation Awareness (adapted from (Endsley and Robertson 2000))

Nevertheless, this definition of TSA still draws heavily on the product dimension of the construct, being more concerned with the SA held by the individuals than with the consideration of the processes by which it is achieved.

A broader conceptualization of TSA requires understanding that teams have specific structural properties (e.g. assigned roles) and dynamics (e.g. task interdependencies) that inherently bound the information flows among the team members and therefore shape how TSA is developed as a collective asset (Fiore, Salas et al. 2003).

Salas et al (Salas, Prince et al. 1995) posit that the interactions among team members mutually inform each other's SA. These interactions are thought to comprehend the four main interwoven factors presented in Figure 3.5. Two of such factors rely in the realm of individuals: information processing (e.g. comprehension of exchanged information and attention allocation); and mental models (including the cumulative knowledge of past experiences, training and team culture). The two other factors are positioned at the team level: team characteristics (e.g. cohesion, trust, and maturity); and team processes (e.g. collaboration).

Acknowledging the constructivist nature of the SA process, Salas et al. emphasize the paramount role of team communications by recognizing that individual SA limitations can be supressed through information exchanges. In fact, significant research has focused on team communication as the key dimension of analysis for TSA e.g. (Kiekel, Cooke et al. 2001; Salas, Burke et al. 2001; Kennedy and McComb 2010).

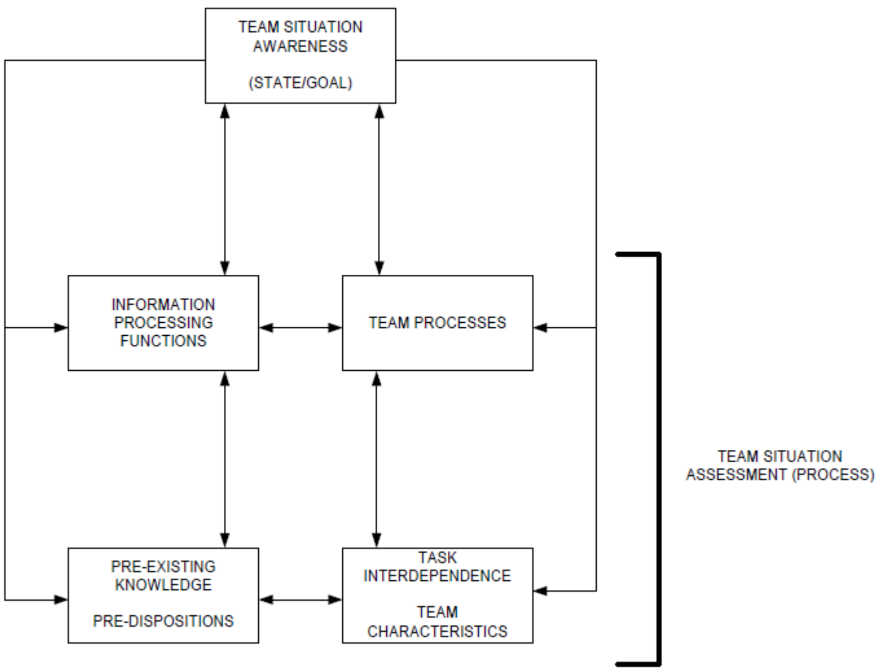


Figure 3.5: Team Situation Awareness model (from (Salas, Prince et al. 1995))

An alternative formulation of TSA, under the designation of Distributed Situation Awareness (DSA), has been developed from Distributed Cognition (DC) theory. The early developments of TSA from a DC lens was conducted by (Artman and Garbis 1998), who suggested that when considering team performance in complex systems, it is necessary to focus on the system as a whole. When defining TSA as DSA, they suggest that TSA is distributed not only throughout the team members but also through the artefacts used by the team to accomplish its goals.

In this perspective, TSA extends the previous account of interactions between team members to the artefacts that support their individual and team processes. Therefore, TSA is fuelled by an emergent cycle of activity borne out of the interactions between system agents whether they are human or not.

According with Artman and Garbis and Stanton et al. (Stanton, Stewart et al. 2006), the team members' SA can be overlapping, compatible or complementary. Thus, the DSA approach purported by Stanton et al., moves the unit of analysis to the system level relying on the notions of compatible SA and SA transactions.

Compatible SA means that although each team member held SA may be different in contents to the other members, it complementarily address the collective needs of the team. This view leads to the notion of SA transactions, which concerns the exchange of information necessary to reach SA compatibility between system's agents.

To date the most integrative model of DSA regarding its tenets is the one offered by Salmon et al., which is presented on Figure 3.6. The model also draw upon Neisser perceptual cycle, for explaining the cyclical and pro-active nature of the SA transactions.

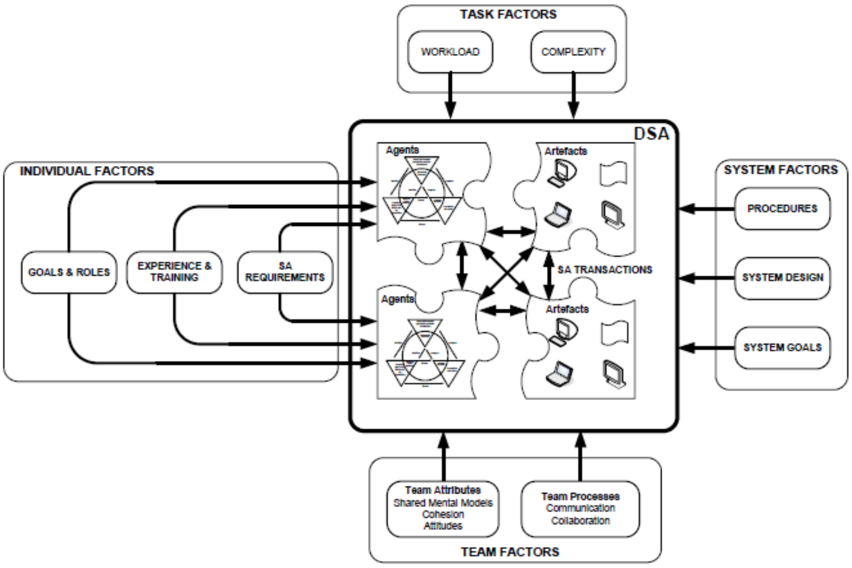


Figure 3.6: Model of Distributed situation Awareness (from (Stanton, Stewart et al. 2006))

The DSA model had been mostly used to develop propositional networks (which are grounded on the notion of concept maps (Crandall, Klein et al. 2006)) for informing the analysis of TSA requirements. Propositional networks are graphs that allow to unveil SA transactions by exposing the contribution of, and usage of, different pieces of information by different agents and the role of artefacts in such transactions.

Nevertheless, they have been devised mostly from post-mortem analysis based upon data derived from observations, task and standard operation procedures analysis. Therefore, its value for early stages of supportive systems design remain residuary. Second, data collection methods lack a more systematic definition and sensitivity to capture fine-grained operational behaviour that underlies TSA development in contingent settings.

Moreover, propositional networks provides mainly a qualitative outcome and despite it constitute a valuable framework for the analysis of TSA requirements, at is present state, it does not support the definition of TSA quantitative measures, for instance to provide a more readily comparison between SA related interventions.

3.1.2 Situation Awareness Measurement Techniques

The disparity between existing SA models has been motivating a great debate over which techniques are more suitable for SA measurement (Cooke and Gorman 2006; Salmon, Staton et al. 2008). As it might be expected, the measurement techniques are strongly linked to the operational definitions of SA. Additionally, the considered application domain will dictate what a SA measurement technique should account for. Goal-Directed Task Analysis (GDTA) e.g. (Endsley, Bolté et al. 2003) and Hierarchical Task Analysis (HTA) e.g. (Kassner, Baumann et al. 2011) had been used for conducting prior analysis of domain related SA requirements and thus unveil what are the relevant elements and behaviours underlying SA development for a particular domain/task context that should be captured by the adopted SA measurement technique.

Existing SA measurement techniques may be categorized as Direct, in the sense that SA is directly enquired from individuals, or Indirect, in the case that SA is inferred from observable activities. Direct SA measurements can be further divided into Objective, as those that elicit factual SA properties, or Subjective, as those that rely on personal judgments. A taxonomy of SA measurement techniques is presented in Table 3.1.

Table 3.1: Taxonomy of Situation Awareness Measurement Techniques

Direct	**Objective**
	Freeze Probes
	Real Time Probes
	Subjective
	Observer Rating
	Self-Rating
Indirect	Performance Analysis
	Communication Analysis

The Freeze Probe technique encompasses freezing the task and presenting to the operators a set of questions regarding the on-going situation. Since it involves freezing the task, the method is mostly suitable for use in simulation and training rather than in real situations. The Situation Awareness Global Awareness Technique (SAGAT) (Endsley 1988) is the most widely used and validated freeze probe technique, and has consistently demonstrated reliability and validity in a number of domains (Jones and Kaber 2004). The technique, which was originally developed by Endsley, concerns the product dimension of SA and therefore the questions included in the freeze probes typically inquires about the three cognitive levels of Endsley's model (perception, comprehension and projection). The main drawback pointed to the freeze probe techniques is the intrusion upon the primary task. Furthermore, since it is focussed on the product dimension of SA, it will assess very little about how SA was achieved.

Another technique involves the use of real-time probes. This technique requires administering questions during the task but without freezing the task. An example is the Situation Present Assessment Method (SPAM) (Durso 1999). While real-time probes do not freeze the task, they still represent an intrusion to the primary task. Furthermore, as in any query-based method, the queries may direct the operators' attention to specific elements/aspects of the situation, which could lead to biased results.

Like the freeze probe method, assessing the SA process is not explicitly addressed. Although recall techniques have been extended to teams by administering SA probes to all team members (Bolstad, Cuevas et al. 2005), they do not account for the team processes leading to SA.

The intrusiveness of probe recall techniques can be overcome by observer-rating SA measurement techniques. Such techniques are particularly suited to assess SA in real-world tasks. Observer-rating techniques typically involve Subject Matter Experts (SME) observing participants executing the task and then providing an assessment of each participant's SA according to the observed behaviour. These ratings are usually based upon pre-defined and observable behaviours, devised for instance from GDTA or HTA analysis.

The Situation Awareness Behaviourally Anchored Rating Scale (SABARS) is an observer rating technique that has been used to assess SA of infantry personnel during field training exercises (Matthews, Pleban et al. 2000). The method involves domain experts observing participants during the task and rating them on 28 observable SA related behaviours.

Clearly, the main advantage of such methods is their non-intrusive nature. However, the extent to which observers can accurately assess SA is questionable, since the relation between SA and observable behaviours cannot be unequivocally defined. Two additional drawbacks can be identified. First, by acknowledging that they are being observed, the operators may diverge from a naturalistic behaviour. Second, the technique may require multiple SME, which may not always be consistent. One should account for an additional limitation when considering distributed teams, since the observers may only get a partial view of the team members' interactions and task context.

An alternative to observer rating techniques is the use of self-rating techniques. In these cases the participants provide a subjective assessment of their SA via a rating scale and are usually administrated on a post-trial basis. The most popular self-rating techniques is Situation Awareness Rating Technique SART (Taylor 1990), originally developed for the assessment of air pilots' SA.

The primary advantages of self-rating techniques are their ease of application and their non-intrusive nature, since they are administered post-trial. However, the collection of post-trial data may suffer from several drawbacks, e.g. poor recall, operator rationalization, and biased judgments by correlating SA with enacted performance.

Endsley et al. (Endsley, Selcon et al. 1998) conducted a comparison between SAGAT and SART, and reported that the operators were poor at reporting detailed information about past events. Furthermore, post-trial questionnaires only capture SA at the end of the task. Additionally, the authors argues that the participants' ability to rate their own SA is questionable, as they may not be able to accurately rate their poor SA.

Once more, whether at individual or team levels, self-rating techniques, by focusing on querying about SA, disregard the understanding of team processes that contribute to TSA.

Indirect techniques appear more suitable to grasp the process dimension of SA. Indirect techniques rely on the inference of SA based on the operators enacted activities. Although the non-intrusiveness of such approach is very appealing, it requires establishing a clear relationship between the operators' activities and the extent that those are actually related with operators SA. This class of SA measurement techniques had been extended to team level, typically by manipulating teams' tasks and assessing the corresponding impact. For instance, the Co-ordinated Assessment of Situation Awareness of Teams (CAST) (Gorman, Cooke et al. 2005) assessed TSA development by introducing glitches in some of the team members communication channels.

Although, communications among team members can be quite revealing of their SA, further considering their performance (e.g. actions enacted over the environment or over the artefacts/systems that support individual and teamwork) is acknowledged to bring more insights to feed SA analysis. Finer grain performance analysis, had also relied on data collected from operational settings. Data collection may be accomplished by several means, ranging from observations to work support systems' logs.

The more holistic approach to TSA provided by DSA is more aligned with the work process tracing based techniques. Nevertheless, as previously noticed, the main outcome of DSA is a propositional network, which hold two main limitations. First, unlike other approaches such as SAGAT and SART, the propositional network does not yield quantitative measures of SA which make it harder to compare and generalize the findings across different tasks settings or application domains. Although practitioners of this approach suggest that metrics from social network analysis (SNA) may be used (e.g. centrality) over the propositional network, that work remain to be done. Second, since propositional networks are mainly developed on retrospective basis, the underlying data collection methods that support the propositional networks development (e.g. interviews, observation transcripts) suffers from the inherent drawbacks already pointed to observer and retrospective based techniques.

Therefore assessments on team communications and work process tracing (which may be more readily accomplished by work support systems' log analysis) seems to provide the more effective means to unveil TSA in its plenitude, although this will require an informed definition of what data should be logged in order to inform representative measures of TSA. Such is inevitably a domain/task bounded endeavour.

3.1.3 Discussion

Much of the debate over the formulation of SA and how it can be measured has been divided between product and process dimensions of the construct. The product dimension concerns the knowledge of the situation that individuals hold. The process dimension refers to how such knowledge is achieved and the extent that operational activity my inform SA inferences.

The definition of team Situation Awareness (TSA) has also experienced different formulations. Those emphasizing the importance of Shared Situation Awareness (SSA) tend to be more focused the product dimension, while others had favoured the process dimension.

SSA had framed team level situation awareness on the overlap of team members' SA. Nevertheless, (Klimoski and Mohammed 1994) had reported that completely overlapping SA in heterogeneous teams may become dysfunctional with regard to team performance. Accordingly, (Cannon-Bowers, Salas et al. 1993) pointed that whereas a certain degree of overlap among TM SA is needed for effective coordination, there is also a point at which too much overlap will lead to group think bias.

The consideration of these arguments drove the need for a more broader formulation of TSA which posit that it is more likely that it will be complementarily distributed through the team.

This acknowledgment had directed the attention for the principled role of team members interactions in complementary inform each other SA. Additionally, the systems oriented perspective on team situation awareness brought by DSA, claims that SA is distributed at a system level, and therefore humans and artefacts perform SA transactions. However given that SA inherently holds a cognitive facet, perhaps a more proper formulation of such conceptualization would be that SA development revolves around exchanges of environmental information among team members and between team members and the artefacts that support their work, and that in order to become SA, information must be available, perceived, attended to and subject to higher level cognitive processing (Hourizi and Johnson 2003).

Thus a comprehensive formulation of TSA measurement, bringing together the shared and distributed views, should assess the transactions between the team members and between the team members and the artefacts supporting their work.

It is, in fact, this set of requirements that makes TSA measurement a complex endeavour. This complexity has been addressed either by direct probes or through indirect inferences made over the analysis of team performance and communications. Probing techniques, if done simultaneously with operational work, present some degree of intrusiveness and are a possible source of results bias. If conducted through post-trials, it may lack sensitivity. The indirect inference techniques draw upon operational data collected through observation and work support systems' logging. Although these approaches may be less intrusive, they also entail some shortcomings. Observations are prone to subjective results, especially if it is not possible to unequivocally establish a relationship between TSA and observable behaviours. This may be also true when the techniques rely on logged data, since the definition of the type and granularity of the data that should be logged is inherently associated with the SA measures that they fuel, which the definition may not be straightforward and is inherently bounded by the team task.

The selection of a TSA measurement technique requires evaluating the trade-offs between their strengths and weaknesses given the research aims. Considering that TSA comprises both the individual and the team levels, and also both the product and process dimensions, a combination of measurement techniques appears to be the more suitable approach.

3.2 Evaluating Collaborative Applications

The evaluation of collaborative applications raises many methodological concerns. The evaluation process must consider the moment (design, prototype, finished product), time span (hours, weeks, months, years), local (laboratory, work context), people involved (domain experts, final users, developers), and type of evaluation (quantitative *vs* qualitative or formative *vs* summative) (Antunes, Herskovic et al. 2008). Also, the purpose of the evaluation may address different interests, ranging from technical (e.g. interoperability, security) to the organizational impact (e.g. effects on performance or work processes dynamics) (Steves, Morse et al. 2001; Gauducheau, Soulier et al. 2005).

The following sections review the fundamental dimensions of evaluation (section 3.2.1), exemplary methods supporting the evaluation process (section 3.2.2), and the factors supporting the adoption of microworlds as an experimental paradigm for the evaluation of collaborative applications (section 3.2.3).

3.2.1 Evaluation Dimensions of Collaborative Applications

Regarding its underlying goals, an evaluation may adopt a formative or summative strategy. Evaluations performed during the development lifecycle are classified as formative, when they are intended to support iterative software development. The literature survey conducted by (Pinelle and Gutwin 2000) reports that 56% of evaluations are formative evaluations based on promised prototypes. Conversely, summative evaluations are performed over finished products in order to assess their fit for the given purpose. This strategy, according to the abovementioned survey constitutes 38% of evaluation initiatives.

A typical evaluation approach is through the comparison of promised solutions. Comparisons can support either formative or summative strategies, contrasting different prototypes or end products regarding the influence of their features against what is defined to constitute the success criteria. These criteria are inherently associated with the goals that motivated the development of the promised solutions, and therefore their formulation involves defining either a set of dimensions upon which to grade the solutions or, especially in experimental research, the definition of a set of operational measures that should be collected (Scholtz and Steves 2004).

In the context of experimental research these measures are classified as dependent variables. Dependent variables, therefore, operationalize the evaluation criteria (Howell 2009). Numerous types of dependent variables have been put forward to evaluate collaborative applications (Fjermestad and Hiltz 1999). Several evaluation frameworks organise the dependent variables in a number of high-level categories and evaluation dimensions aiming to establish a more systematic approach to the evaluation process.

Pinsonneault and Kraemer (Pinsonneault and Kraemer 1989) proposed one of the pioneering evaluation frameworks. It accounts for the relationship between technology support and other factors related with group characteristics, group behaviour and work context, using three main dimensions: Contextual variables, Group Process and Outcomes. Contextual variables reflect the factors that underlie group behaviour. Contextual variables are considered in the scope of five major categories: personal, situational, group structure, task characteristics, and technology characteristics (e.g. anonymity and type of communications). Group process variables account for the characteristics of group interactions, including decisional, communicational, and interpersonal characteristics. Finally, is considered the outcome of the group process affected by technology support, including task-related outcomes and group-related outcomes.

This framework has created a valuable foundation for evaluating collaborative systems. The distinction between group process and outcomes highlights two quite different evaluation dimensions commonly found in the literature, the former usually addressing questions of meaning (rooted for instance, on ethnography (Hughes, King et al. 1994) and groupware walkthrough (Pinelle and Gutwin 2002) based analysis), and the latter addressing questions of cause and effect (e.g. value creation (Briggs, Qureshi et al. 2004)). Other evaluation frameworks such as the ones proposed in (McGrath 1995) and (Fjermestad and Hiltz 1999) are based upon these dimensions.

Additionally group task, group characteristics and organizational impact, have been defined as three fundamental evaluation dimensions by (Tung and Turban 1998). Those accommodate a set of categories to frame representative dependent variables. These categories include dependent variables, regarding: the task, decision quality, group cohesion and coordination, and the organizational culture.

The Antunes and Costa (Antunes and Costa 2003) framework considers the task, group, organization, and technology dimensions. This framework defines several dependent variables: efficiency, effectiveness, satisfaction, perceived value, and economic value. Another evaluation framework developed by (Araujo, Santoro et al. 2004) proposes a set of four dimensions to guide the evaluation, which are: group context, usability, achievable level of collaboration, and impact.

These evaluation frameworks, along with others that may be found in the literature, e.g. (Damianos, Hirschman et al. 1999; Neale, Carroll et al. 2004), organize and bound the evaluation space by purposely guiding the collection of relevant measures. Nevertheless, in practice, the list of dependent variables is frequently considerably too extensive and most evaluation initiatives do not consider all of them (Fjermestad and Hiltz 1999).

McGrath (McGrath 1984) has discussed several high level concerns that should guide any evaluation effort and framed the evaluation process in three fundamental goals: Precision, Generalizability and Realism. The first one refers to the precision of the data collected. This goal is inherently linked with the capability to control the experimental setting and encompasses a clear definition of controlled, dependent and independent variables in order to properly select the experimental protocol and data collection methods. Laboratory based experiments are the paradigmatic example of evaluations pursuing a high degree of control over the experimental setting. Generalizability concerns the extent to which the obtained results remain valid beyond the concrete evaluation instance. High set goals on generalizability usually imply adopting large-scale inquiries and surveys, while low generalizability is obtained by interviewing a small audience. At last, realism addresses how closely the evaluation account for real-world conditions, considering the work setting, the population of users, and the task’s stimulus, and associated time and/or stress and workload demands. Laboratory experiments have been criticized for providing low realism, especially regarding collaborative applications, whereas field studies have been considered to score high on realism but low on precision. Overall, the ideal evaluation should maximize the three goals, for instance using multiple evaluation methods and triangulating the obtained results.

Antunes et al. (Antunes, Herskovic et al. 2008) introduced three more concerns that should be considered when evaluating collaborative applications. The authors also put forward that the evaluation process is also bounded by: the detail, scope and time invested in the evaluation. The detail concerns the granularity of collected data. The spectrum of existing evaluation methods provides different sensitivity regarding data collection ranging from, for instance, collecting mouse movements and keystroke-level data, to less fine-grained data such as completion time or number of operations. The scope of the evaluation had been framed by the breadth of the application regarding the number of functions and components that are included in the evaluation.

Finally, invested time accounts for the time necessary to carry out the evaluation. This in fact is not completely independent from the previous dimensions. However, from a practical standpoint, it is an important dimension to consider as it influences the selection of evaluation methods.

The next section presents an overview of exemplary evaluation methods that had been supporting collaborative applications evaluation.

3.2.2 Overview of Collaborative Applications Evaluation Methods

Existing evaluation methods traditionally fall in two main trends: Usability-oriented and Context-based.

Usability-oriented methods focus on the suitability of software features to the user (and group) requirements to perform collaborative tasks. Usability-oriented methods have their roots in the evaluation of single-user interfaces that have been widely adopted in the Human-Computer Interaction (HCI) field. One may distinguish two types of strategies in usability evaluation methods: laboratory experiments and inspection (also called, discount) techniques.

Usability assessed through laboratory experiments have a strong, formal experimental design undercurrent, in the sense that they require carefully designed experiments, necessary to minimize external influences regarding the original hypothesis on the relation between the pre-established usability measures and the software features. Common usability measures include effectiveness, efficiency, and satisfaction (typically assessed from empirical measurements and questionnaires). However, usability assessment through laboratory experiments requires a fully functioning software application.

In contrast, inspection methods do not specifically require a fully functional application; rather they can be conducted over low fidelity prototypes, and thus performed more frequently over the development process. This is particularly suitable for formative evaluations in the context of iterative development processes, since it allows earlier detection of usability problems.

Usability evaluation based on inspection techniques relies on the judgement of a set of experts about the compliance of the software application with a list of heuristics believed to reflect good usability practices. This approach, in collaborative software applications evaluation, extends previous methods of individual inspection techniques. For instance, heuristic based groupware evaluations, builds upon the list of heuristics (e.g. application must provide feedback of users' actions) created for single-user interface evaluations (Nielsen and Molich 1990). Similarly, the Groupware Walkthrough (GW) method (Pinelle and Gutwin 2002) is based upon the cognitive walkthrough method (Polson, Lewis et al. 1992) originally developed for single-user interface evaluation, positing that experts must walkthrough the steps that users will perform to accomplish their tasks.

Both methods extend their roots to the group level by relying on the 'mechanics of collaboration' conceptual framework defined by (Gutwin and Greenberg 2000; Baker, Greenberg et al. 2001). This framework encompass a set of group work primitives (e.g. communicate a piece of information to another user or reserve a resource) that describe general-purpose communication and coordination activities observed in collaborative tasks.

Inspection-based methods allow more readily evaluations of usability issues. However, two major criticisms have been made: first, they need to combine judgments provided by the experts, which are subjective in their very nature; second, they fall short in considering a realistic usage of the software within the context of a real task and work setting, particularly when a software application is intended to be used in complex work environments (e.g. (Cockton and Woolrych 2002)).

This last criticism could well be further extended to any form of usability-oriented evaluation and inspired context-based evaluation methods.

Context-based evaluation methods are grounded on the assumption that social, cultural, workplace and organizational factors are determinants of software use. Furthermore, they emphasize the need for research results to generalize across operational settings, otherwise the research may be of little practical use. As so, research should be conducted under conditions that are representative of actual work domains (Grudin 1988). Nevertheless, there is also a strong need for defensible research results that can lead to a principled understanding of the factors that affect human performance in complex systems. This requirement implies some degree of control over the data collection apparatus, so that the underlying theoretical principles and assumptions can be rigorously tested.

Representativeness can be maximized by recurring to full-scope simulations or field research. Evaluation methods under this more naturalistic trend include evaluative ethnography (Hughes, King et al. 1994) and scenarios based evaluations (Stiermerling and Cremers 1999). Evaluative ethnography requires one or (usually) more observer(s) to be immersed in the work setting under investigation for a prolonged period. Despite being time and resource intensive, both for collecting and processing data, such approach falls short when considering physically distributed work environments.

Scenario based techniques combine field trials with close interaction with end users. Users are first interviewed to extract contextual information (e.g. work practices, roles) and establishing representative work scenarios. After performing a field trial, the collaborative system is then evaluated by the end-users through discussion workshops where possible design flaws are exposed.

Although context-based approaches are closer to realistic/natural work settings, experience has shown that it is very difficult to obtain defensible and statistically reliable results under such conditions, primarily for the lack of experimental control (Baker and Marshal 1988). Maximizing experimental control is most often achieved by simplifying/constraining the work setting, although this can lead to less relevant results (Howie and Vicente 1998). Additionally, the outcomes of context-based methods most often have a qualitative/descriptive nature that may be difficult to interpret objectively, since they typically do not smoothly translate to more objective quantitative metrics, which could further enhance replicability and comparison.

In an attempt to bridge the gap between overly artificial laboratory approaches and the naturalistic approaches lacking experimental control and objective outcomes, a number of authors in Human Factors (HF) research and complex work environments has been adopting Microworlds as an experimental paradigm. Microworlds provide safe, time and cost-effective environments for system evaluation in the early stages of the technology development cycle, particularly when setting up field studies or full scope simulations presents major challenges (Jhoansson, Trnka et al. 2007; Schraagen and Ven 2008).

The following section reviews and discusses the underlying factors and concerns that have been put forward by the adoption of Microworld environments as an experimental paradigm.

3.2.3 Microworlds as an Experimental Paradigm

The rationale underlying Microworlds as an experimental paradigm purports a conciliation between representativeness of complex work environments and some degree of experimental control that can lead to defensible, well informed results regarding the phenomena of interest (Howie and Vicente 1998). An additional motivation for the adoption of Microworlds is the fact that in complex work environments, particularly in those dealing with high stakes, conducting field research may reveal impractical and thus it is desirable to conduct evaluations of interventions addressing promising solutions within a safe environment.

Microworlds can be defined as dynamic, real-time and task-oriented environments used to study human behaviour in simulated scenarios, retaining some of the real world complexity while omitting other aspects deemed superfluous for the purposes of the research (Brehmer and Dörner 1993).

Despite accommodating experimental manipulation and control of the task environment, without removing the naturalistic characteristics of the phenomena of interest, Microworlds must be carefully developed in order to engage users to the point that their behaviour reflects natural practices (Gray 2002).

In experimental psychology, Microworlds were introduced as an experimental paradigm to study the circular relation between human behaviour and a dynamic environment. This relation comprehends the simultaneous use of a variety of cognitive functions like: establish goals (and further decompose them), make decisions, plan for action, manage attention regarding competing demands, retrieve knowledge from existing MM (e.g. trough association with past experiences), perform situation assessment and update SA, and project future states of environmental variables. Therefore the adoption of Microworlds as an experimental paradigm departs from traditional psychology laboratory studies, by addressing the entangled relations of different psychological functions, and not the elementary cognitive functions in isolation (Brehmer and Dörner 1993; Brehmer 2005).

In fact, the Microworld experimental paradigm has been supporting research on several complex phenomena such as dynamic decision making (Gonzalez, Vanyukov et al. 2005), naturalistic decision-making (Chapman, Nettelbeck et al. 2006), organizational learning (Keys, Fulmer et al. 1996), and SA (Wellens and Ergener 1988). Furthermore, it has been adopted in a number of application domains, for instance: naval warfare (Arthur, Day et al. 2010), industrial processes control (Sauer, Burkolter et al. 2008), air traffic control (O'Brien and O'Hare 2007) and fire-fighting (Omodei and Wearing 1995).

One may find in the research literature several terms similar to Microworlds, such as scaled worlds, synthetic task environments, or high-fidelity simulations (Gray 2002). To some extent they all share the concerns with mimicking the complexity of real world work environment and preserving a subset of the functional relationships found in complex task settings, which are considered fundamental to inform research and support experimental manipulation.

Microworlds, as an experimental paradigm, have supported different research purposes, e.g. whether exploratory (hypothesis generation) or hypothesis testing; with distinct focus, e.g. individual or team level phenomena, or on how specific artefacts, task or environment characteristics impact on particular indices of individual or team performance. For instance, Brehmer (Brehmer 2005) expose a group of subjects to a Microworld environment that introduces communication delays while, a control group perform over the same environment without such delays, and thus eliciting the impact of communication delays over task performance.

Further research works concerning complex work environments had drawn upon Microworld as an experimental paradigm and reported interesting outcomes regarding human performance in complex environments. Dörner, D. (1989) ((in German) cited in (Brehmer and Dörner 1993)) reported that individuals exhibit a poor understating regarding the regularities of the work environment when they receive the information in the form of isolated events over time, moreover, nonlinear processes were usually seen as linear. Others, also reporting from Microworld based studies had revealed individuals difficulty in accounting for both, the delays associated with the effects of their actions, as well as, the side effects of those actions (Brehmer and Allard 1990). Such results appear to be consistent with the inherent limitations of human cognition, that ground the cognitive bias that may impair effective situation awareness, as discussed on chapter 2.2.

In fact, a major concern surrounding the adoption of Microworlds as an experimental paradigm has been the extent that whether the evaluation results are reliable. Rolo et al. (Rolo and Díaz-cabrera 2005) contrasted the results obtained from a field study with those coming from an “equivalent” Microworld experiment, based upon the measures of a set of dependent variables (e.g. number of information requests, number of actions, and performance regarding the success on the purported task). The authors point out that most conceptual aspects of human performance could be assessed in both settings. The comparison had particularly focused on the relationship between individual behaviours and performance. The authors report that more successful individuals actively looked for information about the state of the environment (they act less and request more information), assessed the effects of their actions and anticipated work context changes. While, subjects with lower levels of success accomplishing the task, moved away from the above characteristics by conducting more actions but with less information searches, and they also did not wait/account for the effects of their actions and failed to anticipate contextual changes. The iterative nature of the decision-making process and strong reliance on contextual information exhibited by the most successful individuals appears to be aligned with claims brought by the Naturalistic Decision Making (NDM) and Sensemaking theories (previously introduced in chapter 2.2). Such results provide some evidence that Microworlds support performance evaluation under quasi-naturalistic conditions (Weaver, Bowers et al. 1995).

Nevertheless, although Microworlds have been adopted in a number of application domains and have provided insights regarding a number of phenomena, their validity as an experimental paradigm had been subject to two main criticisms.

One is grounded on the assumption that, if Microworlds really emulate the work environment, it would be expected to find differences in the performance of experts and novices in experiments, which has not been the case in several studies (Chapman, Nettelbeck et al. 2006).

There are a number of possible explanations for such fact. First, the motivation exhibited by experts in simulated conditions may be lower than when facing real high-stakes situations. This has been addressed in Microworld experiments by previously engaging participants in motivation sessions (Cannon-Bowers, Salas et al. 1996; Chapman, Nettelbeck et al. 2006). Second, even if the general characteristics of the work environment may seem familiar to experts, the specificities of the Microworld environment (e.g. time compression) may require an adaptation of experts' mental schemas for action (Brehmer 2005).

These arguments lead, to what has been perhaps the biggest criticism levied against Microworlds, which concerns their fidelity regarding the real work environment. Such discussion has been framed in terms of the ecological and external validity of Microworlds as an experimental paradigm.

Ecological validity concerns the extent a Microworld mimics the real work environment, while external validity addresses the extent that findings coming from Microworld experiments generalize to real-world settings.

The fidelity criticism has been discussed in the related literature with suggestions that fidelity must be viewed in respect to the research goals and acceptable trade-offs between ecological validity in regard to manifestations of the phenomena of interest. Thus, fidelity is not a single feature that is high-or-low for a particular Microworld, since although the Microworld may provide faithful replications/representations of the elements of the operational environment (e.g. resources and artefacts), its development is in large extent bounded for the consideration of the factors that are characteristic of the operational practice (e.g. workload, time-pressure), in order to promote the manifestation of the phenomena of interest.

Experiments with Microworld for training purposes suggest that, when designed for practice of specific behaviours, they require a high degree of fidelity, but those designed for training higher-level cognitive functions (like context-driven decision making) can be simpler and more abstract (Brehmer 2005).

The above arguments lead the question of ecological validity to depart from the assumption that, in order for the Microworld hold validity, it is required a high level of fidelity.

On the other hand, regarding the generalization of the results, the discussion has been framed, in a broader sense, by positing that, it is not a question of transferring concrete results from an experiment to reality. Such transfer may involve the mediation of a theory. So, in that sense, external validity concerns whether Microworlds are useful for developing and testing theories that are informative for supporting generalizations (Kozlowski, DeShon et al. 2004).

Besides fidelity and generalization, (Rolo and DÍaz-cabrera 2005) and (Sauer, Wastell et al. 2000) propose that Microworld validation also requires comparing findings with other methods such as field studies and full-scope simulations, although this may be difficult in some domains because of some already discussed constrains, which have principled grounded the impetus for the adoption of a Microworld based approach.

The conduced literature review had further pointed that in order for Microworld evolve as an experimental paradigm, it is required the definition of what should be its main building blocks, since existing Microworlds based research had experienced a proliferation of domain (and phenomena) dependent initiatives, leading to a lack of a cumulative body of knowledge. This is particular pertinent in the field of (collaborative) applications evaluation, where its adoption is still getting-off.

Sauer et al. (Sauer, Wastell et al. 2000) suggest that the development of a Microworld environment must account for three foundational concerns. First, establishing a rich understanding of the domain and its archetypal tasks (either through participatory design or by engaging domain experts). Second, devise a set of independent, dependent and control variables focussed on the phenomena of interest. The clear definition of such variables informs the Microworld design and supports traceability from requirements to data collection. And third, supply user interfaces that minimize usability problems and enhance user participation.

Nevertheless, these guidelines are very generic and do not significantly contribute to develop and apply the Microworld approach in a systematic way. As previously mentioned and consistently reported in the research literature, judging the validity of a Microworld approach still has to be done on a case-by-case basis, analysing the relationships between a particular Microworld design and the specific phenomena of interest.

The present work addresses this issue in the following chapter, where the fundamental building blocks guiding the development of a Microworld are put forward.

3.2.4 Discussion

In this section existing methodological concerns regarding the evaluation of collaborative applications have been presented and discussed. Several evaluation frameworks have been put forward highlighting a number of dimensions that should be considered when defining evaluation criteria. The most common set of dimensions that have been related to the evaluation of technology support of group work, involve consideration for the group, task, and associated processes characteristics, task context and impacts yielded by the introduction of technological artefacts.

These dimensions guide the selection of measures and measurement methods, supporting comparisons between promised solutions, either using a formative or a summative strategy.

The strengths and weaknesses of several evaluation methods have been discussed, considering in particular their feasibility and the constrains posed by the application domain.

Usability-oriented methods have been contrasted with context-based methods. Usability-oriented evaluation, whether done through laboratory experiments or inspection techniques, may provide earlier insights about the use of a collaborative application; but they disregards the role of the work context in which use takes place. Usability-oriented evaluation out of the laboratory particularly, regarding Mobile Applications and especially on early design stages is still not a common procedure (Kjeldskov and Graham 2003).

Although some works (e.g. (Carter, Mankoff et al. 2007; Froehlich, Chen et al. 2007)) rely on logging users mobile application use, such techniques are more adequate to assess individual users rather than to develop a comprehensive account of collaboration processes within teams.

The impact of the work context is particularly pertinent in complex work settings, such as CIRM, considering it is bounded by factors like stress, information uncertainty and time pressure that most often lead to depart from established work procedures. Thus the evaluation of (collaborative) applications supporting this type of work should account for emergent and unanticipated usage flows.

Conversely, context-based evaluation methods acknowledge the role of the work setting as determinant of collaboration. Nevertheless, these methods are time consuming, resources demanding, and the extent that they hold external validity has been questioned. Moreover, in time and safety critical environments and considering spatially distributed teams, they may not be feasible, which is a strong argument for considering other approaches, particularly when the research addresses the ubiquitous nature of mobile collaborative applications.

The section concludes by presenting Microworlds as a research instrument capable to cope with the above limitations. Microworlds entail a trade-off between the realism of field studies and the overly artificial laboratory experiments. The fundamental concerns surrounding the adoption of Microworlds as an experimental paradigm are directed to its ecological and external validity. It has been pointed out that accounting for both draws heavily on carefully crafting the Microworld in order to support phenomena of interest manifestations on a quasi-naturalistic basis.

4 Microworld Reference Model Specification

As discussed in chapter 3.2.3, Microworld environments have been acknowledged as a valuable experimental paradigm, given the constrains that surround the research regarding some particular target application domains and phenomena of interest, as those considered by this dissertation.

Nevertheless, the conducted literature review suggests a lack of a reference model to inform Microworlds development, so they provide a systematic and consistent instrument for experimental research, particularly regarding the evaluation of collaborative applications.

Microworlds, as synthetic task environments, are intended to support experimental sessions that address the operational level of human activity in quasi-naturalistic contexts. Based on a characterization of the operational level of teamwork, particularly considering the demands posited by Critical Incidents Response Management, this chapter presents a set of fundamental building blocks that should be accounted by Microworld instances. These building blocks are comprehensively specified in a domain independent manner, in order to provide a valuable resource for researchers when adopting Microworlds based research.

4.1 Foundations for Developing a Microworld Environment

4.1.1 Portraying the Operational Level of Critical Incident Response Management

As introduced in chapter 1, the uncertainty surrounding the dynamics of CIs postulates that CIRM is an ill-defined emergent process requiring teams to adapt to exceptional work demands (Orasanu and Connolly 1993).

In CI occurrences, teams often face rapidly evolving situations that require collective and continuously enacted activities, concerning both the diagnosis of the operational context and the planning/re-planning of their response strategy. Both diagnosis and planning entail a collaborative dimension, leveraging individual contributions and balancing collective effort (Sinnreich 2008). Therefore, the operational level of teamwork brings forward that team members continuously shift between individual and collective work (Dourish and Bellotti 1992).

Although each team member may have her/his own area of expertise and pre-established role and responsibility, happenstance, emergent needs, contingencies and the nature of specific tasks lead teams to balance between different work coupling modes. The interdependence of teamwork has been classified as pooled, sequential, and reciprocal (Thompson 1967). Some tasks may be more prone to loose coupling, which indicates that each team member provides a discrete contribution to the whole by collating (or pooling) information concerning her/his own task work. Other tasks may be sequential, positing that team members may have to wait for outputs from the others to accomplish their work. Tighter coupling tasks holds a reciprocal interdependence and occur when the team members must continuously resolve the mutual constraints between their work. These tightly coupled tasks are those that most likely benefit from enhanced SA.

Traditionally, CIRM teams tend to coordinate their activities according with a hierarchic command and control (C2) structure (Berrouard, Cziner et al. 2006). C2 structures comprehend specific roles impersonated by team members and established communication strategies to bridge the gap between the higher level perception of the situation and the actual knowledge hold in field by lower-level team echelons (Wybo and Latiers 2006). The main reason leading C2 team structures is that, by decomposing the information and operational requirements in different levels, it becomes easier to assign, roles, tasks and establish concrete communication strategies for supporting coordinated activity, which helps managing the overall complexity of CIRM.

For instance, a common fire-fighter teams organization involves, an Incident Commander (IC) and several small groups of fire-fighters, each one lead by a captain. The role of the IC is to organize and understand the overall situation and plan how to deal with it, making decisions about which global actions should be enacted to cope with the current emergency and how to manage the available resources in the specific situation context (Jiang, Hong et al. 2004).

It has been noticed, however, that in CIRM contexts teams may experience some difficulties following hierarchical decomposition, especially regarding the strict levels of responsibility and autonomy imposed by such structure. In its pure form, a C2 structure hardly accounts for the emergent properties of CI, since it favours a concentration of decision-making and puts too much emphasis on hierarchical communications and pre-established courses of action. It is not uncommon that the urgency and exceptional demands posited by CI lead team members to overcome predefined operational protocols and enact real-time contingency solutions framed on improvised activities (Mendonça and Wallace 2007).

Aligned with resilience principles drawn from High Reliability Organizations (HRO) (discussed in chapter 2.2), disaster sociologists emphasize that more effective CIRM teams do exhibit openness to self-initiative, which has been considered paramount to overcome gaps in established procedures and to address urgent and situated action requirements (Drabek and McEntire 2003).

Some research results support that flexibility in team structure regarding the autonomy of its echelons facilitates team coordination and decision making in stressful and time-critical situations (Urban, Bowers et al. 1995). Under such settings the team members become more receptive to contributions put forward by other team members regardless of their roles and hierarchical levels (Driskell and Salas 1991).

This distributed responsibility has been found to be more prone to assemble, filter, collate, disseminate, and share information (Sinnreich 2008) and to accommodate qualitatively different communication and coordination strategies (Johnston, Cannon-Bowers et al. 1997; Serfaty, Entin et al. 1998). (Thompson 1967) puts forward that coordination may be accomplished through three main modes: Standardization, Plans and Mutual Adjustment. While standardization posits the existence of established coordination rules or routines, plans purport the identification of key information elements that should be exchanged between team members without specifically detail the fine-grained coordination activities, which may be hardly devised in some task environments. In more extreme cases where the specific information requirements may not be previously known or are highly dynamic, team coordination is supported through continuous mutual adjustment of team members on-going activities.

While the first two coordination modes entail a more implicit coordination mode, as purported by the C2 orientation and thus puts a smaller overhead on team communications, they becomes feasible if team's structure and processes are well established and the task is fairly familiar (Serfaty, Entin et al. 1998; Fiore, Salas et al. 2001). Conversely, mutual adjustment coordination is more likely to emerge under more exceptional tasks, and requires more intensive team communication to articulate work (Entin and Serfaty 1999; Stout, Cannon-Bowers et al. 1999; Klein, Adelman et al. 2008).

As discussed in (Dessalles, Ferber et al. 2008), the coordination of teamwork in emergent settings will be hardly approached purely bottom-up, based on feedback provided by team echelons, or purely top-down based on pre-devised operational procedures. CIRM teams are more likely to exhibit a combination of C2 behaviours and those that are purported by more distributed arrangements of team members' autonomy. CIRM teams do hold a set of standard operational procedures and definitions of specific roles for operational activity, which constitute a valuable team resource to deploy the CIRM, nevertheless, the emergent demands posited by unexpected CI requires that teams dynamically incorporate specific information of the environment and actual team activity is framed in the context of the current situation, and thus adapt their pre-established arrangements and complementarily embrace alternative coordination strategies (Eby and Dobbins 1997; Kasl, Marsick et al. 1997).

Despite the afore discussed role of team structure and the inherent authority, leadership and coordination modes, further team characteristics, such as expertise, maturity, cohesion and the existence of shared mental models, developed through training programs and experience (generally introduced in chapter 2.3 and further considered in chapter 3.1.1.2), research literature had consistently reported that effective teams, particularly when performing beyond anticipated scenarios, exhibit two influential behaviours. Those are: mutual performance monitoring and backup behaviour (Endsley and Kiris 1995; Salas, Sims et al. 2005).

Mutual performance monitoring implies being aware of and keeping track of other team members' work, at least regarding the subset of work that is closely coupled with own work. As more complex a task is, more important monitoring will be for a coordinated response.

On the other hand, backup behaviour is the ability to balance team workload and compensate awareness and knowledge gaps of team members (Carroll, Neale et al. 2003). In the case of spatially distributed teams, both monitoring and backup behaviours are heavily dependent on team communication. Under CIRM, it has been observed that breakdown of team communication yields lower levels of performance whenever coordinated activity is required

(Xiao, Hunter et al. 1996; Neuwirth, Morris et al. 1998; Coiera, Jayasuriya et al. 2002; Kennedy and McComb 2010).

The above characterization of the operational level of teamwork in CIRM establishes the basis for a set of core components for developing Microworlds addressing the operational level of teamwork. Those are detailed in the following section.

4.1.2 Devising Microworld Core Components

In the present work, drawn from the systemic perspectives on CI discussed in chapter 2.1, specifically building upon STAMP, which more readily tackles the operational level of CIRM, one conceptualize the previously described factors that underlie the CIRM process according the traditional control loop defined by control systems theory (Ogata 2010) as depicted in Figure 4.1.

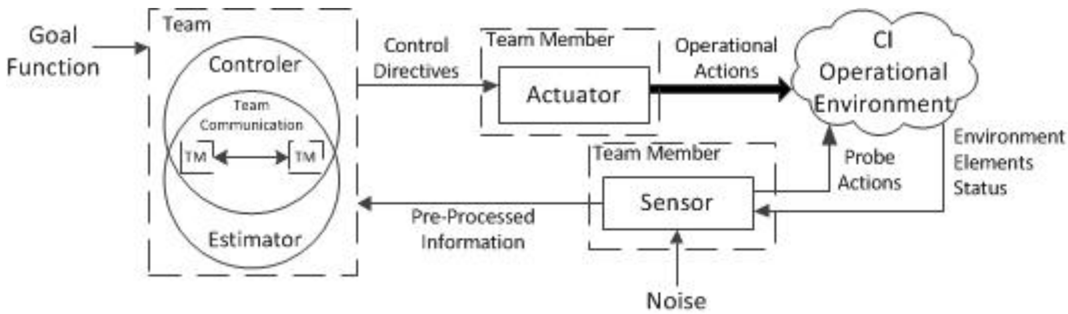


Figure 4.1: Critical Incidents Response Management control loop.

CIRM teams' operational activity is intended to bound (control) the dynamic process of CI escalation, with the goal function of refraining CI major disruptive consequences. Considering how CIRM teams operate, one acknowledges that team members complementary perform the role of sensors and actuators in the work environment.

While impersonating the sensor role, team members probe the environment about the impacts of CI, following the tenets of Sensemaking and Naturalistic Decision Making research that had been presented in chapter 2.2. The team members' perception of the work environment is inevitably framed by a number of external factors (e.g. information availability) and internal factors (e.g. attention), as pointed by SA models discussed in chapter 3.1.1. Moreover, cognitive biases may introduce some distortion on how such information is sensed, as discussed in chapter 2.2.

Although the sensed information is inherently a source of SA, such information is subject to further processing at team level through communication, in order to enhance collective TSA, considering the overlapping, distributed and complementary dimensions of the construct discussed in chapter 3.1.1. Given the uncertainty and availability (or not) of some pieces of information, team communications also provides the basis for estimating and projecting the CI evolution. Thus, pre-existing referents and established operational procedures are revised and may be replaced by emergent ones in order to address the contingent requirements of the CI.

The actions devised from the developed coping strategy are operationally enacted by the team members while assuming their actuator role regarding the operational environment. As posited by the theories of Situated Cognition and Ecological Psychology, which highlight the opportunistic nature of human behaviour, pointed out in chapter 2.3, and the more dynamic models of SA formulated in chapter 3.1.1, the team members performance entails a pro-active facet. Thus, rather than merely probe and deploy, collective action is also self-motivated. As pointed out in the previous section, team communication has been extensively considered as the fundamental mean to support monitoring and back up behaviours for a coordinated response. The role of team communication achieves a particular importance, in operational settings comprehending spatially distributed team members.

This conceptualization leads to the consideration that in order to account for the dynamic control loop, a Microworld must accommodate the interactions that team members enact in the work environment, in the realm of their sensor and actuators roles and also, the means that mediate team communications.

The fundamental building blocks for developing a Microworld are presented in Figure 4.2.

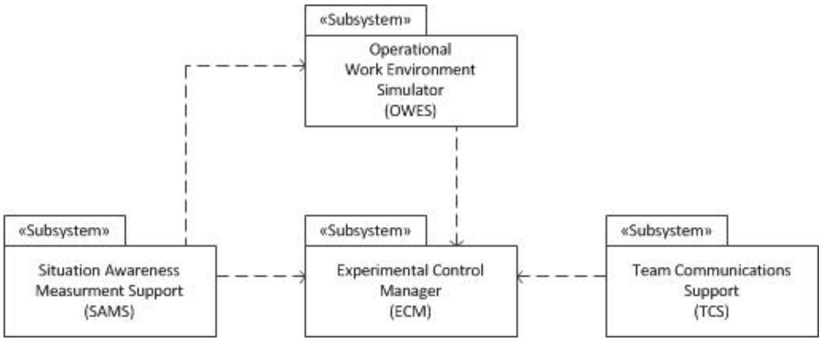


Figure 4.2: Microworld reference architecture.

The proposed reference architecture purports four core subsystems: Operational Work Environment Simulator (OWES), Team Communication Support (TCS), Situation Awareness Measurement Support (SAMS), and the Experimental Control Manager (ECM).

OWES provides a virtual representation of the work environment under which CIRM teams perform. Its main purpose is to provide the means that afford teams to mimic their real actions. Since team communication constitute a fundamental dimension of CIRM, the TCS subsystem concerns the instantiation of communication channels used by CIRM teams.

The definition of SAMS subsystem becomes natural, given that TSA constitutes a major dimension of analysis of team performance. SAMS addresses the TSA measurement techniques reviewed in chapter 3.1.2.

At last, adopting Microworlds as an experimental paradigm, the ECM subsystem tackles the inherent aspects of experimental research. It incorporates experimental protocol requirements managing controlled, dependent and independent variables.

The relations among the subsystems depicted in Figure 4.2 establish their logical dependencies. SA measurement in inevitably linked to operational activities and experimental protocol, which posit what to measure and when. The operational work environment simulator is linked to the team tasks purported by the experimental protocol, which are intended to promote the manifestation of the phenomena of interest. Finally, the affordances and constrains of OWES and TCS are also linked to the experimental protocol, since they may ground the control and independent variables of the study of teamwork. These subsystems are further specified in the remaining sections of this chapter.

4.2 OWES - Operational Work Environment Simulator

The Operational Work Environment Simulator (OWES) holds a virtual representation of the work environment. Its main purpose is to afford that teams accomplish their operational activities toward the elements that constitute the working environment in the real setting.

The definition of the elements that comprise the operational work environment depends on the target application domain (Endsley, Bolté et al. 2003). As discussed in chapter 3.1, the original domains that fuelled early SA research were military aviation and air traffic control, where SA evaluation depends on the knowledge hold by the operator regarding elements in the work environment, which includes for instance, the location of nearby aircrafts. SA research had been extended to other application domains and, for instance, in fire-fighters domain, SA had been evaluated grounded on the emergency response teams' management of vehicles dispatching given their inherent aptitudes and the emergency situation instance requirements. In this case, the vehicles constitute relevant Operational Work Environment Elements (OWEE).

Figure 4.3 presents a model to inform the development of the OWES. The model uses the Unified Modelling Language (UML) notation since due to is graphical orientation it provides an expressive representation of OWES considered constituents and their relations.

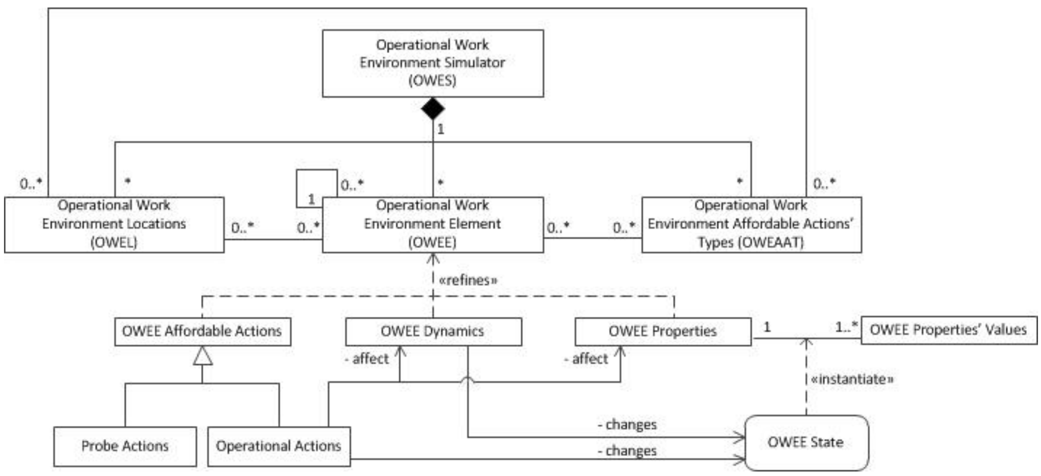


Figure 4.3: Operational Work Environment Simulator (OWES) model.

As shown in Figure 4.3, OWES is composed by three main abstract classes: Operational Work Environment Locations (OWEL), Operational Work Environment Affordable Actions types (OWEAAT), and Operational Work Environment Elements (OWEE). The notion of abstract class in UML posits that such classes are not intended to be directly instantiated as they offer a high-level conceptualization that guides the development of concrete classes that address specific domains requirements.

The OWEL abstract class, was included in the model given the focus on spatially distributed work. Considering that team members may operate in different locations (e.g. in different rooms or buildings), OWEL accounts for the virtual representation of those locations in OWES.

While operating in the field, team members are able to perform a myriad of physical activities (e.g. moving between locations and shifting physical equipment from one place to another). Microworld support to the virtualization of such activities is accomplished by the OWAAT abstract class. The third proposed abstract class, OWEE, addresses the representation of the elements that comprises the work environment (e.g. physical equipment). Recurring to the UML abstraction relationships, which defines dependencies between model elements at different levels of abstractions, the proposed model uses the *<<refine>>* dependency to depict that the OWEE definition is accomplished by three constituents: its representative properties, affordable actions and intrinsic dynamic.

OWEE properties are typically used to represent real world attributes. The scope and level of detail of OWEE properties in bounded by the OWES requirements regarding, for instance, its face validity, given the phenomena of interest that is being scrutinized, as discussed in chapter 3.2.3. Thus if high face validity is required the OWEE properties must account for fine-grained attributes such as size, shape or colours. Conversely, if face validity may be relaxed those properties may be neglected.

The consideration of OWEE affordable actions is of paramount importance in a synthetic task environment, since they it will frame the actions that team members can execute. In the present model, one further distinguishes those, regarding their nature. OWEE supports both probe operations (e.g. inspecting/collecting OWEE properties) and operational actions (e.g. push a reset button, or switch on or off a piece of equipment).

However, the characterization of some OWEE may require the consideration of its intrinsic dynamics in order to reproduce real world behaviour (e.g. how some properties change through time and in respect to some actions and interactions).

Since in the scope of teamwork, several end users (team members) must simultaneously operate the OWES, the client-server paradigm offers a suitable approach for the OWES architecture. As depicted in Figure 4.4, the OWES architecture draws upon the typical three-tier structure of software applications: 1) presentation layer, which is deployed on end-user client computers, provides the Graphical User Interface (GUI); 2) the application’s logic holds the OWES engine, which is responsible for keeping the overall OWES state consistent and synchronized among the several clients, and 3) the data layer, holds the information necessary to support the OWEE (i.e. their properties and respective range of values) and persisting the operational values in the context of the OWES use in experimental trials.

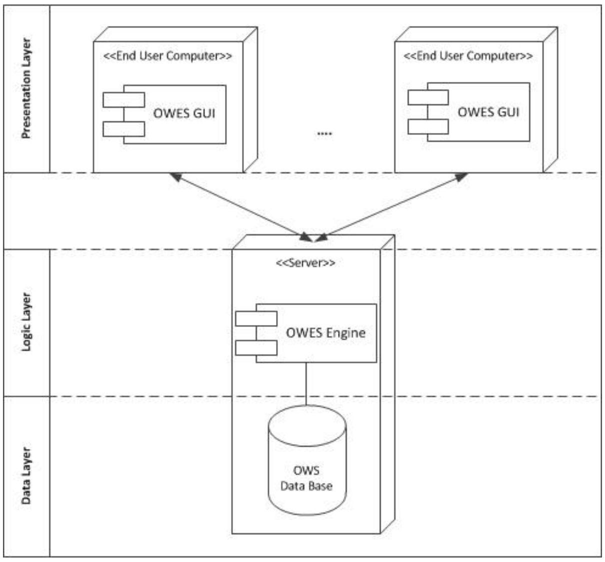


Figure 4.4: Operational Work Environment Simulator (OWES) architecture.

The specification of the OWES Engine rests on a state space (SS) approach. The state space paradigm has been used in software engineering for specifying high-level conceptual requirements and rigorously inform design and implementation (Lee 2004; Wagner, Wagner et al. 2004; Wagner 2006).

(Buchner and Funke 1993) suggest four main strengths of adopting the state space paradigm in the context of computer-simulation research of human performance in dynamic task environments. First, it affords a rigorous, systematic formulation of environments that can be reused. This argument addresses the criticism that Microworld environments are typically developed for particular application domains.

Second, by representing the environment through its internal states, it smoothly accommodates the manipulation of such states to promote specific manifestations of phenomena of interest, which have been pointed as a fundamental tenet in experimental research.

Third, since the elements of OWES may themselves be described in terms of their intrinsic SS, the SS representation of the overall operational environment accounting for the net effects of each element state changes, contributes to the ecological validity of the simulated environment.

And fourth, considering that in the context of complex dynamic tasks users continuously interact with OWEE, framing such interactions in the overall environment's current state provides a white-box approach that may complementarily inform the measures obtained with the black-box paradigm (like using questionnaires and performance indices), by supporting performance tracing framed on the overall context/state under which the operational actions were enacted.

Given the above characterization of OWEE, the definition of an OWEE State is based on a set of its properties and corresponding values, defined according to the target application domain. Table 4.1 illustrates an exemplary definition of OWEE states.

Table 4.1: Operational Work Environment Elements (OWEE) state definition.

	States			
OWEE Properties	State A	State B	…	State N
Property 1	Value = X	-		Value = X
Property 2	Value = Y	Value = W		-
…				
Property M	-	Value = Z		-

State transitions may be accomplished by three means. One derives from the OWEE intrinsic dynamic laws implemented through the methods of OWEE class instances (which accounts for instance, time-related triggers). Other is derived by the operational actions performed by team members over the OWEE (which may change one or more values of its properties). The relation between actions and the OWEE state space must be accounted by the OWES engine and is herein illustrated through the matrix presented in Table 4.2.

Table 4.2: Operational Work Environment Elements (OWEE) state transition matrix regarding affordable operational actions.

OWEE Current State	A	B	…	N
Operational Actions	**OWEE Next State**			
Operational Action 1	B	C		A
Operational Action 2	-	A		-
…				
Operational Action K	C	-		B

The third considered mean by which OWEE may switch between states addresses the ecological validity of the OWES, by acknowledging that events in an OWEE may also constitute a source for state changes on others due a net effect. This requires the OWES engine to keep a representation of the linkages between OWEE. As an example, an instance of the OWEAAT (Operational Work Environment Affordable Actions Type) class could be, *MoveResource,* which prescribes that some linkages of a Resource (an instance of the OWEE) to other OWEE may have to be updated, considering for instance a new location. Moreover, the available OWEE affordable operational actions may also impact their web of associations; an example would be switching off a piece of equipment, which may dictate an update on the states of related OWEE.

The specification of how this ecological system is supported by the OWES engine requires consideration for the systemic structure and dynamics of OWEE relations. The representation of the systemic structure of OWEE is supported by the matrix presented in Table 4.3. According to the exemplary matrix of Table 4.3, OWEE 1 holds an association to OWEE 2 and OWEE 3, and OWEE 2 is further associated with OWEE N.

Table 4.3: Operational Work Environment Elements (OWEE) systemic structure matrix.

OWEE	OWEE 1	OWEE 2	OWEE 3	…	OWEE N
OWEE 1		ADM_{12}	ADM_{13}		-
OWEE 2	ADM_{21}		-		ADM_{2N}
OWEE 3	ADM_{31}	-			-
…					
OWEE N		ADM_{N2}			

The dynamics of the OWEE associations is accounted for, in a set of Association Dynamics Matrixes (ADM_{ij}), which reflect the net effects of a state change in $OWEE_i$ on $OWEE_j$, as depicted by Table 4.4.

Table 4.4: Operational Work Environment Elements (OWEE) Association Dynamics Matrix (ADM).

$OWEE_j$ Current State / **$OWEE_i$ Current State**	A	B	…	M
	$OWEE_j$ Next State			
A	B	-		A
B	-	-		B
C	-	A		A
…				
N	B	-		-

The presented OWES specification provides a domain independent referent for informing implementations. An immersion on the specific application domain will lead its instantiation towards a functional environment. A concrete example is presented in chapter 5.3.1 regarding Help Desk Teams application domain earlier introduced in chapter 1.3, as a representative target application domain for this work.

4.3 TCS - Team Communication Support

4.3.1 Team Communications in Distributed Settings

Spatially distributed teams do not have the communicational advantages of those operating in a collocated environment, where the activities of team members may be directly perceived. When collocated, teams can get a great deal of information through non-verbal communication relying, for instance, on facial expressions and gestures, or by directly observing the on-going activities. For instance, (Xiao, Mackenzie et al. 1998) emphasize the paramount role of non-verbal communication in collocated medical teams. The participative report of on-going activities constitutes a fundamental asset to support team performance in distributed settings (Kreijns, Kirschner et al. 2003; Janssen, Erkens et al. 2007).

Accordingly, in (Taylor, endsley et al. 1996) the authors studied teams in their adaptation to dynamic tasks and found that team members of effective teams actively check information against each other, perform specific communication towards the coordination of collective information needs, establishing priorities for action as a group, and exhibit a group norm of questioning each other. The authors report that poorly performing teams exhibit a group norm of not volunteering pertinent information, rely on individual expectations, do not check others' expectations, fail to prioritize their tasks as a group, and perform more uncoordinated.

In line with these results, (Klein, Zsambock et al. 1993; Citera, McNeese et al. 1995) also found out that efficient teams communicate more in order to improve situation assessment, keeping track of other team members and collectively working to build a shared understanding of roles and functions. These research studies stress the relevance of team communication when dealing with exceptional work settings, as put forward in section 4.1.1.

Effective communication is not just a matter of transferring information between sender and the receiver, since the receiver may not readily understand the meaning or relevance of the conveyed information. However, what distinguishes teams from other groups of people working together, is the knowledge that each member possesses regarding differentiated responsibilities, expertise and roles hold by the others, as well as the team's expected performance (Salas, Dickinson et al. 1992). Without such common ground, communication may become inconsequential, which has been related to lower performance (Bolstad and Endsley 1999; Stout, Cannon-Bowers et al. 1999). Acknowledging that just as experience and training for accomplishing individual tasks increases resilience in high-stress environments, training specific team communication skills, furnishes teams with deliberate strategies to cope with information requirements under unexpected situations. This issue has been addressed by several training programs such as cross-training (Blickensderfer, Cannon-Bowers et al. 1998) and interposition training (Johnston, Cannon-Bowers et al. 1997; Duncan, Rouse et al. 1999).

In the case of spatially distributed teams, team communications has traditionally placed the burden of information transmission through a variety of systems, ranging from audio communication channels (e.g. phone and radio) to groupware tools (e.g. chat and shared displays applications), to boost the sharing of pertinent information and to compensate for the constrains brought by the fact that team members are not co-located.

In a number of application domains such as air traffic control (Berndtsson and Normark 1999), control rooms (Heath and Luff 1992), emergency dispatch (Pettersson, Randall et al. 2002), and network troubleshooting (Whittaker and Amento 2003), where some work is characterized as time-critical, it has been emphasized the principled role of speech communication. Field studies involving mobile professionals like service technicians (Fallman 2003), police patrols (Nulden 2003) and fire-fighters (Landgren 2005) suggest that collaborative work in these settings is primarily supported by speech communications (supported on phone calls or radio communications). Furthermore, it has been stressed out that in uncertain work situations the primary means of coordination is speech and that in emergency situations an increase of spoken communications is experienced (Dunn, Lewandowsky et al. 2002).

In the scope of teamwork under CIRM, persisting communication and making it available at team level have been pointed out as key requirements to support accountability of performed actions, collaboration and TSA (Landgren 2006). Despite supporting information sharing and persistence, groupware systems also support other cognitive enhancements like perception and attention, using notifications of pieces of information and/or displaying information according with salience, thus relieving many working memory limitations (Norman 1991).

Although situated information report and retrieval, and real-time information sharing, affordances brought by groupware systems, particularly mobile collaborative applications (MCA), are apparently very appealing for information assemble and dissemination to support continuous team performance monitoring and adaptation, they also hold some inherent limitations. Interactions mediated by MCA generate only a fraction of the information conveyed through speech-based channels. Thus, a more realistic perspective about the role of MCA in emergency situations is that they offer a complementary, redundant communication channel. Multiple communication modalities can increase information throughput, since they draw upon different cognitive resources (Wickens 1984). While time-critical information requirements may lead teams to rely more on speech communication, MCA may nevertheless lower cognitive load regarding information retrieval, assemblage from multiple sources, and storage.

The Team Communication Support (TCS) subsystem is defined by the two subsystems depicted in Figure 4.5: one is intended for Speech Communications Support (SCS), while the other considers the Mobile Collaborative Application Emulator (MCAE) necessary to implement the functional features of MCA.

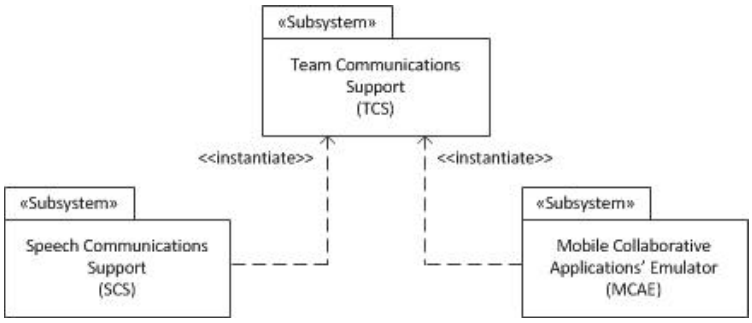


Figure 4.5: Team Communication Support (TCS) subsystems

The SCS may be developed in multiple ways. The SCS implementation, presented in detail in chapter 5.3.2, was developed upon the Team Speak project (TeamSpeak), a Voice over Internet Protocol (VoIP) engine, which supports one-to-one (emulating phone calls), one-to-many and many-to-many (emulating conference calls and radio communications) communication modes.

The next sub-section builds upon the existing body of knowledge to propose a set of MCA functional requirements that should be available on the MCAE to support team communication.

4.3.2 Mobile Collaborative Applications Functional Requirements

As it has been noticed previously, efficient performance of distributed teams draws heavily on individual pro-active volunteering of information to the benefit of the team. One may frame these communication initiatives on a team member acknowledging the pertinence of informing the others about her/his perception of an Operational Work Environment Element (OWEE) state and enacted operational actions toward it, or on the requirements derived from managing work coupling. Accordingly, (Salas, Prince et al. 1995) had model teamwork in two dimensions: task work and team management. While task work refers to the team's operational activities, team management addresses the overhead necessary to coordinate the team to perform as an integrated unit.

These two dimensions of teamwork root the logical architecture of MCA, sown in Figure 4.6, and put forward three main constituents: 1. Operational Work Environment Elements Report Support (OWEERS); 2. Team Management Support (TMS); and 3. Information Dissemination Support (IDS). Those are intended to frame the specification of high-level functional requirements of MCA.

The specification of functional requirements in OWEERS addresses the need to report some relevant individual information derived from team members operational activity to team level. Even though, such reporting posits some overhead work, inevitably team members enacted operations over OWEE will only have the potential to feed TSA if explicitly reported to a common medium such as the one provided by the MCA. Figure 4.7 presents an UML use case diagram depicting the underlying functional requirements for supporting OWEERS. As portrayed, in the use case, MCA must provide the means to report both over OWEE properties (or states) and enacted operational actions over it.

Since teamwork also inherently requires team management, MCA functional requirements toward coordination support are put forward in the TMS use case, as shown in Figure 4.8. The TMS use case expresses the consideration of team coordination in its more open form, as discussed in section 4.1.1, which acknowledges that team management is driven by contingent/emergent needs. Therefore, MCA must afford real-time task assignments derived from the consideration of the work context (e.g. individuals and OWEE location and states).

Despite support for real-time assignments and acknowledging the opportunistic nature of CIRM, the present TMS specification also purports that team members may explicitly assign individual responsibilities in OWEE interventions, and that such assignments should be made noticed to team level.

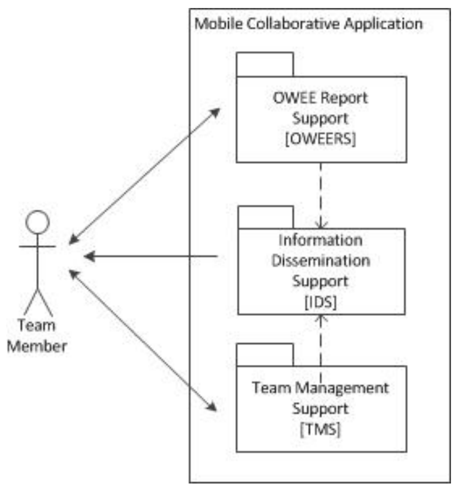


Figure 4.6: Mobile Collaborative Application (MCA) main components

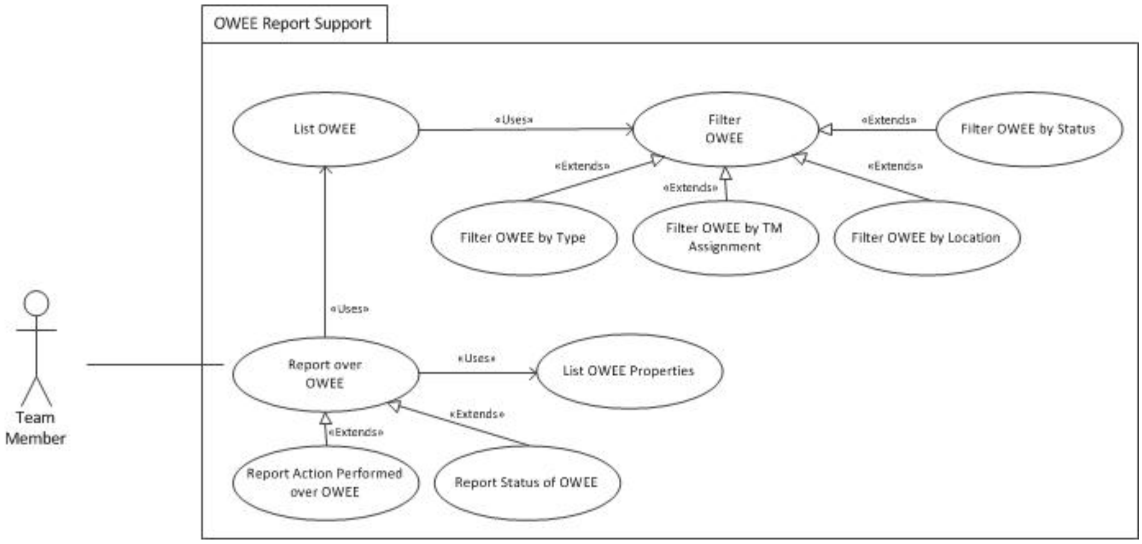


Figure 4.7: Use Case for the Operational Work Environment Elements Report Support (OWEERS)

Since the work environment may comprehend a multitude of OWEE (each further holding several properties) and that teams are composed by several members (each characterized by a number of attributes such as expertise, assignments and current location), both the OWEERS and TMS use cases express the need for information filtering mechanisms.

As already noticed, information exchange mediated through MCA posit some overhead work. Thus, information retrieval should be facilitated, to alleviate MCA usage. As attention and working memory are limited, the way MCA filters and displays information impacts individuals ability to process and integrate it. Excessive display density, particularly in mobile devices (with limited screen dimension), can slow down information search and retrieval.

The effect of overloaded displays has been reported to negatively impact SA (Bolstad and Endsley 1999). Being capable to narrow the displayed information range, by framing it on a team member criteria according current needs is an important usability requirement. The filtering criteria specified in the use cases diagrams do not pretend to constitute a closed set but rather provide representative examples and guidance for MCA developments, which is ultimately framed by the target application domain.

Reporting information regarding operational work and the emergent work arrangements derived from real-time assignments will only constitute an effective asset for acquiring and maintaining TSA if such information become readily available at team level. This consideration has driven the reasoning for introducing the Information Dissemination Support (IDS).

The IDS functional requirements are related with the global use of MCA affordances. Therefore, one address these requirements by first discussing the MCA reference architecture illustrated in Figure 4.9. Acknowledging that several team members must operate the MCA, and given that information must be synchronized, the proposed MCA architecture adopts the client-server paradigm. Moreover, it follows the three-tier structure adopted in the development of many software applications, which allows discussing the MCA functional requirements according with the specific concerns of each of these tiers.

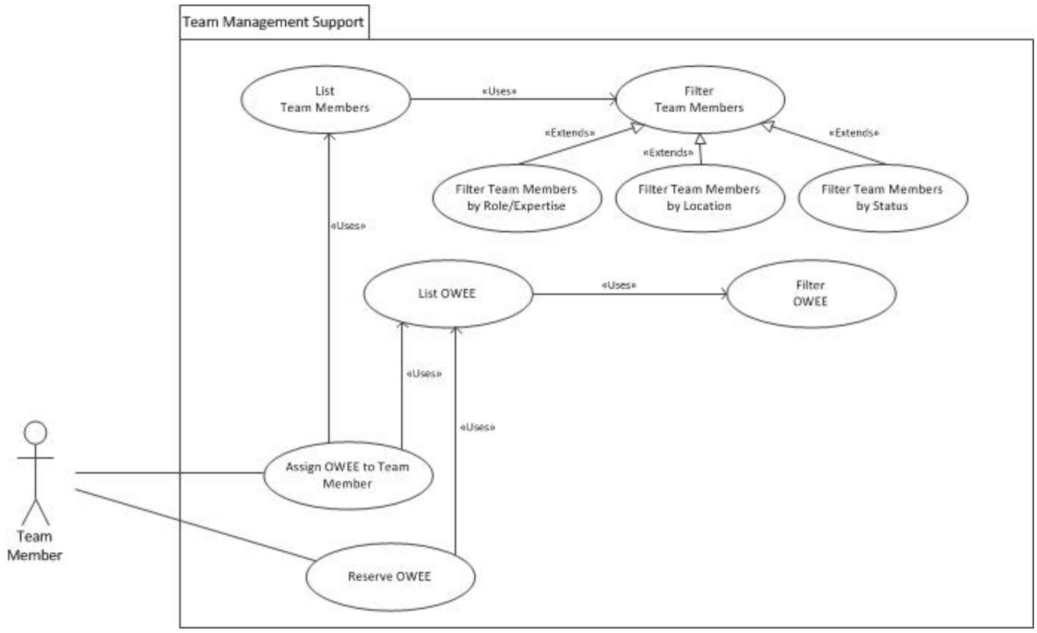


Figure 4.8: Use Case for the Team Management Support (TMS)

The presentation layer (Figure 4.9) offers a navigational model between the promising main screens of the MCA. The navigational model purports that MCA must refrain from excessive “menuing” and “windowing” and should minimize logical branches which, particularly in highly dynamic work contexts, may delay access to some information pieces and thus contribute for SA narrowing (Endsley, Bolté et al. 2003). Another, Graphical User Interfaces (GUI) design principle to consider, is the standardization of layouts, display modalities and interactive controls that constitute the presentation layer. Consistency across displays layouts aids the operator to develop a reliable mental representation of the MCA affordances, which allows her/him to predict the meaning and location of the displayed information across the available GUI. This predictability reduces cognitive load and, among other things, the time spent on accessing and operating the Graphical User Interface (GUI) (Endsley, Bolté et al. 2003).

Research on SA oriented design, has been pointing out that it is preferable to provide higher level global presentation of information (e.g. through a control panel), and that detailed information elements should only be provided upon request (Jones and Endsley 1996). Nevertheless, the higher level global presentation of information must accommodate the most salient aspects of information being conveyed, since they provide critical cues necessary to evaluate relevance and thus frame users attendance to them (Kaplan and Simon 1990).

Accordingly, OWEE displayed on the control panel must be furnished with their salient properties (or states) current values, since they will promote a more contextualized account of them; while leaving details, to be displayed in specific screens where such information granularity level becomes more pertinent (e.g. a more detailed presentation of an OWEE properties, may be required in the assignment screen in order perform an informed assignment). Determining just what critical cues should be conveyed in an overview screen is often quite difficult. Furthermore, such decision will be domain/task dependent, it is thus an effort that is in the realm of the target application domain immersion that will ground an MCA instance implementation.

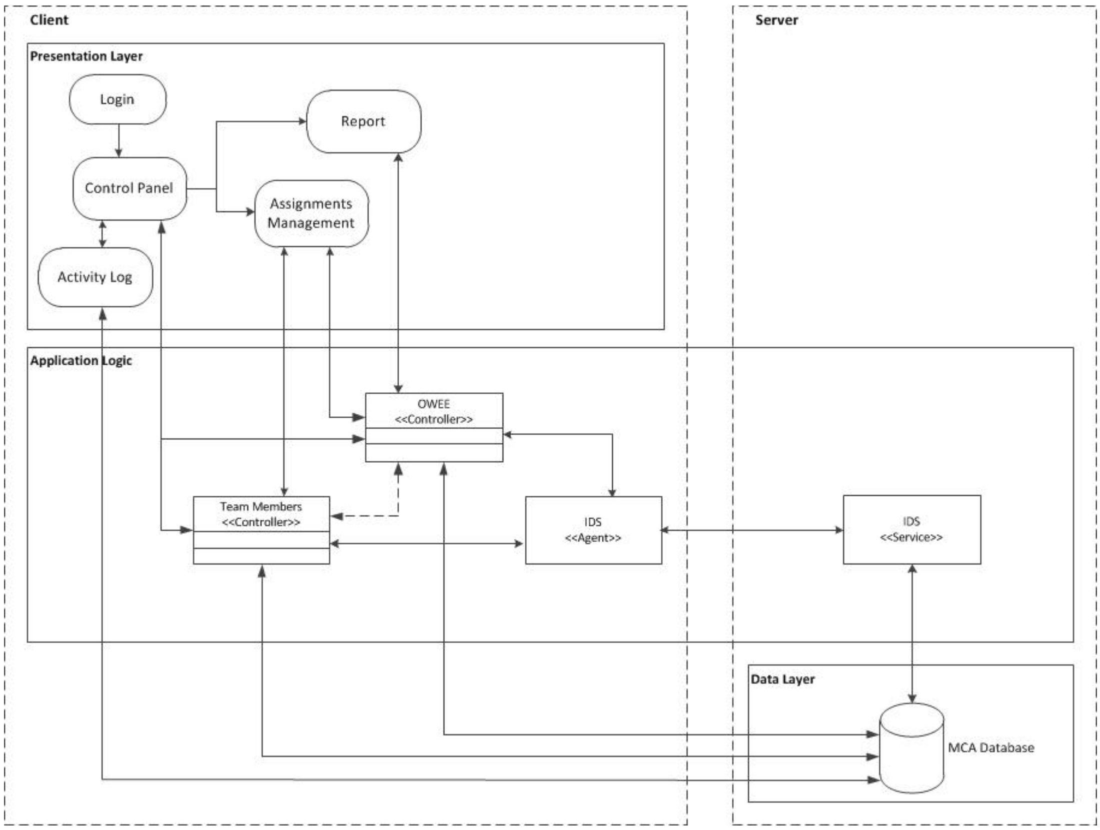


Figure 4.9: Mobile Collaborative Application (MCA) reference architecture

The MCA logic layer is composed by two control classes and a client side agent with is counterpart service on the server side. The Team Member Controller and the OWEE Controller classes are intended to keep the information on the MCA clients' presentation layer synchronized. Through these controller classes, operations like assignments and OWEE related reports are persisted into the server database and as clients access screens that load related information, it is readily retrieved form the server database.

Both assignments and reports are monitored by the IDS agent. The role of the IDS agent is to generate notification messages and passing them to the IDS server, which is responsible for delivering them to all clients, to provide accountability of on-going operational activity. This notification mechanism constitute a feedthrough information communication flow provided by the MCA. The role of feedthrough flows in applications that mediate collaborative work have been highlighted as paramount for establishing group awareness and construct meaningful contexts for collaboration (Hill and Gutwin 2003). Without this mechanism, team members would have to explicitly communicate with others to account for their actions positing an additional work and cognitive overhead (Khoshafian and Buckiewicz 1995).

Figure 4.10 illustrates the notification mechanism specified for the IDS using an exemplary UML sequence diagram. The diagram shown in Figure 4.10 illustrates the main workflow derived from the high-level functional requirements defined by the OWEERS use case diagram presented in Figure 4.7. According to the example provided, reporting information regarding an OWEE is accomplished by selecting it on the control panel, which will bring up a report screen with detailed information and report options (grounded on OWEE affordable operations, e.g. change one or more properties' values). The MCA persists the reported information on the server's database and issues an event through the IDS (client) agent to the IDS (server) service, which in turn broadcasts it to the other IDM agents. Receiving IDM Agents notify the end users (other team members) of their team mate report.

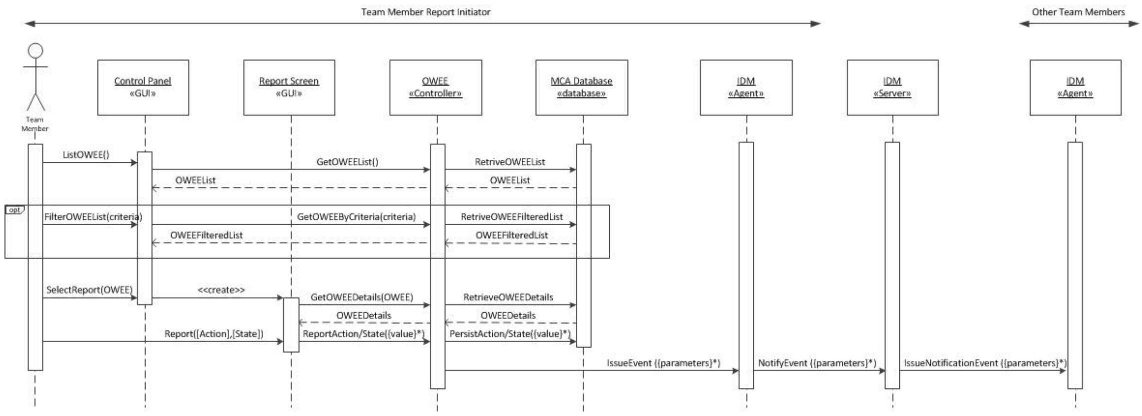


Figure 4.10: Information Dissemination Support (IDS) sequence diagram

Although notification mechanisms have been pointed out as a critical asset, particularly in tasks that exhibit tightly coupling (Turoff 1991), their implementation must minimize the disruption of on-going activities. Ultimately, it will be the operator's choice whether to readily account for the notification details or not, given its relative importance framed on the current information requirements and work context.

Individuals adopt several prioritization strategies regarding the attendance to incoming information, considering their perception of its relevance for own needs (e.g. immediately attend it, queuing it for latter attendance or disregarding it) (Hollnagel and Woods 2005). Moreover, considering the work and cognitive load, and potential stressors of the operational environment individuals may fail to notice notifications (Endsley, Bolté et al. 2003). Performance errors have been associated to the lack of SA where, although needed information was delivered, it was not attended to by the operators (Jones and Endsley 1996).

Therefore, despite notifications are delivered in real-time, MCA must provide the means for retrieving issued notifications for latter attendance. This justifies the introduction of the Activity Log Screen in the MCA navigation model shown in Figure 4.9.

Activity logs have constituted main tools to make sense of CIRM. Although they have been usually used for post-mortem analysis, their role in technological systems as a dynamic record of the unfolding activity provides the basis for their consideration in assisting TSA in real-time (Hale 1997; Turoff, Chumer et al. 2004). In order to become a valuable operational resource, MCA activity logs most provide a rich representation of the course of action (Endsley, bolté et al. 2003).

The MCA presented specification can support the study of MCA use in CIRM settings by integrating the derived functional requirements on the MCAE. This is elaborated in chapter 5.3.2.2.

4.4 SAMS - Situation Awareness Measurement Support

As discussed in chapter 3.1, there is no single overreaching SA measurement technique, since all existing ones present strengths and weaknesses, and their choice is bounded by the formulation of what constitutes SA and how it is developed on a given domain. As also put forward in 3.1.3, a more realistic approach to SA measurement is to combine several techniques in order to grasp the multi-dimensional nature of the construct. This involves its product dimension (i.e. the current mentally hold model of the situation) and its process dimension (i.e. the way the current model of the situation is achieved).

The product dimension has been mainly measured based on the administration of questionnaires (e.g. freeze and on-line probes). Addressing the process dimension can be accomplished by assessing the operational actions done by the actors using tracing techniques (e.g. work support systems logs analysis and team communications analysis).

Questionnaire-based techniques provide quantitative measures of SA, while tracing techniques may provide qualitative or quantitative measures. Although qualitative outcomes provide valuable insights for the analysis of SA, as those purported by propositional networks developed in the context of Distributed Situation Awareness (DSA) approaches to TSA (see chapter 3.1.1.2), quantitative measures offer a more readily mean to perform comparisons across experiments, particularly if they can be informed from fine-grained operational data.

However, the development of SA quantitative measures at the team level (TSA) had experienced a significant debate. The dichotomy between shared SA (in the sense of overlapping SA) and distributed SA (in the sense of the complementarity of SA) has fuelled a discussion over what constitutes representative TSA measures and what would be an effective measurement approach. Based on the existing literature, TSA measures may be framed on three categories: Shared, Collective and Holistic (Cooke, Salas et al. 2004).

Whereas Shared SA measures aim to assess the extent that individual SA overlaps; Collective measures posit that TSA measures are driven from the aggregation of individual SA; conversely, holistic TSA measures pursue representative measures that account for the role of team processes in SA development (e.g. communication, coordination). In contexts where each team member have independent skills and roles, it becomes particularly pertinent the consideration of holistic measures rather than aggregations or accounting only for the overlaps (Rentsch and Hall 1994).

The main concern surrounding the Collective TSA measures is, given the individual SA measures, how can they be aggregated to account for the team level. Mostly this had been accomplished by averaging individual scores (Cooke, Salas et al. 2000). However, the danger in averaging is that, due to team member variance, particularly in heterogeneous teams, the average result may be unrepresentative of the team members and the team as a whole.

Alternatively, TSA measures had also been devised from accuracy of the responses given by the majority of team members to SA elicitation questionnaires. Nevertheless, (Gualtieri, Fowlkes et al. 1996) measured the similarity of the knowledge associated with the situation model among team members with the same role on different teams, and found that within-role similarity was greater than within-team similarity, supporting the criticisms regarding the collective measures. The above arguments support that there is no definitive method for aggregating individual data to yield a team level SA measures since individual data is framed accordingly the SA requirements for each individual performance in the context of her/his task on a given situation/work context.

The endeavour of devolving holistic TSA measures is more recent and had strived for new operational TSA measures and respective assessment methods (Cooke, Salas et al. 2004). If teams are to be considered the unit of analysis than, measures that account for the team dynamics should be further developed. This consideration suggests that team-level properties may not be directly determined by team-member properties since team processes are a ripe source of variance specific to team level. Therefore, approaches that rely on individual team members, in order to determine TSA measures, will be insufficient for informing holistic TSA indices (Cooke and Gorman 2006). Assessing holistic TSA measures appears to be more suitably accomplished through work processes tracing techniques by embedding the collection of SA related information within the context of the task and teamwork.

The stated examination of the major concerns surrounding the definition of TSA measures and what constitutes proper measurement techniques, had informed the specification of the Situation Awareness Measurement Support (SAMS) subsystem purported by the reference architecture of the Microworld environment introduced on section 4.1.2 of this chapter.

The SAMS subsystem is specified by considering the support of the three types of SA measures: individual, shared and holistic. According to the most mature SA measurement techniques, individual SA measurement relies on the use of questionnaires and work process tracing. Shared SA measures can be accomplished by relying on the overlapped answers to questionnaires by the team members. Nevertheless, in order to depart from the perspective that all team members must have the same knowledge regarding the situation context, the overlap must be assessed by the consideration of only those answers to SA questions that refer to the knowledge that should be held in common. Finally, the support to holistic TSA measures is accomplished by work process tracing. Table 4.5 depicts the three types of measures and companion measurement techniques that informs the specification of the SAMS subsystem. The SAMS subsystem architecture is illustrated in Figure 4.11.

To support work process tracing, the SAMS requires that the OWES subsystem implements an interface which allows collecting the operational actions executed by the team. This information is persisted on the Operational Activity Log database to support SA measurements based on the operational level of teamwork.

Moreover, the current state of OWES is of high value to inform the formulation of relevant SA questions. Whether the questionnaires are administrated in real-time or through freeze probes, the questions that comprise them should be dynamic, if they are intend to elicit the currently hold SA regarding the current operational context.

Figure 4.12 presents an exemplary mock-up of a SA questionnaire. As afore mentioned, both the questions and the range of answers must be dynamically generated considering the context of the OWES.

In order to minimize the disruptive effects of questionnaires' administration, the answers should be facilitated by departing from open field text controls toward the use of radio buttons, check boxes and list boxes providing alternative answers. This also relieves grading the answers and scoring devised SA measures. Toward these ends, the delivered question statement, range of answers, given responses and their accuracy regarding the current OWES state are persisted in the questionnaires database.

The definition of SA questions and SA measures will be, as already pointed out, inevitably bounded by the target application domain and team task. In the experimental design presented in chapter 6, a set of SA measures addressing the individual and team levels presented in Table 4.5 were devised as dependent variables of the study.

Table 4.5: Specifications for Situation Awareness Measurement Support (SAMS): Measures and Measurement Techniques

Measure Type		Measurement Technique	Requirements
Individual Level	Individual SA	Questionnaires	- Incorporate questions regarding the task and team context - May be administrated on-line/real-time or through freeze probes
		Work Process Tracing	- Collect individual operational activates representative of her/his SA
Team Level	Shared SA	Questionnaires	- Requires the identification of common questions regarding other team members
	Holistic SA	Work Process Tracing	- Collect individual operational activities accounting for the overall team operational activity

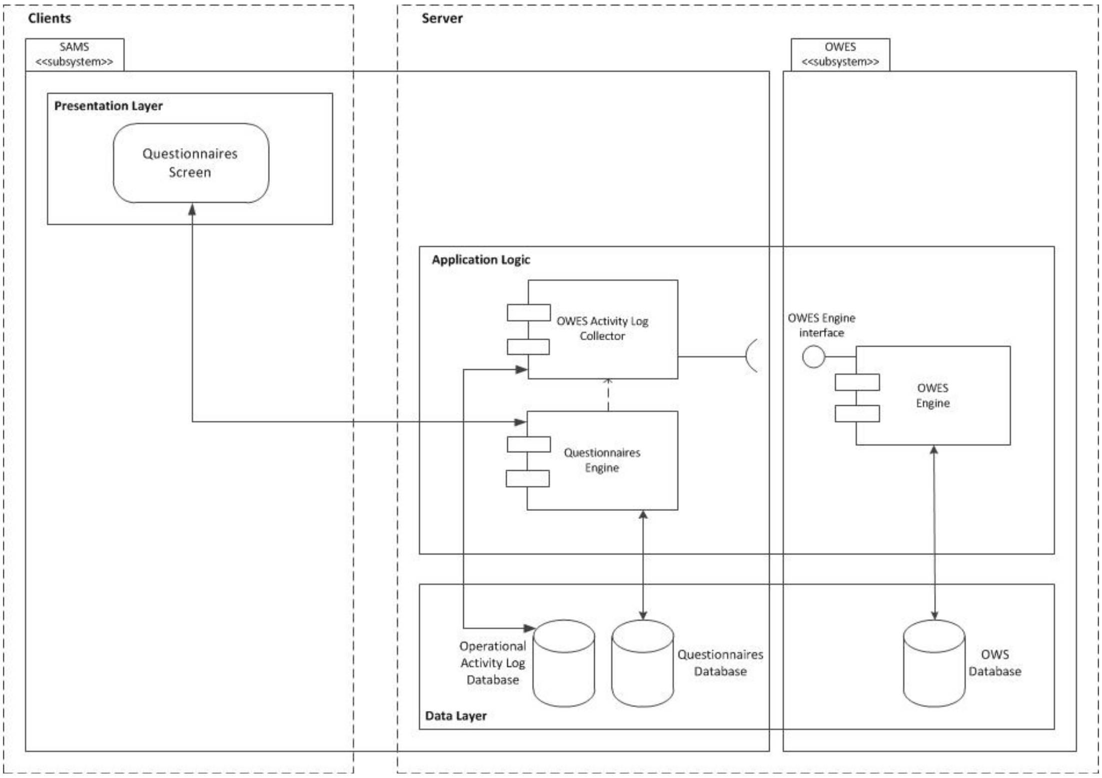


Figure 4.11: Situation Awareness Measurement Support (SAMS) subsystem architecture

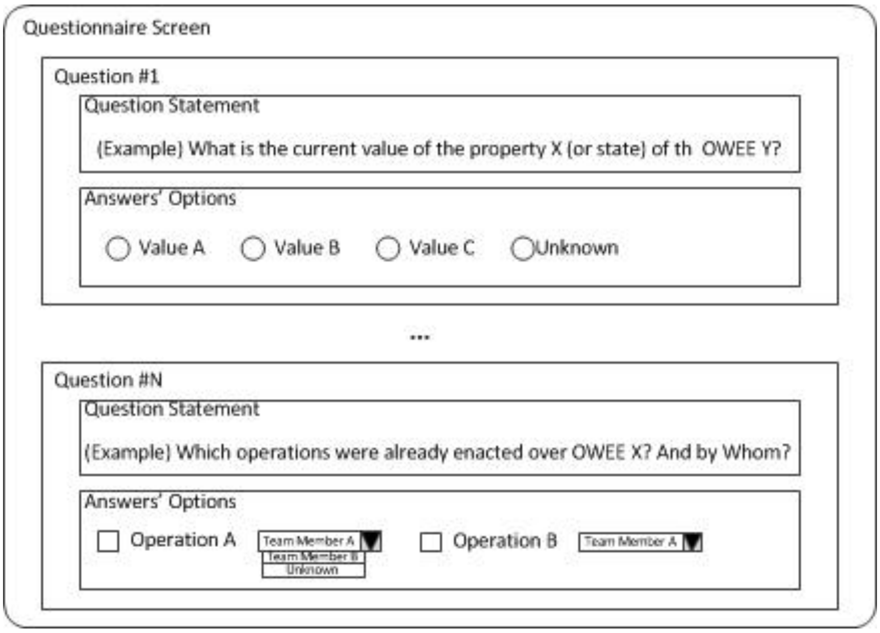


Figure 4.12: Mock-up of the SA questionnaires user interface

4.5 ECM - Experimental Control Manager

As an experimental instrument, Microworlds must support experiments underlying protocol. At least to some extent, given that experimental protocols have a broader reach, since they encompass, among other things, the definition of materials to be used, characterization of participants profiles, contents of briefing sessions, data collection and analysis methods, description of the experimental apparatus, and aims of the experiment. Moreover, experimental protocols also frame the definition of controlled, independent and dependent variables that guide the study of the phenomena of interest. This matter is deeply explored in chapter 6.

Nevertheless, some experimental conditions purported by the experimental protocol can and should be controlled by the Microworld. One discusses them in the context of the specification of the Experimental Control manager (ECM) subsystem.

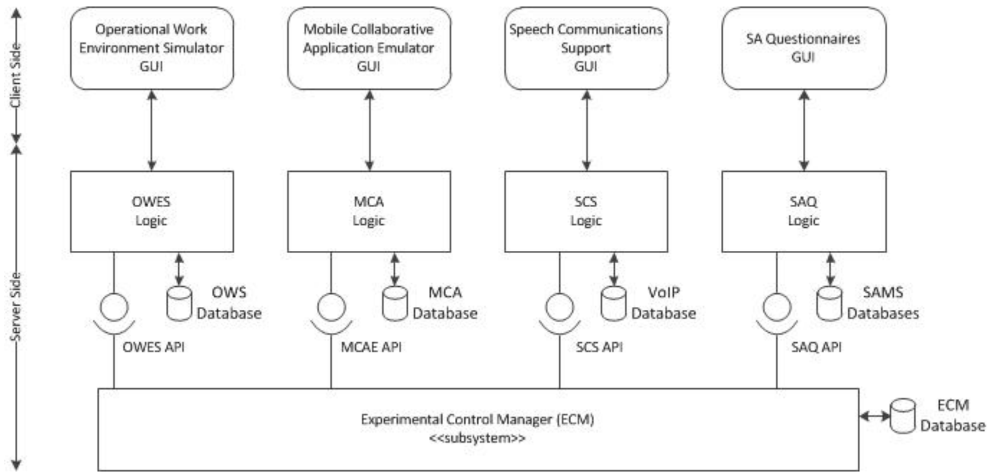


Figure 4.13: The Experimental Control Manager (ECM) subsystem in the context of the overall Microworld environment

Figure 4.13 situates the ECM subsystem on the overall Microworld environment. According to the diagram, the ECM subsystem require that each of the Microworld's subsystem specifies an Application Program Interface (API) at its logic layer.

A natural requirement for these APIs, common to all subsystems, is that they must provide the means for deploying the number of clients according to the number of team members participating in the experiments.

Additionally, as discussed in chapter 3.2.2, experiments often entail a comparison between features of promised solutions. Therefore, MCA functional features, available through the MCAE Graphical User Interface (GUI), may be enabled or disabled given the experimental protocol. As an example of how the experimental protocol may guide these requirements, one can bring up the experiments described in chapter 6, contrasting the performance of control teams, which had only available the SCS subsystem to support team communications (thus deprived from the MCAE), with the performance of teams that besides the SCS had the MCAE available to assist teamwork in the experiments.

Moreover, the control of the experimental apparatus purported by the experimental protocol may constraint the functions provided by the OWES GUI and the SCS GUI to the participants. For instance, some OWES affordable actions' type (OWEEAAT) or OWEE affordable actions (probe and operational actions) introduced in section 4.2 may be enabled or disabled in the OWES GUI. This allows the operationalization of experimental control exerted over the team task. Likewise, the SCS subsystem may be configured for supporting conference calls or one-to-one conversations, as put forward in section 4.3.1, given the target application standard communication channels.

Regarding the SA questionnaires API, it must also provide the means to follow the experimental protocol, which may posit if whether the SA questionnaires will be continuously available or delivered at specific moments (either simultaneously with the team task or though freeze probes). Moreover, in the case of administrating the questionnaires in specify moments, it requires the support of the triggering condition. For instance, it may be on specific moments in time or given the overall state of the OWES after some triggering action. Toward that end the OWES API, must provide the means to collect the OWES state and on-going activity.

An additional requirement must be considered in the specification of the ECM subsystem. If Microworld experiments are intended to promote the manifestation of teamwork, it must deliver to the teams a task scenario, purported by the experimental protocol.

Different task scenarios will exercise different aspects of collaborative work and technology usage. As discussed in chapter 3.2.2, those scenarios may be devised to account from social interaction, regarding on how participants interact with each other, to the role of mediation effects of collaborative work support systems (such as MCA) (Damianos, Hirschman et al. 1999).

Representative task scenarios musts be carefully crafted to accommodate a characterization of the exogenous dynamics towards the OWES, leading it to depart from its expected endogenous dynamics specified by its state space (Bai, Tsai et al. 2002).

The present characterization of task scenarios builds upon three main constituents. The first one holds a description, in natural language of the task context, that is conveyed to the team through the OWES GUI (which constitute an additional requirement for the definition of the OWES API). For instance, it may give notice that one or more OWEE (e.g. a critical piece of equipment) is malfunctioning or became unavailable, and thus triggering the team CIRM process.

The second constituent is the set of initial states that the OWEE holds in order to reflect the purported task scenario.

And the third constituent holds a collection of events that constrain the OWES endogenous dynamics. Such events directly affect the OWES state-space by overriding the OWEE state transitions defined in the OWES engine state matrixes. This naturally constitutes an additional requirement to be supported by the OWES API. As an example, a particular OWEE (e.g. a physical equipment) affordable operations (e.g. reset or switch on) that expectedly lead the OWEE to a specific state (e.g. operational or active) may contrary direct the OWEE to an unanticipated state (e.g. by keeping it inactive), if this stands for the pursued exogenous dynamic that characterizes the underlying dynamics purported by the task scenario.

The task scenario is hold by the ECM database, which provides the support for persisting several scenarios. The characterization of the task scenario’s events that bound its underlying dynamics, is further clarified in Table 4.6 through the consideration of three main attributes: the targeted OWEE, the event pre-conditions and the event post-conditions.

Table 4.6: Task Scenario's events characterization

Task Scenario's Event Attributes	Description
Targeted OWEE	Defines the OWEE impacted by the event
Event Pre-Conditions	Establishes the triggering criteria for the event
Event Post-Conditions	Defines the outcomes of the event

The targeted OWEE attribute holds the identification of the OWEE over which an external influence on its state space is intend to be exerted by an event of a given task scenario. The triggering criteria that establishes if the event overrides the OWEE intrinsic dynamics is hold by the event pre-conditions attribute. This criterion may have different natures. It can establish a specific moment in time to activate the event, it can follow from the enactment of a particular action over the OWEE, or it can be based on particular states exhibited by OWEE in the OWES. For instance, if a particular OWEE changes from state A to B, this may constitute the reasoning for an external intervention on the state space in the targeted OWEE. The task scenario's event post-conditions attribute comprise the outcome consequences proceeding event's occurrence. One put forward that, according to the intended task scenario underlying dynamics, the outcomes may be three-fold: changing the targeted OWEE state, inhibiting (or re-establish) some of the targeted OWEE affordable actions (and thus preventing/allowing it to achieve some states) or changing the states of related OWEE.

The above description on how task scenarios must be accommodated by the Microworld provides a set of valuable requirements to inform the definition or the Microworld subsystems' APIs, which are summarized in Table 4.7.

The ECM subsystem specification and associated requirements for experimental control presented in this section, concludes the Microworld reference model specification.

Table 4.7: Requirements for the Microworld subsystems APIs

Microworld subsystems APIs	Requirements
Operational Work Environment Simulator (OWES)	• Allow control of the number of clients to be deployed • Deliver task scenario description • Retrieve OWEE states • Retrieve actions enacted over OWEE • Enable/Disable OWEE affordable actions • Change OWEE state
Mobile Collaborative Application Emulator (MCAE)	• Allow control of the number of clients to be deployed • Enable/Disable MCA functional features • Enable/Disable MCA availability
Speech Communication Support (SCS)	• Allow control of the number of clients to be deployed • Enable/Disable SCS functional features • Enable/Disable SCS availability
SA Questionnaires (SAQ)	• Allow control of the number of clients to be deployed • Enable/disable SA questionnaires according the triggering conditions

4.6 Discussion

This chapter puts forward the specification of a set of building blocks that should be accounted for developing a Microworld. Although two of the proposed building blocks are heavily coupled with the present research interests, namely the Mobile Collaborative Application Emulator (MCAE) and the Situation Awareness Measurement Support (SAMS), their specification still remains domain independent and brings out considerations that constitute valuable guidance for Microworld development initiatives addressing the analysis of teamwork on other contexts/domains.

A fundamental tenet of Microworld based experiments is promoting team operational activities as close as possible to the ones performed in the field. Thus, experimental control departs from more traditional laboratory experiments, which highly restrict (script) the participants' activities. This level of control may be adequate in certain cases, e.g. usability studies, but hardly addresses the openness of human behaviour in broader contexts such as collaboration in CIRM settings.

In the context of collaborative applications (such as mobile collaborative applications) development and evaluation, Microworlds allow the collection of application's usage related data framed on the demands posited by an illustrative task scenario of the considered domain. However, scenarios must be carefully crafted to promote the manifestation of phenomena of interest. Phenomena of interest, in the context of teamwork, are for instance: team efficiency, coordination, and communication patterns, which given the current work aims had been considered in regard of team situation awareness (TSA).

Besides addressing the product dimension of TSA (the knowledge that team members hold regarding the task context) through direct TSA measurement techniques, that mostly had relied on the use of questionnaires (whether administrated simultaneously with the task or through freeze probes), the devised Microworld reference specification also accounts for individual and team work processes tracing in order to supports holistic measures of TSA based on the operational level of teamwork, which had been pursued by SA related research.

These holistic measures are accomplished due the capacity of the Microworld environment of capturing large sets of experimental data that may range from, the key-stroke level to speech communications enacted between the team members, while preserving such data in the context of the teamwork, task scenario and operational environment state.

This fine-grain contextual information also affords assessing the impact of promised mobile collaborative applications (e.g. whether it increases or decreases TSA), as well as, a keen understanding on how particular application features (or how they are used) constitute a valid contribution for SA support. This last endeavour is particularly valuable since it enriches the experiments beyond the more traditional emphasis on causal epistemology, bringing forward an understanding of the factors that underlie hypothesis acceptance or rejection and thus informing the design-evaluation cycle of (mobile collaborative) application development.

The Microworld specification presented in this chapter guided the implementation and deployment of a Microworld for a particular application domain, Help Desk Teams, earlier introduced in chapter 1.3, which is further described in the next chapter.

5 Case Study - Help Desk Teams

In this chapter it is presented an implementation of the Microworld specification provided in the previous chapter 4. Such implementation was done by a consideration of a concrete target application domain, which, as introduced in chapter 1.3, refers to Help Desk Teams operating in Critical Incident Response Management contexts. The motivation for the select target application domain is discussed in terms of its representativeness regarding the current research aims. The chapter details the adopted domain immersion methodology, which had the ultimate goal of instantiating the presented Microworld specification in the application domain. The functional features of the each subsystem that comprises the proposed Microworld reference model are depicted, since they support the conducted experimental trials on the target domain, later described in chapter 6.

5.1 Motivation

The commonly referred Help Desk Teams (HDT) are the operational organizational units responsible for maintaining organizations' Information and Communication Technology (ICT) infrastructure operative. The ICT infrastructure is composed by a multitude of equipment's such as, networking devices, backup units, printers, servers and desktop computers, as well as, their respective hardware parts, firmware, software and operating systems.

The selection of HDT as the target application domain for grounding the case study, is herein discussed by putting forward the main factors that establishes its representativeness regarding the present research aims and the respective relevance regarding the research question.

In medium and large organizations the complexity that surrounds the work performed by HDT lays on the relation between the ICT infrastructure and the diversity of organizational processes that it supports. Some of the work enacted by HDT may imply significant consequences to the all infrastructure operational state and therefore for organizational activity stability.

A large number of disruptive events are perceived as critical to organizations that heavily support their activity on the ICT infrastructure, for instance major server failures, critical software services break downs (e.g. email, web proxy, domain naming) or loosing network connectivity, particularly those that occur from more untypical, unprecedented or unanticipated factors. In such contexts HDT work departs from their routine and standardized activities and is usually characterized as highly pressed by time constrains and impose high stress levels, which may impair the ability of HDT to make an effective integrated response to contain and mitigate disruptive events consequences.

Despite the HDT hold trained personnel and business continuity and contingency plans to address some disruptive situations, two main issues may strongly condition their operational effectiveness in CIRM operational contexts. First, standard procedures typically do not accommodate the whole variety of contingent factors that may arise. Second, many disruptive situations require bringing together tacit knowledge from multiple participants who, as a team, may be more capable to assess the situation and develop creative solutions or temporary workarounds.

As so, the operational activity of HDT in CIRM contexts, requires that team members make use of the collective expertise and experience and therefore, real-time collaboration, constitute a fundamental requirement for carrying out operational teamwork. Moreover, since some disruptive events may involve the loss of connectivity in some network segments or physical equipment damages, the typical remote control panels and other systems' administrators tools that rely on remote communications may reveal useless. Such cases, require that team members move toward physical locations (where the infrastructure affected resources are placed) and diagnose in situ what has occurred or perform some local recovery actions (e.g. replace a RAID disk on a server, crimp a network connector on a network cable or replace some equipment, or parts). This distributed work setting with situated and real time operational information communication/management requirements to promote Team situation Awareness (TSA), had provide the reasoning for the study of the role of Mobile Collaborative Applications (MCA) to assist HDT operational work in CIRM and further explore their role regarding TSA development.

The evaluation of promising MCA to assist such work contexts is challenging, since one cannot naturally evaluate their usage in real CIRM instances. Although, field simulations could be considered, they will require the use of a number of observers, as many as the displaced team members, in order to frame MCA usage in the context of the operationally enacted activities. This fact, along with the shortcomings of observation and field based evaluation techniques, particularly acute in early stages of promising solutions design, which deeply discussed in chapter 3.2.2, had driven the motivation for the adoption of a Microworld environment as an experimental paradigm to support the evaluation of MCA introduction on HDT performing in CIRM contexts.

The next section presents the adopted domain immersion strategy that furnishes the Microworld specification, provided in chapter 4, toward its instantiation framed on the HDT operational work domain underlying CIRM.

5.2 Domain Immersion Methodology

Toward a deeper understanding of the target application domain three domain immersion techniques were complementary used with domain practitioners. Domain immersion has thus relied on the use of semi-structured interviews e.g. (Smith 1995; Davis, Dieste et al. 2006), questionnaires e.g. (Lloyd, Rosson et al. 2002) and workshops for discussing operational scenarios e.g. (Stewart and Shamdasani 1990). Each was applied to two HDT from two medium size organizations. Such teams present the following composition: the first of the teams comprehended one coordinator, two senior technicians and two junior technicians; while the second team consisted of one team coordinator and two senior technicians.

Table 5.1 outlines the three phases that comprise the devised domain immersion methodology and further detail the respective goals of each one.

Table 5.1: Domain immersion methodology: phases, techniques and goals

Phase	Technique	Goals
1.	Semi-structured interviews	• Identify what are considered critical incidents in the scope of HDT work • Unveil actual/current work practices in CIRM • Elicit existing CIRM support systems and its usage
2.	Questionnaires	• Rate a set of requirements derived for informing MCA main promising features implementation that account for HDT operational work
3.	Workshop	• Understand the interactions at team level in the scope of CIRM through CI scenario discussion • Inform the implementation of the Microworld Environment's specification on the domain

The set of questions that guided the semi-structured interviews are presented in Table 5.2.

Table 5.2: Semi-structured interviews questions

Question #	Question formulation
1	Which situations may be described as critical incidents?
2	Which preventive and diagnosis practices are currently being used?
3	Which formal and informal procedures have traditionally been deployed?
4	Which communication systems are used during disaster recovery?

The conducted interviews indicate the most critical incidents are related with server failures (mostly derived from disk failures) and connectivity losses in specific network segments (mostly derived from router/switch failures) which compromise a wide variety of network services. A concern with more untypical and extreme problems was also reported, "[...] like a flood in the basement where some of the equipment is located [...]".

Such critical incidents may arise or scale from a number of vulnerabilities that were pointed out. For instance, not all pieces of equipment have a spare stock or coverage of Service Level Agreements (SLA) with suppliers.

It was also pointed out, that existing preventive practices include monitoring active network elements through control panels (e.g. fed by Simple Network Management Protocol – SNMP – messages), and having alerts displayed and emailed to the team members. Nevertheless, as previously stated, team members confirmed that many disruptive situations require them to move to the physical location of a particular infrastructure's resource to better understand its state and/or enact specific actions only afforded locally.

The team that exhibit both senior and junior members made such distinction on the basis of their experience and technical skills. For instance, both teams report that some members may handle Windows and LINUX based software services and applications, while others are proficient in only one operating system; moreover some team members hold certifications in particular domains such as database administration or network management. Their expertise consequently frame the credentials to manage or operate specific infrastructure's resources.

It was perceived that it was in fact experience and expertise that most often ground some temporary workarounds that will minimize the impact of the CI until a more definite solutions can be deployed. The strategies to maintain the required level of TSA to coordinate the CIRM include mainly quick informal meetings and phone calls.

Although both teams use some Trouble Tickets software in their more routine operations, they also realized that such software applications are almost irrelevant during non-routine situations. In such cases, Trouble Tickets are sometimes used for incident opening and only occasionally for post mortem annotations to close incidents. Nevertheless, team members stated that such usage does not accommodate a fine grain documentation of the CI and deployed CIRM, therefore provide a minimal impact on planning or supporting future situations, which will rely on the developed collective memory.

The lack of a medium for situated, real-time, up-to-date information sharing and persistence, lead teams to consistently report the concern of not documenting the fine-grain activities underlying CI diagnosis and response strategy enactment. The persistence of such information was considered a valuable asset to build organizational memory and provide future guidance, especially because human resources in HDT may rotate and past experiences often become, at least to some extent, partially lost. Accordingly, it was pointed out that the use of a MCA may provide such medium to capture in real-time situated enacted operational activates. Thus, MCA despite potentially promote TSA, although both teams had notice the high value of speech communications (through phone calls) to that end, they may more readily addresses the concerns of document the CIRM.

As previously posited, one had relied on the use of a questionnaire to capture the pertinence of a set of high level requirements to further support the implementation, in the HDT domain, of the MCA specification provided in chapter 4.3.2. Table 5.3, presents the requirements accounted in the questionnaires administrated to each of the team members of the two teams engaged, in a total of seven of the eight interviewed participants (since the coordinator of the second team was not available). The rating scale for each requirement was from 1 to 4, meaning: 1 – Not perceived as important; 2 – Less important; 3 – Important; and 4 – Very important. The questionnaires result scores are presented in Table 5.4, order from the highest to the lowest score.

Table 5.3: Requirements that composed the questionnaires administrated to HDT members

Req. #	Requirement	Description
1	Provide minimal usage overhead	Minimize the number of screens and logical branches between them
2	Facilitate monitoring	Afford notifications of operational actions and display notifications history log
3	Communication support through shared displays	Use a shared screen to all the team members
4	Assist understanding of teamwork context	Display team members assignments
5	Help perceiving who is involved	Display logged-in team members
6	Improve diagnosis time	Visually represent the equipment states
7	Document fine grain operational activities	Support reporting enacted operations over equipment
8	Support attendance for several incidents simultaneously	Allow switching between different CIRM instances contexts

Table 5.4: Requirements questionnaires scores

Req. #	Number of answers in each step of the rating scale				Average	STDEV
	1	2	3	4		
1			**1**	**6**	**3.86**	**0.38**
4			**2**	**5**	**3.71**	**0.49**
6			**2**	**4**	**3.67**	**0.52**
7		**1**	**2**	**3**	**3.33**	**0.82**
2			**6**	**1**	**3.14**	**0.38**
8	1		5	1	2.86	0.90
3		2	5		2.71	0.49
5		4	3		2.43	0.53

Considering a threshold of *3*, for distinguish the most pertinent requirements, it can be noticed that the requirements number 1,4,6,7 and 2 were considered the most relevant. Requirement number 1 mainly addresses practical usability concerns, naturally given the demands posited to teamwork in CIRM the MCA usage overhead should be minimized and therefore avoid complex user interfaces. Also in the usability domain, the score of requirement number 6 points for the need to readily perceive the state of the equipment involved in the CI (e.g. by representing their states through representative icons). Regarding the scores of the requirements number 2, 4 and 7, they appear to meet the desire for the MCA to support tracking of the operational activities enacted by team members. Apparently, given the score of requirement 8, the support for attendance of simultaneous incidents was not perceived as most relevant, possibly due if a really critical incident occur in will have a top priority on the HDT work agenda. The low score of requirement 5 may be explained by the fact that the consulted HDT are not very numerous and thus they already know who should/will be engaged in CIRM for a CI instance. The also lower score of requirement number 3 is in line with some research results that had put forward that the use of shared displays does not directly positively correlate with team performance and awareness (Bolstad and Endsley 2000), in heterogeneous teams, since each team member hold specific information requirements to accomplish her/his work.

The last phase of the adopted domain immersion methodology comprised a workshop for the discussion of a concrete CIRM scenario. To depart from the rationalization exhibited in the interviews when discussing hypothetical cases, the workshop scenario was previously developed with the collaboration of team coordinators targeting to provide to the teams a concrete situation and elicit their collective process on addressing it. Table 5.5 presents the scenario description provided to the teams and the discussion triggering question.

Table 5.5: Discussion workshop scenario description

Scenario
"From several rooms, were reported the loss of network connectivity. Some technicians were notified by email, while others received several complains by phone. The senior technician that received some of this complaints suspects from the central switch located on the main building." What will be the best course of action to diagnose and recover considering the usual practices and resources?

Despite notes were taken regarding the team on-going scenario discussion for future reference, one had provided participants with a set of paper matrixes that were intended to unveil the web of relations and interactions that exists at different dimensions of the CIRM endeavour. This approach was rooted on works on CI analysis that had adopted the Systems-Theoretic Accident Modelling and Processes (STAMP) framework (Leveson, Daouk et al. 2004) (introduced on chapter 2.1) and on the propositional networks purported by the distributed situation awareness oriented analysis (Stanton, Stewart et al. 2006) (introduced on chapter 3.1.1). Both initiatives target to provide concept graphs that establishes the relations (links) among elements (nodes) at different abstraction levels (accommodating for instance, operators, goals, procedures, resources, etc.) that bound the collective task work.

Although the use of the provided paper matrixes follows similar goals, they hold a more confined scope and are more directed to capture fine-grain relations that bound the operational level of HDT teamwork. Figure 5.1 presents a photography of a set of matrixes that were filled by the participants during the scenario discussion workshops.

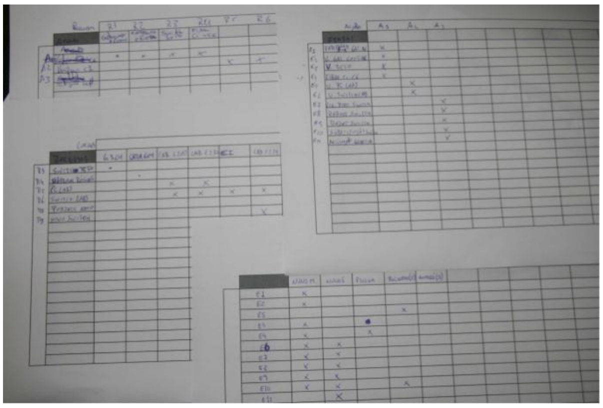


Figure 5.1: Paper matrixes used to unveil the dimensions of HDT operational work

Participants freely fulfil a number of matrixes as they (and the researcher) acknowledge the relevance of unveiling the relations between different dimensions (accounting different abstract levels), of their collective practice regarding the provided scenario.

Participants filled one matrix relating the physical equipment (e.g. servers, routers) to their current locations, which had supported the discussion on how team members may be distributed to address the CI scenario. While filling this matrix it was noticed that more junior technicians hold a less precise model of the overall ICT infrastructure, and thus are the ones that may benefit more from an explicit indication of the current location of a piece of equipment. Following that discussion an additional matrix was created by the participants for exposing the connections between the network's equipment, which had provided a model of the part of the ICT that should be accounted. This had grounded the debate regarding the diagnosis endeavour (e.g. what could the faulty equipment or a broken link).

The reasoning beyond task assignments had led participants to develop an additional matrix relating team members, with their specific expertise, since as put forward in the interviews their experience, certifications and authentication credentials frame their ability to operate specific equipment. Through an additional matrix, the tasks were further drilled down to the operational actions that should accomplish in order to check and/or re-establish a (possibly) faulty equipment (e.g. check router status' led indicators, perform a reset).

Although, the discussions were carried out with team members collocated it was evident that the CI scenario impose a tight level of work coupling, since the outcome of each assignment will inform the next step of the diagnosis and recovery efforts undertaken in the CIRM, and thus team communication plays a fundamental role. Nevertheless, has it was pointed in interviews the considered HDT teams rely on phone calls to that end, and unlike other CIRM teams (e.g. firelighters or some industrials maintenance teams) do not adopt conference calls or multi-point radio communications.

The matrixes that emerged from the scenario discussion lead to some considerations regarding the implementation of the Microworld environment, which are put forward in the next section along with the description of the developed Microworld environment instance.

5.3 Microworld Instance

According the characterization of the operational level of work of HDT put forward in the previous section, the Operational Work Environment Simulator (OWES) should accommodate a representation of the several rooms over which the ICT's equipment is placed. Those locations comprise the instances of the Operational Work Environment Locations (OWEL) abstract class depicted on the OWES model (chapter 4.1.2). Moreover, in order to support a virtualization of the team members locomotion between this rooms, an instance of the Operational Work Environment Affordable Action types (OWEAAT) abstract class must be accomplished to endow the OWES with the functional feature: "move to a room". The practical implementation of such classes is more readily understood by the description of the OWES provided Graphical User Interface (GUI) conducted in the section 5.3.1 of this chapter.

The ICT's infrastructure equipment, will constitute the instances of the Operational Work Environment Elements (OWEE) abstract class. Following the OWES specification, their characterization is presented in Table 5.6.

Table 5.6: Operational Work Environment Elements (OWEE) instances on the HDT domain

OWEE	Properties [Range of Values]	States	Affordable Probe Operations	Affordable Operational Actions
Server	Operative [True, False]	- Operating (Operative=True) - Malfunctioning (Operative=False)	Check state	- Restart - Update - Replace -(Re) Connect
Router				
PC				

The main OWEE that grounds the team operational work were considered, according those elicited from the matrixes devised from the workshops discussion (presented on the previous) were, the *servers*, *routers* and end users *Personal Computers* (PC). Although, they could be more comprehensively characterized by accounting for a broader number of properties that characterizes their real attributes (e.g. memory and hard drives capacity, firmware version, etc.), the workshops discussion regarding teams' CIRM deployment had unveiled that what is relevant is whether if the equipment is operative or not, and thus one had considered the *Operative* OWEE property. Given this simplification of the OWEE characterization and the claim that relevant information to be passed to the team level is the report of team members on OWEE current state, the property readily translate into the OWEE considered states: *Operating* or *Malfunctioning*.

In order to support a virtualization of the probe operations that are enacted toward the diagnose of OWEE state (e.g. attended to status led indicators of routers or PCs' network cards) the OWES implements the Affordable OWEE Probe Operations, purported by the OWES model through a *check state* functional feature, which will deliver to the team member whether the equipment is on the *Operating* or rather on the *Malfunctioning* state.

Following the specification of the OWES model, the intrinsic dynamic of the OWEE is framed according two main sources of influences on its state space.

The first, is the inherent outcome of the operational actions enacted over it. Each of the equipment considered Affordable Operational Actions may change its state.

The matrixes from the workshops' discussion had brought the consideration of four main operational actions that team members undertake while diagnosing and recovering an equipment. Thus the considered Affordable Operational Actions were: 1. Restart, which addresses the common procedure of restarting an equipment and check if the malfunctioning state still hold; 2. Update, this operation is intended to mimic another common procedure which is updating some of the devices' software or firmware versions; 3. Replace, which pretends to express the substitution of some equipment (or parts, e.g. network card, hard drive); and at last 4. (Re)Connect, that supports the (re)connecting the equipment to other one (e.g. change a PC connection to another router). The linkages, as well as, the underlain dynamic impacts that frame the web of connections underlying the ICT infrastructure are kept by the OWES by accounting for the specifications provided on Table 4.3 and Table 4.4 in chapter 4.2. This web of connections support the net effect of the impact of a state change in one equipment regarding the others that are linked to it (e.g. if a router is on the *malfunctioning* state the PCs linked to them will lack network connectivity, framing their state as also *malfunctioning*).

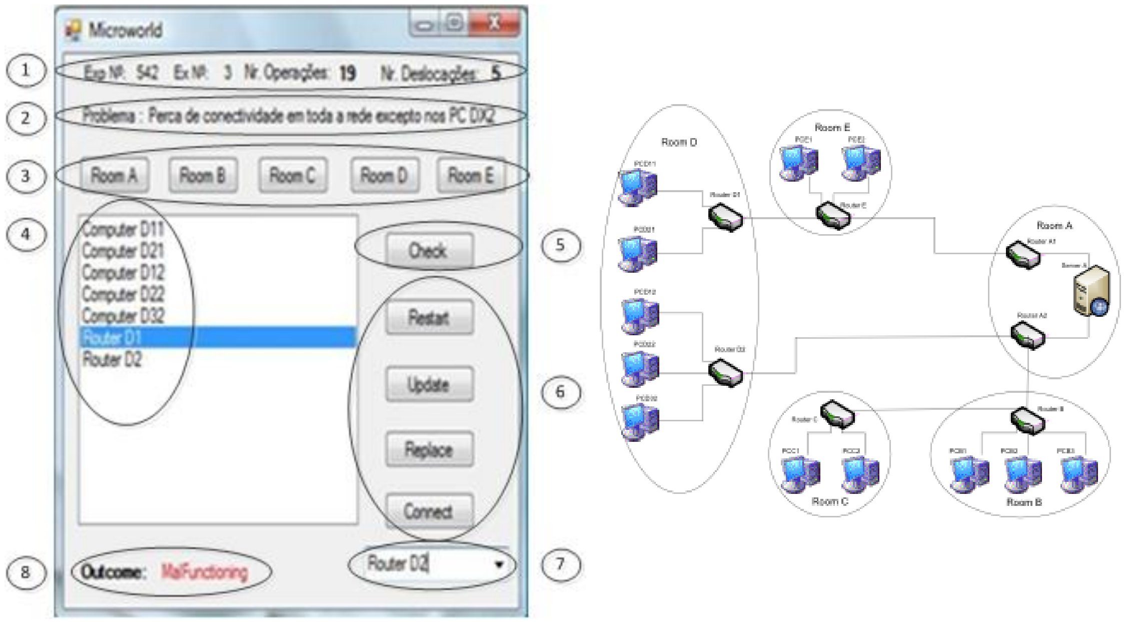
The second source of influence on the equipment state space will come from the external influences brought out by a particular CI task scenario, which as described in chapter 4.5, will constrain its intrinsic dynamic, leading it to unexpected states. Both the initial states and collection of events purported by a particular task scenario will lead the overall ICT's infrastructure beyond its regular operative state. The devised task scenarios that grounded the experimental trials enacted over the herein presented Microworld instance, are further depicted on chapter 6, that provides a comprehensive description of the conducted experiments.

The presented instantiation of the Microworld specifications, framed on the HDT domain, grounded the development of the instances of the Microworld subsystems purported by the Microworld reference model, each of those is further detailed in the next subsections.

5.3.1 Operational Work Environment Simulator

Given that the Operational Work environment Simulator (OWES) was comprehensively specified on chapter 4.2, and that the implementation followed the presented specification, framed on the HDT domain, through the considerations provided in the previous section, one support the presentation of the OWES functional features on the clients' Graphical User Interface (GUI).

The OWES clients' GUI Graphical User Interface (GUI) in Figure 5.2a. The OWES affords to load different network architectures, as the example depicted on Figure 5.2b.



a. b.

Figure 5.2: Operational Work Environment Simulator (OWES): a. client graphical user interface; b. Exemplary network architecture loaded on the OWES

The description of the OWES GUI is supported on its main constituent parts highlighted be an elliptic surrounding identified by a side number.

The first GUI part holds, by the presented order, respectively: the number of the experimental trial (this identification is only for informative purposes for the researcher), the task scenario number (this concerns to the identification of the task scenario that is loaded), the number of operations and the number of locomotion's made by the team members were also presented to the team members, since the experimental design, later described in chapter 6, posit them as criteria that teams should minimize while accomplishing the experimental trial, to endow the experiments team task endeavour with an enhanced realism regarding the operational practices.

The CI's task scenario description was delivered to the team in the GUI part number 2.

The third GUI part, concerns the implementation of the locomotion of team members between the rooms derived from the underlying network architecture. By pressing a room button the equipment that is placed in that room is listed in GUI part number 4. This action freezes the screen for a while for two main reasons: accounting that such locomotion in real settings are not instantaneous and penalize teams (regarding their completion time and number of locomotion) if team members engage in fruitful visits to several rooms.

By selecting an equipment in the GUI part number 4, team members can perform the probe action, to check their current state (GUI part number 5), as well as, the affordable operational actions (GUI part number 6) which their outcome will be displayed in GUI part number 8.

The equipment that may be (re)connected to other network devices, will bring out the list box (GUI part number 7) with those devices that the currently loaded network architecture afford connecting with. Each of this actions also freezes the screen since they are not as well instantaneous in real settings.

Table 5.7 presents the times associated to each of the locomotion and operational operations, that were configured to run the experimental trials, grounded on some preliminary pilot experiments, as reported in chapter 6.

Table 5.7: Lag time associated to each of the actions afforded by the Operational Work Environment Simulator

Operation	**Time lag** (in seconds)
Check	0.5
Restart	1
Update	2
Replace	4
Connect	6
Switch room	8

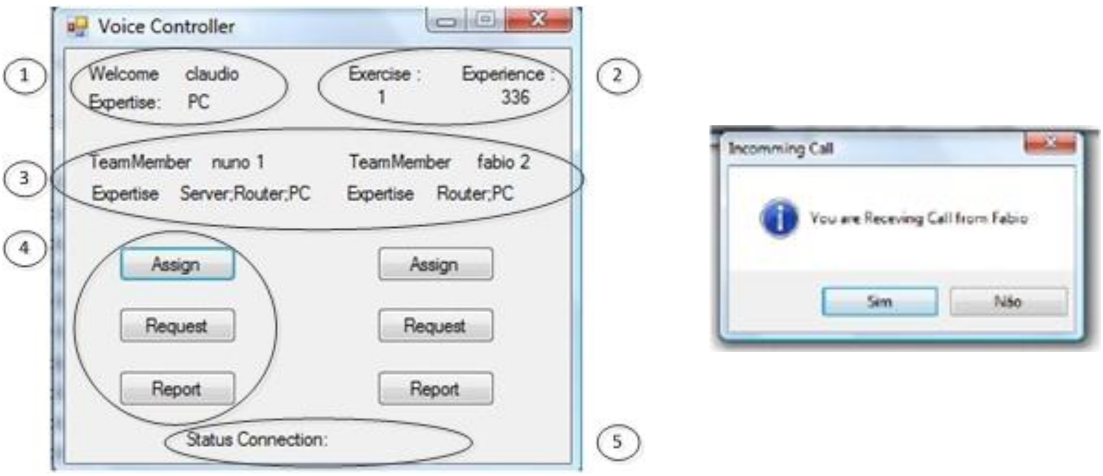
Having presented the OWES main functional features, in the next section it is described the functional features of the Team Communication Support (TCS) subsystems.

5.3.2 Team Communication Support

The implementation of Team Communication Support (TCS) envisioned by the Microworld reference model is described in the next two subsections, respectively addressing the Speech Communications Support (SCS) and the Mobile Collaborative Application Emulator (MCAE).

5.3.2.1 Speech communication support

The Speech Communication Support (SCS) subsystem was implemented by integrating the Team Speak project (TeamSpeak), which as earlier introduced in chapter 4.3.1 constitutes an engine for supporting Voice over Internet Protocol (VoIP) communications. According the consulted Help Desk Teams (HDT), team communications in CIRM contexts draws heavily upon one-to-one phone calls. Figure 5.3a presents the Graphical User Interface (GUI) provided to team members by which they can emulate phone calls communications.



a. b.

Figure 5.3: Speech Communication Support (SCS) subsystem: a. Graphical user interface; b. Notification of an incoming phone call

One draws on the use of the main constituents parts of the SCS GUI, identified by the numbered elliptic highlighting's, to describe its functional features.

The elements on the GUI part number 2, are common to those already explained in the OWES GUI part number 1, in the previous section; they hold the identification of, the number of the experimental trial, and the loaded task scenario on the Microworld environment.

The GUI part number 1, provides the identification of the logged-in team member by displaying its name and expertise. The GUI presented on the Figure 5.3a, expresses that the team was composed by three elements. Therefore the other two elements' identification was displayed on GUI part number 3, also through their name and expertise. The option of referring each team member's expertise was made by the consideration that if team members are assigned different roles, those must be made explicit since they frame the types of operations that they can enact, and thus may root the reasoning behind the communications initiatives.

A speech communication between the team member and the targeted recipient is triggered by pressing one of the buttons on GUI part number 4, immediately bellow the targeted team member. The distinction of the three different buttons to perform a speech communication was made in an attempt to code the communication content. For clarification, if the team member C*laudio*, wants to phone team member *Nuno* with the purpose of assigning a task, he should press the button *Assign*. Conversely, if he intend to report activity, than he should press the button *Report*, or if he is requesting information it should press the *Request* button. This still holds if the targeted team member of the phone call was *Fabio*, by respectively pressing the appropriated button under the *Fabio* name, given the intent of the communication. The receiver of the phone call is prompted with a modal window identifying the caller (Figure 5.3b) and may accept or reject the call.

The GUI part highlighted in Figure 5.3a with number 5, displays the state of the phone call, if it is on-going, meaning the phone call had been established or if it was rejected or the line was occupied.

5.3.2.2 Mobile collaborative application emulator

Besides the specifications of the functional requirements underlying the study of Mobile Collaborative Applications (MCA) role in CIRM, that should be incorporated in the Mobile Collaborative Applications Emulator (MCAE) subsystem, put forward in chapter 4.3.2, and their consideration framed on the Help Desk Teams (HDT) domain; one must notice that the MCAE implmentation follows two non-functional requirements. These are: the display area should be made similar in size to the one that mobile devices present and the interaction modalities should also follow the limitations and more typical user interface controls provide by applications in mobile devices, which posit the refrain from the adoption of more open text inputs toward more stylus (or touch) based interactions (which in the MCA emulator are mimicked by the mouse pointer).

According its specification, the MCAE delivers four main MCA Graphical User Interface (GUI) Screens: 1. Situation Monitoring, which is intended to provide a control panel regarding the overall CIRM context, 2. Assignments Management, which supports issuing task assignments to team members, 3. Information Report, that provides the means for reporting enacted operational activities and 4. Notifications History Log, delivering the issued notifications log. Each of those are briefly reviewed, below, regarding their main functional features.

Figure 5.4, depicts the main GUI parts of the Situation Monitoring Screen.

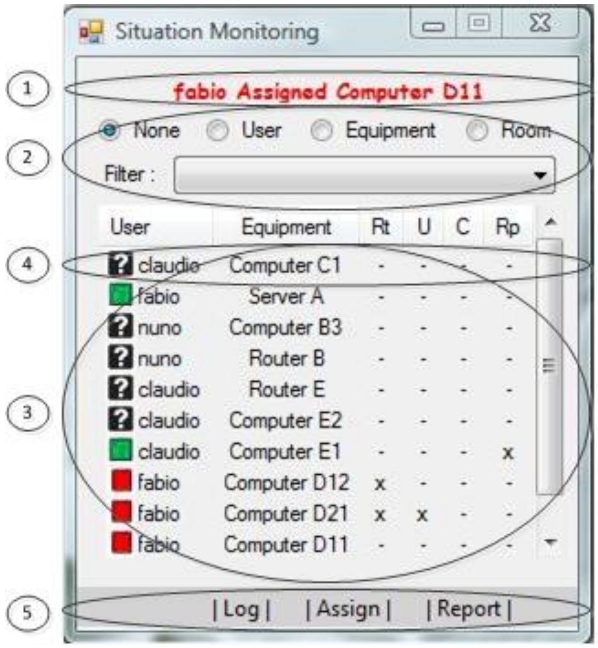


Figure 5.4: Mobile Collaborative Application Emulator (MCAE) Situation Monitoring Screen

The GUI part number 1, holds the notifications area. According the MCA specification, the operational activities enacted by team members (e.g. task assignment, equipment states reports) are broadcast to team level (exception is made to the MCA client of the team member that originated the notification).

The major area of the Situation Monitoring GUI (GUI part number 3) is occupied by the list of the equipment that composes the HDT operational work environment. Each of the elementary elements that compose this list (highlight number 4), holds the following information: the relation between a team member and an equipment devised from the task assignments, four columns that if marked with an *X*, express the operations that already have been reported regarding that equipment (respectively, Rt – Restart, U – Update, C – (Re)Connect and R – Replace), and at the beginning of the line a colourful icon representing the equipment currently known state (green if *operating*, red if *malfunctioning* and *?* if unknown/unreported). This more explicit visual cue draws from studies that reported that people infer missing information, often mistakenly, by assuming that it would be in agreement with their expectations (Banbury, Selcon et al. 1998; Bolstad and Endsley 1998).

Considering that such list (GUI part number 4) may provide a very crowded display, and that information requirements for a particular team member at a given moment should more readily addressed, following the MCA specification, a dynamic filter mechanism was made available (GUI part number 2). It allows to set a number of filter criteria namely, by team members, by equipment type (e.g. see only routers) or by location (list only the equipment that are placed on a given location).

The bottom of the screen (GUI part number 5) holds the navigation bar between MCA screens, the user can directly navigate to each of the provided screens, minimizing MCA logical branches.

The MCAE Assignments Management screen is depicted in Figure 5.5. It hold some parts in common with the previously presented Situation Monitoring screen, namely the notifications area (GUI part number 1) and the navigation bar (GUI part number7), which were already described.

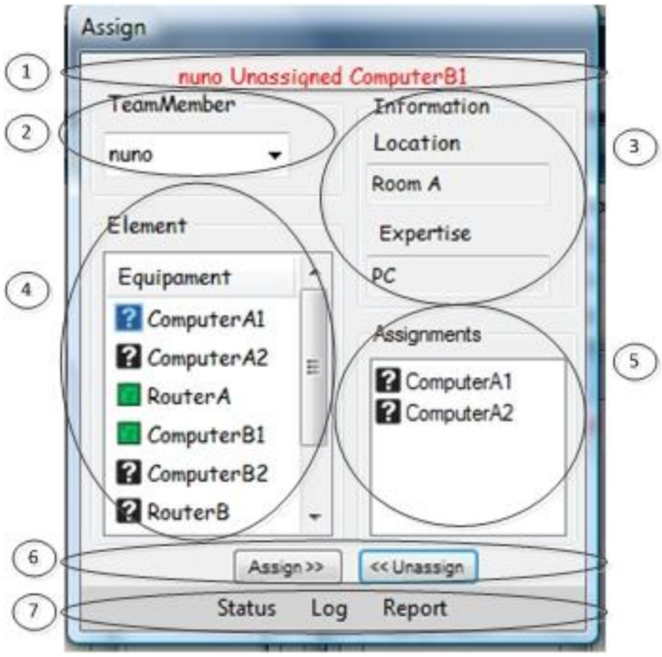


Figure 5.5: Mobile Collaborative Application Emulator (MCAE) Assignments Management Screen

Task assignments are accomplished by selecting the team member that will receive the assignment (GUI part number 2). This selection updates the fields in GUI part number 3, providing additional information for the last reported location of the selected team member and the expertise that frame her/his currently hold role in the team. This information may assist the reasoning for an assignment issuing. As earlier described equipment's description are furnished with a colourful icon, revelling their last reported state. Additionally the already issued assignments for that team member, are presented in the listing that constitute the GUI part number 5. By selecting one of those, an un-assignment can be issued through the respective button placed in GUI part number 6.

Assignments are done by selecting an equipment on the listing that comprise the GUI part number 4 and pressing the assign button also presented in GUI part number 6. If it passes a validation mechanisms it will figure on the assignments list for the selected team member (GUI part number 5). The validation mechanism, addresses the situation that the selected equipment had in the meantime been assigned to other team member without being noticed by the user, since the assignments list only exhibit the assignments for the currently selected team member. In this case, the user will be prompted if she/he want to override the existing assignment, if so the MCAE will trigger an un-assignment notification to the formerly assigned team member along with the notification of the assignment to the selected team member.

Both the validation mechanism and the functional feature of supporting un-assignments, supports two distinct operational requirements. One concerns that when the assignment initiator realizes that it will be better to re-assign a task rooted on the urgency of the current assignment, the current situation evolution (which may motivate a revision of priorities and workload balance) or on the acknowledgement that the alternative team member current location or expertise qualifies her/him better to the current assignment in disregard of the previous one. Secondly, the receiver team member may use the *unassign* button to decline issued assignment, due her/his work context (e.g. excessive workload or her/his current location).

Regarding, the Information Report screen, it may be triggered by two means. One is by selecting an equipment in the Situation Monitoring screen and then press the *Report* button on the respective navigation bar. The other is by pressing the *Report* button, without any selection from the main listing of the situation Monitoring screen. This second functional feature addresses the opportunistic basis of the CIRM activity, where a particular team member may undertake some operational activity on an equipment without a previously issued assignment, and make notice of that to the team level. In either cases an information report notification will be issued to the team (naturally with the exception of the reporting team member). Figure 5.6, presents the Information Report screen, which as the previous ones is described through is main GUI parts.

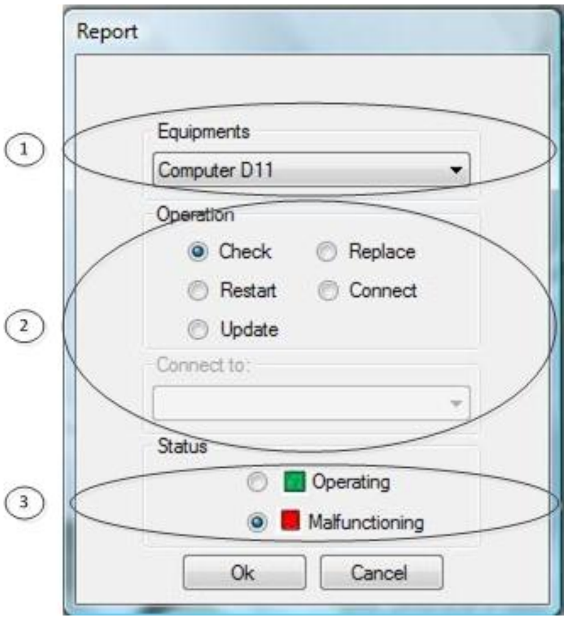


Figure 5.6: Mobile Collaborative Application Emulator (MCAE) Information Report screen

The first GUI part provides a list box with the equipment, which if, the reported screen was triggered from a selection on the Situation Monitor Screen, will already exhibit the respective selection, if not the user will have to select the equipment over which she/he want to report information over, from the list box.

The information reported give account for the enacted operational actions performed over the equipment (e.g. restart or replaced the equipment), which comprise the GUI part number 2 and the respective outcome sate achieved (*operating* or *malfunctioning*), which is supported by the radio buttons hold in GUI part number 3.

As it had been discussed in the MCA specification, considering the high demands of the CIRM process, the issued notifications may lack attendance from team members, due their current focus and/or high work and cognitive load. This had grounded the motivation of delivering the notifications' history log through the *Notifications History Log* screen. The corresponding screenshot is present in Figure 5.7.

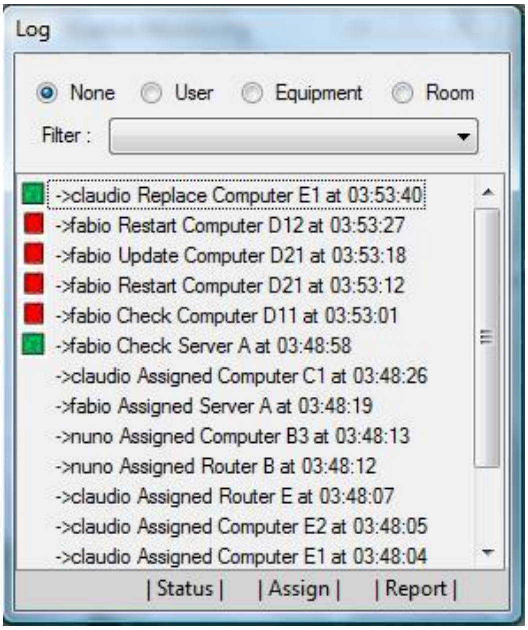


Figure 5.7: Mobile Collaborative Application Emulator (MCAE) Notifications History Log screen

The Notification History Log screen may be accessed by the navigation bars present at the Situation Monitoring and Assignments Management screens. It displays all the issued notifications with their respective timestamp. If a notification was driven from an operational information report, it is further furnished by a colourful icon expressing the outcome of the enacted operational activity, regarding the state of the equipment over which the report was done. Since the log display may become too crowded and the user information needs may be very specific, this screen also provided a filtering mechanism to promote a more readily information retrieval.

This section had overviewed the implementation of the main functional features of the developed Team Communication Support (TCS) subsystems, which comprise the Speech Communication Support (SCS) and the Mobile Collaborative Application Emulator (MCAE) subsystems. The next section describes the implementation of the Situation Awareness Measurement Support (SAMS) subsystem purported by the Microworld reference model.

5.3.3 Situation Awareness Measurement Support

As discussed in the Situation Awareness Measurement Support (SAMS) subsystem specification (in chapter 4.4), the present work had addressed situation awareness measurement by furnishing the Microworld environment with two main mechanisms. One is by logging all the team activity either on the Operational Work Environment Simulator (OWES) as well as, respecting to the Mobile Collaborative Application (MCA) usage, in the context of a task scenario. Such data informs SA analysis by feeding a set of dependent variables addressing more holistic measures of TSA, supported by teams' work processes tracing. These variables along with the task scenarios that were delivered to the teams are described in chapter 6, where the experimental design is deeply depicted.

The other mechanism, to feed an additional set of SA measures, those that addresses the individual and shared dimensions of the SA construct, had relied on the administration of SA questionnaires. The screenshots of exemplary SA questions delivered to team members by the SA Questionnaires Management subsystem implementation, are presented in Figure 5.8.

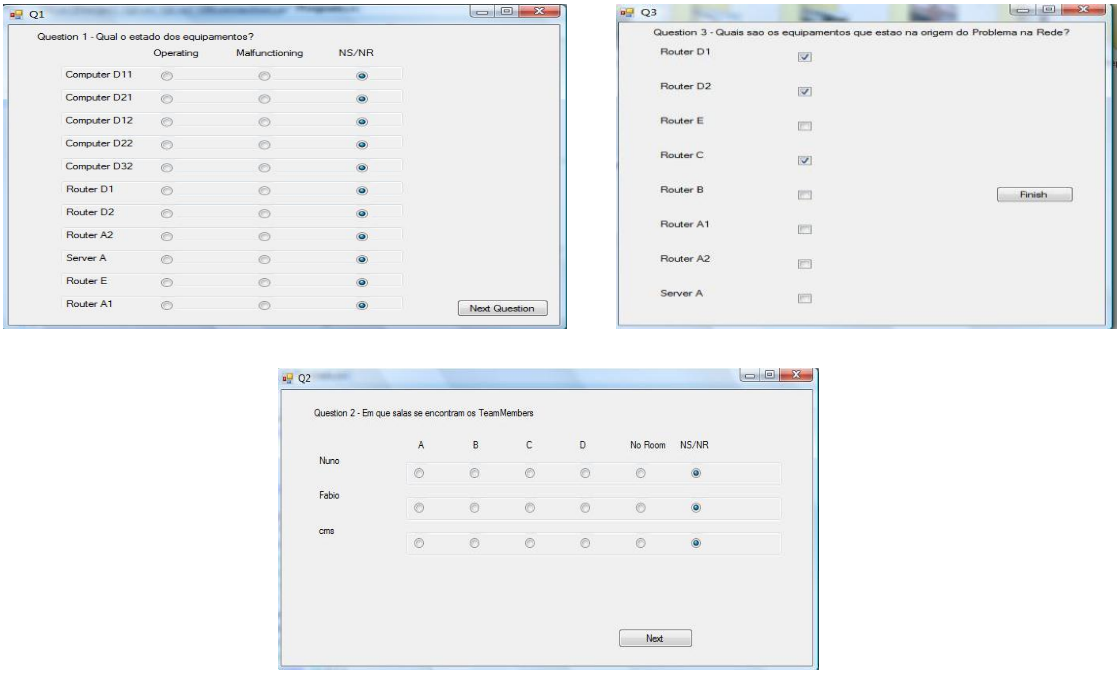


Figure 5.8: Exemplary SA questions delivered to team members through the Situation Awareness Questionnaire Management subsystem

The presented examples on Figure 5.8 support the discussion on how the SA questions are generated by the SA Questionnaires Management subsystem. SA questions answers' options are dynamically generated with the consideration of the currently loaded network architecture on the Operational Work Environment Simulator (OWES), through its provided Application Program Interface (API), the OWES API, which its main requirements were discussed on chapter 4.5. For illustrative purposes one provide a brief overview on how the SA question answers' options are generated.

The SA question on number 2 (Q2), holds the following statement: *"Where are currently located each of the team members?"*, thus the answers options will be, naturally, all the rooms over which the network architecture is distributed. A similar generation process is followed in respect to the SA question number 3 (Q3), which have its statement formulated as: *"What* *is/are the equipment constraining the network connectivity?"*, since it must provide as answers' options, only the existing network devices (routers and servers) of the currently loaded network architecture.

SA question number 1 (Q1), provides an example of a more subtle generation of the SA question answers' options. Its statement is: *"What is the currently hold state of the following equipment?"*. Since it was not reasonable that the answers' option will account for all the equipment that comprises the network architecture, they are restricted to those that are in the same room that the tem member is currently located and those that are in the path of that ones, since it is reasonable to expect that the respondent team member had exchanged information with other team members regarding that ones to frame her/his operational activity. Therefore the generation of such answers' options is grounded on the operational context of the on-going teamwork over the OWES delivered by the OWES API.

This context is of paramount importance to assess the accuracy of the answers, the answers given by each team member are persisted along with, what would be the correct answer for the time when the SA question was administrated.

5.4 Discussion

This chapter had presented an implementation of a Microworld environment by framing the specifications provided on the chapter 4, on the selected target application domain, the Help Desk Teams (HDT) addressing disruptive events on organizations' network infrastructure. The relevance of the research question on the target domain was discussed, as well as, the domain representativeness regarding the present research aims.

The adopted domain immersion methodology, proved to constitute a valuable guidance for the development of the Microworld instance. More specifically, the first stage of the domain immersion strategy, which consisted on conducting a set of interviews to two HDT from two medium size organizations, had proved crucial for establishing a common ground between practitioners and the researcher. Also the scores of the administrated questionnaires, in the second stage of the domain immersion strategy, had furnish researcher's understanding of the salient features that enrich the instantiation of the mobile collaborative applications' (MCA) specifications in a concrete implementation for HDT work support. The last stage of the domain immersion methodology, the scenario discussion workshops, was very fruitful since it had deeply unveil the operational level of HDT in CIRM contexts. The operational level of CIRM may be thought to be distributed in two main dimensions: spatially, regarding the various locations on which the team operational activity is enacted, given that the equipment that comprise the networks' infrastructure is spread through different places; and cognitively, since the knowledge, experience and skills required for addressing critical incidents (CI) rely distributed at team level. While the first dimension posit that Microworld based experimental trials constitute a safe and affordable mean for the study of promising MCA functional features, the second dimension points for the motivation of studying the underlying practices regarding TSA development adopted by HDT operating in Critical Incident Response Management (CIRM) contexts, and the consideration of the role of MCA in such endeavour.

The implementation of each of the subsystems that compose the Microworld reference model, presented on chapter 4, had been presented and discussed in terms of their functional features.

The herein presented implementation of the Microworld environment instance, had supported the set of experimental trials described in the next chapter.

6 Experiments

Toward the development of a keen understanding regarding the role of Mobile Collaborative Applications (MCA) on assisting teamwork in Critical Incident Response Management (CIRM) a set of experimental were conducted engaging a total of 23 participating teams distributed through two rounds of experimental trials. The experimental trials were supported on the Microworld environment implementation presented on the previous chapter 5. The experimental protocol and the overall experiments apparatus that were devised and bounded the experimental trials are initially introduced. The achieved results are presented and discussed in the light of the derived experiment's dependent variables.

6.1 Experimental Design

It was adopted a repeated measurements design (Howell 2009) and each participating team was submitted to two experimental settings.

Considering that the ultimate goal of the experiments was to unveil the role of Mobile Collaborative Applications (MCA) in assisting teamwork, their availability to the teams performing the provided task scenarios underlying each experimental trial, constituted the independent variable of the experiments. Therefore, each team perform two trials exposed to two distinct experimental settings. Accordingly, the distinction between the two experimental settings lays on the availability or not of the Mobile Collaborative Application Emulator (MCAE) in the Microworld environment, which are hereafter referred, for short, as the "w/" (with MCAE available) and "w/o" (without MCAE available) setting, this latter one affords the collection of the experimental data that constitutes the control data to ground the comparisons between experimental settings.

Under both experimental settings, speech communications between pairs of team members were allowed through the Microworld's Speech Communications Support (SCS) subsystem, that as discussed in chapter 5.3.2.1, was implemented over a Voice over Internet Protocol (VoIP) engine. Thus, the emulation of the typical phone communications that Help Desk Teams (HDT) perform to coordinate their operational activity, recurred to the use of headsets, since although participants were located far apart in three separated places of the laboratory while operating the Microworld environment, the use of the headsets eliminates the possibility of overhearing speech communications between other team members, that were not intended to be delivered to them. This constitute a fundamental experimental control for two interwoven reasons: First, in real settings teams operate distributed through several places, and thus only the communications that are directed to each own should be provided; and secondly, a very important dimension of the analysis of the teamwork, in the present research, is Team Situation Awareness (TSA) development and thus if overhearing speech communications between other team members was not controlled this will lead to results bias on the devised dependent variables addressing TSA measurement.

Each experimental setting delivered to teams a network infrastructure and a task scenario, through the Microworld's Operational Work Environment Simulator (OWES) subsystem. Task scenarios posit network connectivity problems in some of its segments. The team task was to diagnose and re-establish the overall network connectivity.

Participants engaged in the experimental trials consisted of final year students of undergraduate courses in Informatics. Extra course credits and a prize money were offered to the best performers to encourage deeper engagement. The participants were informed that their performance was evaluated by their minimization of three main factors: the time to accomplish the task, the number of operations enacted and the number of dislocations that they perform over the distributed (virtual) places that were required to visit.

In order to be eligible for participation, it was mandatory that the students had completed the course on computers' networks, so that they have a deeper understanding regarding the nature of the problems underlying network's infrastructures connectivity management.

Participants apply for participation in groups of three, that given the HDT consulted (chapter 5), was considered a reasonable team size for the first level of the CIRM operational response on small to medium size organizations.

Each participant was assigned a team member number that had an expertise profile associated. Team member number one (TM1) could operate any of the existing network equipment (servers, routers and PC), team member number two (TM2) could only operate network routers and personal computers and team member number three (TM3) could only operate over network's personal computers.

Before each experimental session the teams received a manual (provided in annex B of this thesis) describing the experiment's goals, the type of tasks they will be prompted to accomplish, the team composition and leadership arrangement that bound their affordable operational procedures, and the Microworld functional features.

At the beginning of the experimental session a briefing was conducted to clarify any doubts regarding the experiment or software usage and a consent form (which was attached to the participants manual and is also available in annex B) was signed by each participant, stating her/his commitment to the experiment and authorizing data collection. Figure 6.1 shows two photographs taken during a briefing session.

A registration number was assigned to each selected team for participation and teams were divided by their registration number for the purpose of defining the order that each team was exposed to the two experimental settings.

The odd teams started with the "w/o" experimental setting. In this condition, they perform two trials, the first for training purposes, to get familiar with the setting, and the second for effective experimental data collection. After that, the odd teams were subject to the "w/"

experimental setting and again performed two trials, the first for practice and the second for collecting the data associated with the “w/” experimental setting.

Figure 6.1: Photographs of a briefing session

Even teams performed in the reverse order. They were first exposed to the “w/” setting, and then to the “w/o” setting, and have also performed two trials in each setting, with the first being for practice purposes and the second for data collection. This experimental design, which also purport that, in each of the four trials performed by each team, the delivered task scenario was different, was intended to avoid experiments’ results bias. Table 6.1 depicts the experiments’ controlled conditions.

A further account of two of the presented experiment’s controlled conditions worth to be made. Concerning the team leadership arrangement, which purports that teams will perform under a flat hierarchy mode, was adopted informed on the discussion presented in chapter 4.1.1, that puts forward that in Critical Incidents Response Management (CIRM) contexts, at the operational level of teamwork, teams often departs for pre-established command and control structures, given the contingencies posited by a particular Critical Incident (CI) instance, toward improvised behaviours and that the empowerment of team members regarding their authorship and responsibility for the enacted operational actions is more likely to promotes flexibility and resilience regarding the emergent demands.

As it had been put forward in chapter 4.4, SA questionnaires constitute one of the means to collect SA measures to inform the devised experimental dependent variables that support TSA analysis. Thus, regarding their administration mode, the current experimental design had relied on the adoption of freeze probes. Considering that on-line/real-time SA questionnaires will direct team members attention to specific matters of the overall operational work environment, and that post trial SA questionnaires are prone to rationalizations and unveiling SA level achieved at the end of the experiments trials, freeze probes appears to better avoid SA questionnaires results bias. This issue was extensively discussed on chapter 3.1.2, and although, one acknowledge, that freeze probes will present an interruption of the on-going task, such was addressed by considering that the required effort and time to take the questionnaires was minimized through the provided means for answering them and thus hold a reasonable trade-off between the avoidance of results bias and an acceptable intrusion level.

Table 6.1: Experiments controlled conditions

Controlled Conditions	Description
Eligible participants	Final year students of an undergraduate Informatics course, which have already successfully concluded the computers' networks course.
Team size	Teams were formed with three team members.
Team composition	Teams hold three differentiated roles distinguished through team members expertise's that frame their affordable operational activities over the equipment that comprise the network's infrastructure.
Teams leadership arrangement	Teams operate under a flat hierarchy mode, meaning that each team member is empowered to enact her/his operational activity, only bounded by her/his currently hold expertise profile.
Task scenarios	Different task scenarios are provided for practice and data collection trials in each of the experiments' settings.
Experiment's setting	Teams perform under two settings, given that one holds the availability of the MCAE. Teams with odd and even registration's number are subject to each setting in reverse order.
SA questionnaires administration mode	SA questionnaires are administrated on a freeze probe basis.

Two rounds of experiments were conducted. The first round had comprise 12 experimental sessions engaging respectively 12 teams totalizing 36 participants. On each session the respective participant team was prompted with one network infrastructure delivered through the Microworld environment and faced, for each of the four experimental trials, one of the four task scenarios developed for that network infrastructure, as presented on Table 6.2.

Table 6.2: Experimental trials configuration for the first round of experiments

Experiment Setting / Network Architecture	w/o		w/	
	Practice	**Data Collection**	**Practice**	**Data Collection**
Network #1	Task Scenario #1	Task Scenario #3	Task Scenario #2	Task Scenario #4

On the first round of experiments, the Microworld environment did not delivered all the network's equipment affordable operational actions foreseen by the Operational Work Environment Simulator (OWES) implementation described on chapter 5.3. In Table 6.3, those are reviewed and their availability it each round of experiments is expressed.

Table 6.3: Description of the network equipment affordable actions and their availability on each round of experiments

Network Equipment	Affordable Actions	Description	Availability	
			First Round	Second Round
Servers Routers Personal Computers	Check (Probe Action)	Check equipment state	Available	Available
	Restart (Operational Action)	Restart equipment	Available	Available
	Update (Operational Action)	Update equipment components	Available	Available
	Replace (Operational Action)	Replace equipment components	Not Available	Available
	(Re)Connect (Operational Action)	Re(Connect) equipment to other network devices	Not Available	Available

The constrains regarding the available network's equipment affordable actions, as well, as the fact that all task scenarios in the first round of experiments were toward a single network architecture, were made to minimize the learning effort of the teams participating in the experiments, regarding the multitude of aspects that underlies the experiments endeavour.

In the light of the results achieved in the first round of experiments, discussed in section 6.2.1.1 of the present chapter, which had revealed that more elaborated task scenarios would enhance the experiments outcomes analysis and by the acknowledge that such level of complexity would in did be acceptable by teams elected by the adopted selection criteria, a second round of experiments was conducted. In this round one had introduced two additional network architectures and included all the available network's equipment affordable actions in the OWES. The configuration of the experimental trials conducted in the second round of experiments is presented on Table 6.4.

Table 6.4: Experimental trials configuration for the second round of experiments

Trial's Goal	Network #	Task Scenario #	Experimental Setting	Trial Configuration #
Practice	1	3	w/	1
	1	3	w/o	2
	1	4	w/	3
	1	4	w/o	4
Data **collection**	2	5	w/	5
	2	5	w/o	6
	3	6	w/	7
	3	6	w/o	8

In order to reduce the results bias that may arise from a learning effect, the teams that participated on the first round of experiments could not participate on the second round. The applying and selection process of teams for participating in the second round of experiments followed the criteria already described. The second round of experiments had engaged, 11 additional teams totalizing 33 participants, that perform the experimental trials as previously delineated. Therefore, the four experimental trials of each team (two for practice and two for data collection) were randomly selected from one of the trial configurations presented in Table 6.4, although keeping the concern of achieving similar number of trials in both the “w/” and ”w/o” settings for supporting the comparisons between both settings . For further clarification, again teams holding an odd registration number were first exposed to the “w/o” setting and then to the “w/” setting, while the teams with an even registration number were exposed to each experimental setting in the opposite order.

The description of the provided network’s architectures, as well, as their corresponding task scenarios is done on the sections 6.2.1 and 6.2.2 of the present chapter, respectively addressing those deployed on the first and second round.

The remaining of this section introduces the derived experiment’s dependent variables and discusses the reasoning that grounded their definition. Given that, in the present work, Situation Awareness (SA) development constitutes the main dimension of analysis of the role of MCA in teamwork under CIRM contexts, a set of SA related dependent variables were defined. Those are presented on Table 6.5.

The herein devised SA related dependent variables, addresses the three main dimensions of the construct: Individual, Shared and Holistic. The reasoning underlying their definition and the adopted measurement techniques: SA questionnaires and work processes tracing, are further discussed below.

Regarding individual SA, three dependent variables were devised. The first dependent variable, Individual Questionnaire’s Scores (IQS), addresses the fact that, SA inherently holds a cognitive dimension and thus only to some extent the level of SA possessed by a team member could be inferred from individuals operational activity. As so, the direct elicitation of the level of SA hold by each team member is grounded on the administration of freeze probe SA questionnaires and SA is quantified by the achieved score. More specifically, the IQS is defined by the ratio between the number of correct answers and the total number of questions present on the SA questionnaire. One should notice, at this point, that as discussed in both, the specification of the SA Questionnaires Management Microworld’s subsystem on chapter 4.4 and later on its implementation described in chapter 5.3.3, SA questionnaires are dynamically generated given the team member’s operational context, and thus a ratio seems to be a more proper questionnaire score than the absolute value of the number of correct answers.

By acknowledging that the understanding of the level of SA hold by a team member could also benefit for the account of her/his operational activity, the two additional dependent variables regarding individual SA were drawn from her/his operational enacted work tracing, which is supported by the Microworld ability to log the enacted operational actions framed on the overall task scenario context. Given that, the Individual Diagnosis Efficiency (IDE) dependent variable accounts for the ratio between the redundant probe actions that a team member enact toward the diagnosis of network’s equipment regarding the total amount of enacted probe actions. Network’s equipment affordable probe actions have been considered the *check* action by which a team member can diagnose the equipment current state (*operational* vs *malfunctioning*). A redundant *check*, is thought as one enacted by the team member over a network's equipment that was already been *checked* by other team member, or even if it had not been *checked* by other team member none operational action (e.g. restart, update, etc.) was enacted by any element in the team over the equipment or over one (or more) equipment in its path before the performed *check* action, which make it also redundant.

Table 6.5: Situation Awareness related dependent variables

Scope	Data Collection	Dependent Variable	Dependent Variable Definition
Individual SA	SA questionnaires administrated through freeze probes	Individual SA Questionnaire Score [IQS]	$\frac{\text{\# of correct answers}}{\# \ of \ questions}$
	Work process tracing through OWES logs	Individual Diagnosis Efficiency [IDE]	$1 - \frac{\# \ of \ redundant \ equipment \ checks}{\# \ of \ equipment \ checks \ enacted}$
		Individual Operational Efficiency [IOE]	$1 - \frac{\# \ of \ redundant \ operational \ actions}{\# \ of \ operational \ actions \ enacted}$
Shared SA	SA questionnaires administrated through freeze probes	Shared SA Questionnaire Score [SQS]	$\frac{\text{\# of correct common answers between two TM}}{\# \ of \ questions \ common \ to \ two \ TM} + \frac{2}{3}\frac{\text{\# of correct common answers between two TM}}{\# \ of \ questions \ common \ to \ three \ TM} + \frac{\text{\# of correct common answers between three TM}}{\# \ of \ questions \ common \ to \ three \ TM}$
Holistic SA	Work process tracing through OWES logs	Team Diagnosis Efficiency [TDE]	$1 - \frac{\sum_{i=1}^{3} \# \ of \ redundant \ equipment \ checks \ of \ TMi}{\# \ of \ collective \ equipment \ checks}$
		Team Operational Efficiency [TOE]	$1 - \frac{\sum_{i=1}^{3} \# \ of \ redundant \ operational \ actions \ of \ TMi}{\# \ of \ collective \ operational \ actions}$

The reasoning behind this definition of the IDE, is that if the current state of a network's equipment had been unveiled by any of the other team members and/or none operational actions were enacted since over it (or on those in its path) by the team, than the individual diagnosis endeavour lack SA and thus it will reduce the IDE value.

The definition of the Individual Operational Efficiency (IOE), follows the same reasoning, by acknowledging that any operational action (e.g. restart, update, etc.) enacted by the team member over a given network's equipment toward is recovery, will not be redundant only if none operational action was already enacted over that equipment or one its path that had already lead it to the operational state. For instance, if a team member perform a *(Re)Connect* action on a personal computer (PC), that earlier exhibit a lack of network connectivity, by linking it to another router, but other team member had already established its connectivity by replacing the router where it was originally connected, and that constitute the source of the PC lack of connectivity, then any operational action toward the PC will be redundant and points for a lack of SA and thus reducing the IOE value.

Although the IDE and IOE scores hold to some extent an holistic flavour, since they account for the overall underlying team activity, they are formulated to yield an individual SA score grounded on each team member operational activity that is representative for her/his current understanding of the state of the network's equipment that are related to her/his current task.

Therefore the discussed reasoning behind the IDE and IOE, was naturally extended to yield a team level TSA measure, through the definition of the Team Diagnosis Efficiency (TDE) and Team Operational Efficiency (TOE) dependent variables. TDE and TOE are derived from the same logic of the their individual counterparts, and are formulated by accounting respectively the diagnosis and operational actions of all team members.

Also at team level, it was defined the dependent variable, Shared SA Questionnaire Score (SQS). As it had been pointed out by the discussion of the related literature review, heterogeneous teams (holding team members with specific expertise) performing in complex task scenarios, will hardly achieve or require a common completely overlapped model of the situation. Therefore the shared dimension of the SA construct must be embraced by the consideration that the extent that SA overlap among team members is derived from the coupling of work imposed by the demands of the task.

Since, as already pointed, SA questionnaires are dynamically generated and delivered to each team member according her/his current operational context, they hold some questions that only address each team member SA, but may also hold some questions in common to some other (or all) team members. Accordingly, the SQS scores are computed with the consideration of such common questions following the respective formula presented on Table 6.5.

In order to support a finer grain understanding of the impact of MCA on teamwork an additional set of dependent variables was defined, those are presented on Table 6.6. This dependent variables were divided into two main categories: Team Speech Communications and Mobile Collaborative Application (MCA) usage.

Team speech communications analysis pursue the development of a deeper insight regarding how teams manage operational information and coordinate their activity. The analysis of the impact of the availability of MCA regarding this matters was grounded on the following dependent variables: the *Number of enacted speech communications* (NSC), which supports the assessment of the extent that under the "*w/*" experimental setting teams still rely on speech communications; and the type of speech communications that are been carried out, more specifically, the fraction of speech communications that concerns operational information management (SC_IM) (e.g. reports or requests on the state or enacted operational actions on a given network equipment) and the fraction of speech communications that concerns team management (SC_TM) (e.g. task assignments).

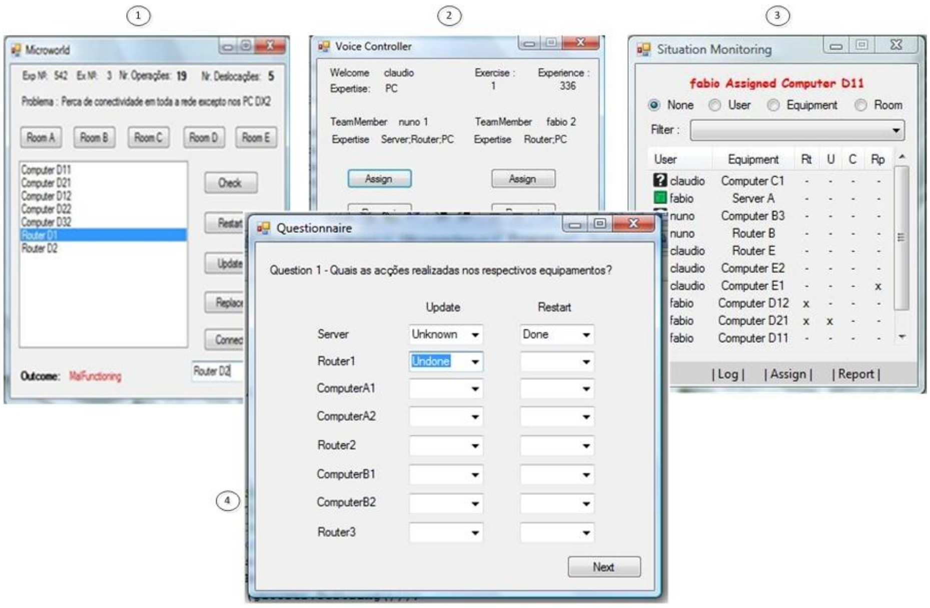
Table 6.6: Devised Team Speech Communication and Mobile Collaborative Application Usage related dependent variables

Category	Dependent Variable	Data collection	Definition
Team Speech Communications	Number of speech communications [NSC]	Speech Communication Support (SCS) Log + Observers' ratings	Total number of speech communications
	Fraction of speech communications for information management [SC_IM]		$\frac{\#\ info\ manag\ communications}{NSC}$
	Fraction of speech communications for team management [SC_TM]		$\frac{\#\ team\ manag\ communications}{NSC}$
MCA Usage	Report Ratio [RR]	OWES + MCA log	$\frac{\#\ of\ operations\ reported\ through\ MCA}{\#\ of\ operations\ logged\ on\ OWES}$
	Situation Monitoring screen usage percentage [SMS_U]	MCA log	$\frac{\#\text{ of visits to Situation Monitoring screen}}{Total\ \#\ of\ visits\ to\ all\ MCA's\ screens}$
	Report screen usage percentage [RS_U]		$\frac{\#\text{ of visits to Report screen}}{Total\ \#\ of\ visits\ to\ all\ MCA's\ screens}$
	Assignments Management screen usage percentage [AS_U]		$\frac{\#\text{ of visits to Assignments Manag screen}}{Total\ \#\ of\ visits\ to\ all\ MCA's\ screens}$
	Notifications History Log screen usage [NHL_U]		$\frac{\#\text{ of visits to Notifications History Log screen}}{Total\ \#\ of\ visits\ to\ all\ MCA's\ screens}$

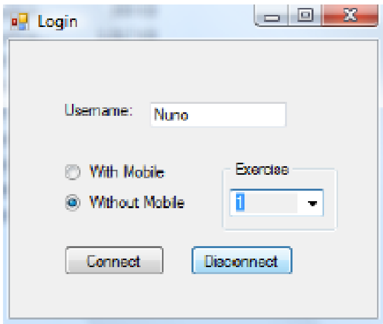
The dependent variables depicted on Table 6.6, regarding the MCA usage category, more readily addresses the development of an informed understanding regarding the real usage of MCA's functional features. This is accomplished by putting forward the following dependent variables: *Report Ratio* (RR), *Situation Monitoring screen usage percentage* (SMS_U), *Report screen usage percentage* (RS_U), *Assignments screen usage percentage* (AS_U) and *Notifications History Log screen usage percentage* (NHL_U).

While the first (RR), expresses the extent that the enacted operational actions are in did reported through the MCA, the four latter ones, more specifically expose what are the functional features that had received more acceptance/adoption by the end users (team members).

For illustrative purposes, one presents in Figure 6.2a the composition of the desktop that is displayed to team members' after logging-in in the Microworld environment. It comprises: 1) the Operational Work Environment Simulator (OWES); 2) the Speech Communications Support (SCS); 3) the Mobile Collaborative Application Emulator (MCAE) and 4) the SA Questionnaires.



a.



b.

c.

Figure 6.2: Microworld environment deployed in the experiments: a. Graphical User Interface of all the Microworld subsystems; b. Microworld login screen; c. An experiment's participant operating the Microworld in the context of an experimental trial

One should make notice, that only in the *"w/"* experimental setting they were all available, since that in the "*w/o"* experimental setting the MCAE is inhibited. As it can be perceived from the login screenshot (Figure 6.2b), participants, following researcher instructions, have to explicitly indicate if the MCAE is to be loaded or not, as well as the task scenario number (through the exercise number list box). Regarding the SA questionnaires, given the adoption of the freeze probes technique, they are only displayed (in both experimental settings) when the triggering condition was met freezing and overlapping all other subsystems Graphical User Interfaces (GUI). The triggering condition was defined at every ten operational actions that the team collective enact over the OWES. This number resulted from preliminary pilot experiments that were conducted for testing the software and users reactions (although none of the teams participating in the experimental trials were used in pilot runs, to prevent results bias derived from a learning effect). In in Figure 6.2c it is exhibited a photograph of a participant performing an experimental trial.

In the following section one presents and discuss the results for each of the dependent variables for the two conducted rounds of experiments.

6.2 Experimental Trials

6.2.1 First Round of Experiments

As introduced the previous section the herein reported first round of experiments engaged a total of 36 participants forming 12 teams.

The network architecture loaded on the Microworld Operational Work Environment Simulator (OWES) for this first round of experiments is presented in Figure 6.3. This network architecture entails that network's equipment are distributed through three different rooms named: room A, room B and room C. Room A and room B mimic what could be two separated offices each one equipped with a router that provide the network connectivity to the two personal computers that are present in each room (room A and room B). Room C represents a server room, which contains the network's server and a router that will distribute the network connectivity to rooms A and B.

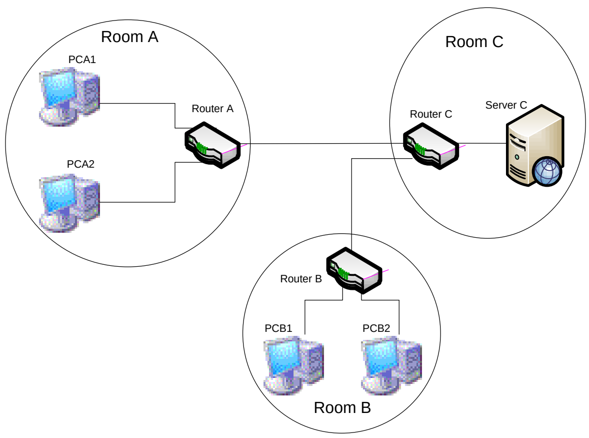


Figure 6.3: Network architecture number one loaded on the Microworld Operational Work Environment Simulator (OWES)

According the experimental design presented in the previous section, in the first round of experiments the data collection experimental trials for the *"w/"* experimental setting relied on the use of the task scenario number 4, while the data collection experimental trial for the *"w/o"* setting was drawn upon the use of the task scenario number 3.

Following the task scenario characterization described in chapter 4.5, task scenarios comprehend a description of the Critical Incident (CI) that will trigger the team's CIRM, a set of initial states for each equipment that comprises the network's infrastructure and a collection of events that lead the network's dynamic purported by a CI occurrence.

While the description of the task scenario number 3 posit to the teams that with the exception of the server, all network's equipment lack network connectivity, the description of the task scenario number 4 states that none of the network's equipment currently have network connectivity. This had rooted the initial conditions for each scenario, in terms of the initial state of all the equipment that comprise the network architecture number one as detailed in Table 6.7.

Table 6.7: Task scenarios number three and four initial conditions

Equipment ID	Equipment Description	Equipment label	Initial states Scenario #3	Initial states Scenario #4
1	Personal Computer A1 (room A)	PCA1	*Malfunctioning*	*Malfunctioning*
2	Personal Computer A2 (room A)	PCA2	*Malfunctioning*	*Malfunctioning*
3	Router A (room A)	RA	*Malfunctioning*	*Malfunctioning*
4	Personal Computer B1 (room B)	PCB1	*Malfunctioning*	*Malfunctioning*
5	Personal Computer B2 (room B)	PCB2	*Malfunctioning*	*Malfunctioning*
6	Router B (room B)	RB	*Malfunctioning*	*Malfunctioning*
7	Router C (room C)	RC	*Malfunctioning*	*Malfunctioning*
8	Server C (room C)	SC	Operating	*Malfunctioning*

The collection of events, for each task scenario, that hold the extraneous manipulation of the networks dynamic is presented on Table 6.8, accounting for the attributes that follows the specification for the scenarios events characterization.

For further clarification, one offer an overview of the underlying issues of each task scenario. Regarding task scenario number 3, the network connectivity problems were derived due the router on room C (RC) holding the *malfunctioning* state, which would only be recovered to the *operating* state if an *Update* action was performed over it. If enacted, such action will naturally re-establish the connectivity in all room A equipment, but connectivity on room B will still be constrained. The connectivity in the equipment of room B, will only be restored if a *Restart* action is performed on its router (RB).

On task scenario number 4, the lack of connectivity on the network was rooted on the server (in room C) initial *malfunctioning* state. If an *Update* is performed over it, the server will recover to the *operating* state and re-establish connectivity in all room B equipment, but room A will still lack connectivity. Connectivity in the equipment placed on room A may only be achieved if an *Update* action was enacted on its router (RA).

Table 6.8: Task scenarios number three and four collection of events that constrain the regular network dynamics

	Event ID	Targeted Equipment (ID)	Pre-conditions	Post-conditions
Task Scenario #3 ("w/" setting)	1	RC (7)	An *Update* action is carried out on RC.	Update the states of equipment: PCA1, PCA2, RA and RC to the *Operating* state
	2	RB (6)	RC holds the *Operating* state and A *Restart* action is carried out on RB.	Update the states of equipment: PCB1, PCB2 and RB to the *Operating* state
Task Scenario # 4 ("w/o" setting)	1	Server (16)	An *Update* action is carried out on SC.	Updates the states of equipment: PCB1, PCB2, RB, RC and SC to the *Operating* state
	2	Router A (11)	RC holds the *Operating* state and An *Update* action is carried out on RA.	Updates the states of equipment: PCA1, PCA2 and RA to the *Operating* state

This section concludes by presenting the questions delivered through the SA questionnaires, depicted in Table 6.9. Given the reduced dimension of the network architecture number one, in this first round of experiments, the SA questionnaires were generated equally to all team members and SA questions number 1 and 3 account for the totality of network's equipment. Figure 6.4 presents the screenshots of the administrated SA questionnaires in the first round of experiments.

Table 6.9: Situation Awareness questions administrated through the freeze probe questionnaires in the first round of experiments

Question ID	Question
Q1	What were the operations performed over each network equipment ?
Q2	To which team member is assigned each network equipment ?
Q3	What is the current state of each network equipment ?

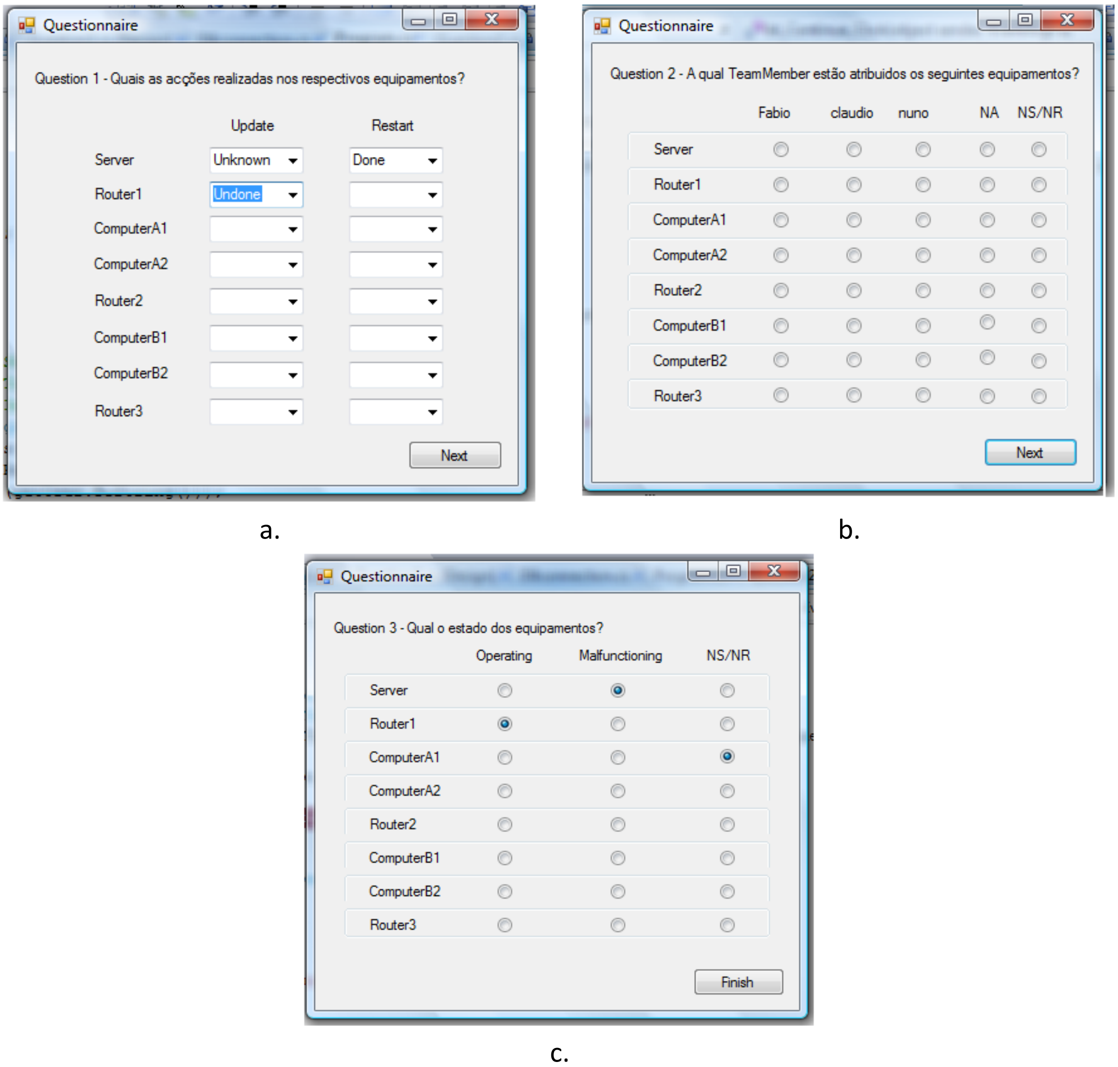


Figure 6.4: Screenshot of the freeze probe situation awareness questionnaires for the first round of experiments : a. SA question number one; b. SA question number two and c. SA question number three.

The following sub-section presents and discusses the results from the first round of experiments, which the apparatus was herein described.

6.2.1.1 Results and Discussion

The discussion of the results achieved in the first round of experimental trials regarding the derived set of experiment’s dependent variables is organized in three main parts.

One start with the presentation and discussion of the SA related dependent variables. The presentation of their results is further divided into those regarding the SA freeze probe questionnaires, namely, the Individual SA Questionnaires Scores (IQS) and the Shared SA Questionnaires Scores (SQS), and those addressing work processes tracing, namely, Individual Diagnosis Efficiency (IDE), Individual Operational Efficiency (IOE), Team Diagnosis Efficiency (TDE) and Team Operational Efficiency (TOE).

The discussion of experiments’ results that moves toward those dependent variables on the realm of team speech communications category, which comprehends the Number of Speech

Communications (NCS), the Fraction of speech communications for information management (SC_IM) and the Fraction of speech communications for team management (SC_TM).

The third, and final part of the results presentation is rooted on the discussion of the Mobile Collaborative Application's usage, which entail the: Report ratio (RR), Situation Monitoring screen usage (SMS_U), Report screen usage (RS_U), Assignments screen usage (AS_U) and the Notifications History Log screen usage (NHL_U) dependent variables.

Given that the data on the dependent variables does not follow a normal distribution, one had relied on the non-parametric, distribution free Wicoxon matched-pairs signed rank test (Tamhane and Dunlop 2000) for assessing the statistical significance regarding the comparisons of the results between the *"w/o"* and *"w/"* experimental settings.

In both experimental settings teams had successfully accomplished the task of recovering the connectivity in all network's equipment. The averages, across the 12 teams, of task completion time and the total number of actions enacted over the network are depicted on Table 6.10, for each experimental settings.

Table 6.10: Average (and standard deviation) of task completion times and number of operations of teams in both experimental settings.

	Experimental Setting		**Wilcoxon p-value**
	w/o	***w/***	
Average completion time	7,94 (3.52)	6,88 (1.57)	0,34
Average # of enacted operations	24,67 (10.97)	25,17 (4.99)	-

Teams under the "w/" experimental setting took in average one minute less to perform the trial but enact just about the same amount of actions over the network. The use of MCA caused little impact on the number of operationally enacted actions, while regarding the task completion time it will be required further research to make a more grounded statement on the exhibited trend, since the difference between the two experimental settings, does not hold statistical significance, given that the Wilcoxon p-value is above the considered statistical significance threshold (*0,05*).

Regarding the results for the IQS for each of the three questions that comprised the SA questionnaires on should notice that, given the triggering condition of the freeze probes was framed on the number of operational activities that the team collective enact, some teams had faced two iterations of freeze probes, while others only were prompted once with the SA questionnaires. Table 6.11 gives further account of this issue.

Table 6.11: Number of SA questionnaires freeze probe iterations achieved by the each team

	Experimental Setting			
	w/o		*w/*	
Freeze Probes	**Teams ID**	**# of Teams**	**Teams ID**	**# of Teams**
Iteration #1	1,2,3,4,5,6,7,8,10,11,12	12	2,3,5,6,7,8	6
iteration #2	1,2,3,4,5,6,7,8,10,11,12	12	1,2,3,4,5,6,8,10,11,12	11

Table 6.12 presents the IQS average of all team members regarding the three SA questionnaire's questions for each freeze probe iteration. The IQS holds similar values across experimental settings, with the exception of those related with the SA question number 3 (awareness of the current state of network's equipment), over which higher scores were achieved in the second iteration of the freeze probes, nevertheless that difference does not hold statistical significance since. Although this may be justified due the more reduced number of data (N=15) that feed that analysis. Thus a more valid result on this matter will require further research.

Table 6.12: Average (and standard deviation) of Individual Questionnaires Scores (IQS) for each iteration on each SA question

	SA Questions					
Experimental Setting	**Q1**		**Q2**		**Q3**	
	it#1 (N=36)	**it#2** (N=15)	**it#1** (N=36)	**it#2** (N=15)	**it#1** (N=36)	**it#2** (N=15)
w/o	0,33 (0.25)	0,40 (0.25)	0,64 (0.22)	0,79 (0.15)	0,40 (0.23)	0,48 (0.23)
w/	0,32 (0.28)	0,44 (0.23)	0,63 (0.25)	0,78 (0.21)	0,49 (0.26)	0,64 (0.3)
Wilcoxon *p-value*	-	-	-	-	-	0.1

The introduction of the MCA appears to have little impact on the individual SA hold by team members. The results point that in both experimental settings, team members exhibit a low level of awareness regarding the operational actions that were enacted by the team, addressed by the SA question number 1. Conversely, higher levels of individual SA are achieved regarding the matters addressed in SA question number 2, which concerns the awareness of the distribution of task assignment among the team members. Given the similarity of the IQS across the two experimental settings conducting a statistical significance analysis of their differences is pointless. Nevertheless, in either experimental settings some progression on the level of SA as the team task unfolds is noticed from the evolution of the IQS through the freeze probe iterations. A deeper analysis on the evolution of IQS is depicted on Table 6.13, where one have introduced the SA Improvement Ratio (SA_IR) between freeze probe iterations to support the analysis. This measure was defined as the ratio between the IQS achieved in the second iteration of the freeze probes regarding the IQS exhibited in the first iteration, for each of the SA questions, assessing this way if individual SA improved (SA_IR > 1) or not (SA_IR < 1).

Table 6.13: Comparison of the Average (and standard deviation) of Individual Questionnaires Scores (IQS) across iterations for each SA question

Freeze Probe Iteration	SA Questions					
	Q1		Q2		Q3	
	w/o	***w***	***w/o***	***w***	***w/o***	***w***
it#1	0,26 (0.18)	0,25 (0.23)	0,57 (0.16)	0,48 (0.23)	0,45 (0.25)	0,40 (0.21)
it#2	0,40 (0.25)	0,44 (0.23)	0,79 (0.15)	0,78 (0.21)	0,48 (0.23)	0,64 (0.30)
SA_IR	1,53	1,76	1,39	1,63	1,10	1,60
Wilcoxon p-value (N=15)	0,02	0,02	0,01	0,002	-	0,018

The analysis of the SA_IR shows that individual SA improves as the task unfold, regardless the experimental setting, since the majority of the results hold statistical significance. This may be indicative that the use of MCA does neither enhance or impair SA development, at least regarding the matters addressed by the three SA questions that constituted the SA questionnaire.

Moving the analysis of SA toward the Shared SA Questionnaire's Scores (SQS), presented on Table 6.14, through the average of the scores achieved by the teams, one may point that the common/overlap awareness exhibited by the team was greater regarding SA question number 2, meaning that the team members shared a common picture toward the task's assignments. Both regarding the enacted operational actions (SA question number 1) and the current state of network's equipment (SA question number 3) the SQS point to lower levels of common awareness. Although the latter one following the IQS trend, entails an exception, since under the *"w/"* experimental setting the SQS presents an higher value in respect to the second iteration of the freeze probes. Nevertheless, none of the differences between the scores of the *"w/o"* and *"w/"* experimental conditions hold statistical significance. Moreover, given the reduced number of teams that commonly accomplished the both iterations of the freeze probes, a further analysis on how SQS evolve between the iterations revealed impracticable.

The data collected from the log of the Operational Work Environment Simulator (OWES) had fuelled the analysis of the, Individual Diagnosis Efficiency (IDE), Individual Operational Efficiency (IOE), Team Diagnosis Efficiency (TDE) and Team Operational Efficiency (TOE) dependent variables which support SA measures derived from work processes tracing. The average of the results of IDE and IOE for the participating team members, as well as, the averages of TDE and TOE computed from the twelve participating teams are presented in Table 6.15.

Table 6.14: : Teams average (and standard deviation) of Shared SA Questionnaires (SQS) for each iteration on each SA question

Experimental Setting	SA Questions					
	Q1		Q2		Q3	
	it#1	it#2	it#1	it#2	it#1	it#2
w/o	0,19 (0.17) (N=12)	0,24 (0.10) (N=6)	0,63 (0.16) (N=12)	0,80 (0.13) (N=6)	0,31 (0.17) (N=12)	0,32 (0.13) (N=6)
w/	0,23 (0.27) (N=12)	0,37 (0.22) (N=10)	0,59 (0.27) (N=12)	0,80 (0.22) (N=10)	0,36 (0.26) (N=12)	0,58 (0.26) (N=10)
Wilcoxon p-value	0,88 (N=12)	- (N=5)	0,65 (N=12)	- (N=5)	0.87 (N=12)	- (N=5)

Table 6.15: Average (and standard deviation) of Individual and Team SA measures derived from work processes tracing

Dependent Variables		Experimental Setting		Wilcoxon p-value
		w/o	*w/*	
Individual SA (N=36)	IDE	0,92 (0.15)	0,93 (0.12)	-
	IOE	0,49 (0.45)	0,55 (0.40)	0,530
Team (holistic) SA (N=12)	TDE	0,91 (0.10)	0,92 (0.10)	-
	TOE	0,53 (0.20)	0,61 (0.21)	0,507

According the results presented on Table 6.15, at individual level it appears that team members diagnosis actions toward network 's equipment states (IDE) are enacted in a more informed manner than those concerning the operational actions enacted toward network's equipment recovery (IOE). While this later result is consistent with the low IQS results regarding SA question number 1 (awareness of enacted operational actions over equipment states) and SA question number 3 (awareness of equipment current state), the high value of IDE appears to be contradictory and thus worthy of further research.

At team level the TDE and TOE results follow a similar trend, and again the lower TOE result, which seems to indicate that teams enact operational actions less coordinated, is more in line with the scores from SA questionnaires, than the TDE result. This considerations holds whether the experimental settings since no statistically significant differences were found between the experimental settings regarding the IDE, IOE, TDE and TOE dependent variables.

Although the performed speech communications could further enrich this analysis, a drawback was experienced regarding how the teams had used the Speech Communication Support (SCS) subsystem of the Microworld. Figure 6.5, recovers the earlier presented Graphical User Interface (GUI) of the SCS subsystem (on chapter 5.3.2.1) to support the explanation of the reasons for its flawed usage.

Figure 6.5: Speech Communication Support (SCS) subsystem Graphical User Interface

During the experimental trials it was perceived that instead of pushing the corresponding button which classify the nature of the speech communication that was intended to be carried out, a large amount of end users (team members), press any button, in disregard of the nature of the communication content, for starting the conversation with the intended recipient team member. Moreover, even then in some cases they will press the right button, the conversation departs from the original intentions toward a broader scope. For instance, when a team member press the *assign* button with the goal of assigning a task to the receiver team member, they further engage in an exchange (request/report) of information, making the adopted codification schema for the content of the speech communications useless.

This fact had impaired the analysis of the Fraction of speech communications for operational information management (SC_IM) and Fraction of speech communications for team management (SC_TM) dependent variables. This aspect was revised and is further discussed in the description of the second round of experiments presented in the next section 6.2.2.

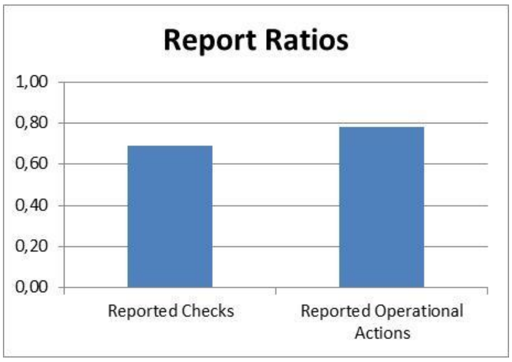
Nevertheless, one was still able to contrast the number of speech communications (NSC) enacted in each experimental setting which are presented in Table 6.16, by averaging NSC between the twelve participating teams. It is notorious that under the availability of the MCA, the *"w/"* experimental setting, teams engaged in substantially less speech communications. This result's statistical significance is supported on the Wilcoxon p-value which is in the acceptance threshold.

Table 6.16: Average and standard deviation of the number of speech communications

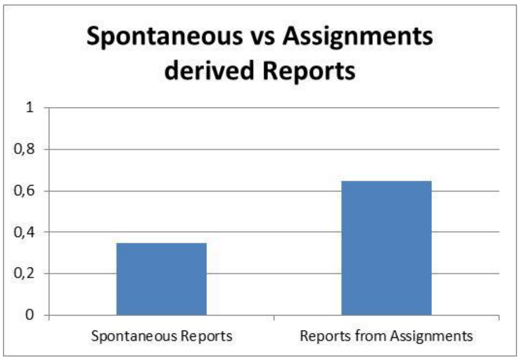
Dependent Variable		**Experimental setting**		**Wilcoxon p-value**
		w/o	*w/*	
NCS	Average	12,5	2,33	0.005 (N=12)
	STDEV	5,44	1.97	

MCA usage was unveiled by the defined Report Ratio (RR), Situation Monitoring screen usage percentage (SMS_U), Report screen usage percentage (RS_U), Assignments screen usage percentage (AS_U) and Notifications History Log screen usage percentage (NHL_U) dependent variables.

The Report Ratio (RR) exhibited an average value of 0.71 (stdev = 0.17), meaning that the twelve teams reported in average 71% of the enacted actions (including the check, restart and update actions), which appears to be an interesting value for supporting the documentation of the CIRM, that had been pointed a valuable contribution by the consulted Help Desk Teams (HDT), as described in chapter 5.2. Toward a deeper understanding of the distribution of such reports regarding their nature, one had further divided them on those concerning probe actions that support diagnosis (*check*) and those operational actions (*restart* and *update*) performed regarding recovery endeavour, this distinction is expressed in the graph presented in Figure 6.6a. According the graph, it appears that teams felt slightly more compelled to report actions concerning the recovery endeavour (78%) than those underlying diagnosis (69%). In Figure 6.6b is contrasted those reports that were driven by an existing task assignment (65%) with those that were performed on an spontaneous basis (35%), i.e., those reports that team members perform to give account to team level of their operational activity that was not derived from an issued assignment (at least not issued through the MCA).



a.



b.

Figure 6.6: Mobile Collaborative Application Information Report functional feature usage: a. diagnosis versus operational activity related reports; b. spontaneous versus task assignment derived reports

The analysis of the remaining MCA functional features usage is grounded on the average values exhibited on SMS_U (49%), RS_U (38%), AS_U (9%) and NHL_U (4%) dependent variables that rooted the graph shown on Figure 6.7.

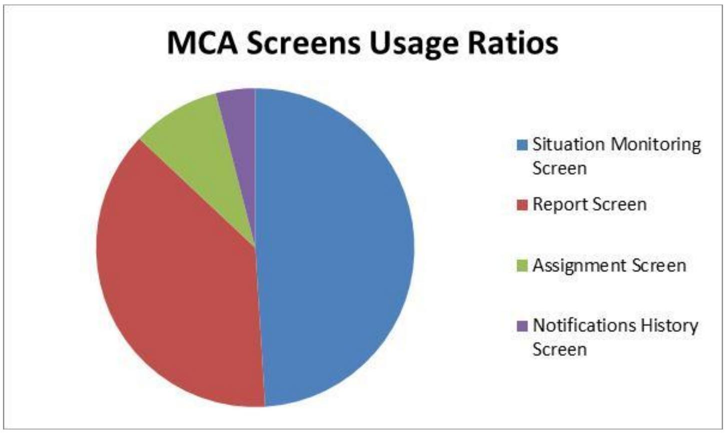


Figure 6.7: Percentage of usage of each Mobile Collaborative Application functional feature

Although, this values holds an inherent bias, since the *Situation Monitoring Screen* was the default main screen of the MCA, it is possible to draw some additional considerations. The *Notifications History Log* screen was the less used one, which may posit that team members had little time to browse the notification history log and were more focused on the momentary demands for which the information present in the S*ituation Monitoring* screen appear to be perceived as enough. Moreover, task assignments were less used when compared to the reporting feature, which lead one to think that the MCA constitute a valuable medium for managing operational information but not as much for team coordination. Although, this statement could be better supported if an analysis of the team speech communications was available. This assessment is recovered on the analysis of the second round of experiments results, where the content of speech communications was captured.

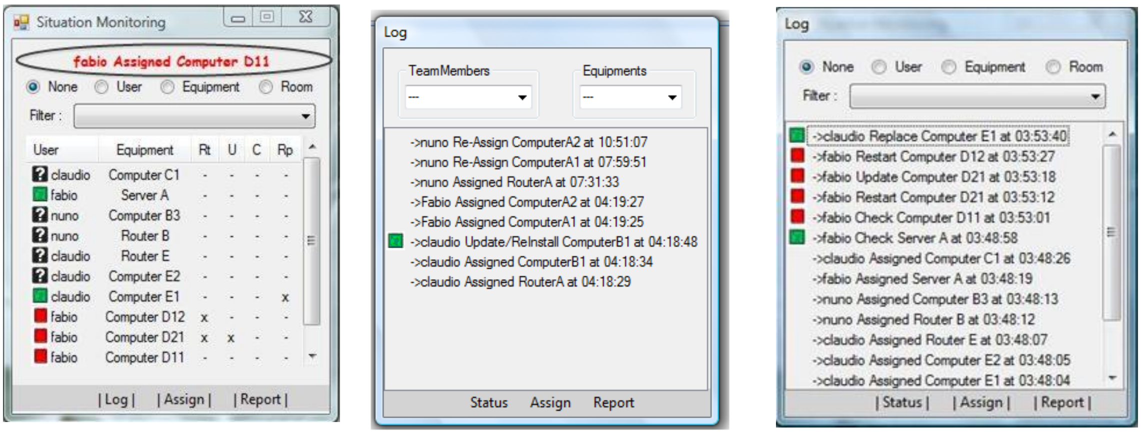
Nevertheless, the unveiled extensive use of the report functional feature appears to be aligned with the slightly improved scores on SA question number 3 (awareness of network's equipment states), at the time of the second freeze probe, exhibited under the “w/” experimental setting.

The lack of statistical significance to support this result, the unavailability of speech communication analysis and the inconsistencies found between some of the process tracing measures of SA regarding the SA questionnaires scores, motivated the second round of experiments that are described in the next section.

6.2.2 Second Round of Experiments

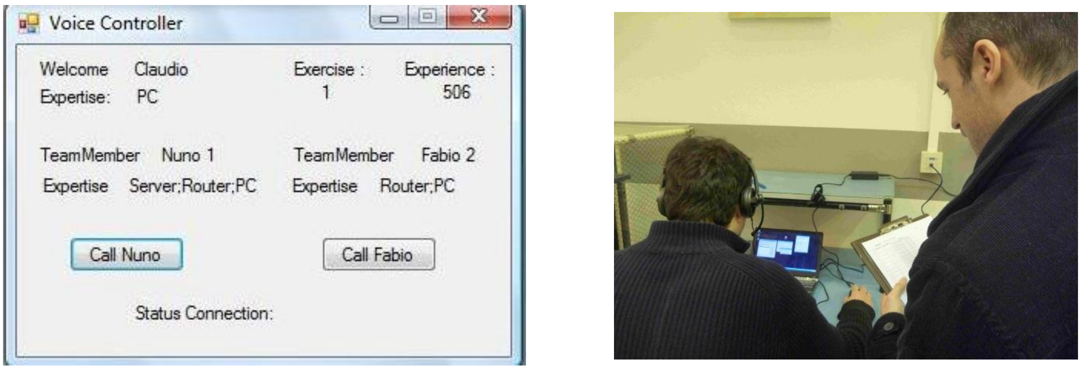
On the second round of experiments two revisions were made on the features of the Mobile Collaborative Application (MCA) delivered through Mobile Collaborative Application Emulator (MCAE) subsystem of the Microworld environment. The delivered notifications of the team activity will remain available in the notification area of the MCA screens (Figure 6.8a) until the end user (team member) click with the mouse pointer over the currently displayed notification or a new notification arrives, which in this case will override the previous one, if it was still being displayed. The earlier version of the MCA, used in the first round of experiments, will remove the notification from the notifications' area when the user browse between MCA's screens. This revision was made given the discussion of the first round of results provided on the previous section, which posit that although a significant number of information reports were done through the MCA's report functional feature, teams appear to be performing under a lack of more effective coordination that accounts for the on-going operational actions enacted by the team members. Also, toward the goal of promoting the readily retrieval of reported information, an enhancement of the filter mechanism of the *Notifications History Log* screen was made, Figure 6.8b and Figure 6.8c contrasts the previous and new version of the filter mechanism. This revision was motivated by the acknowledgment of its minimal usage yielded by the first round of experiments outcomes and considering the increased complexity of the networks infrastructures loaded on the Operational Work Environment Simulator (OWES) in this second round of experiments.

The Speech Communication Support (SCS) subsystem was also simplified since its initial formulation aiming to codify the content of speech communication revealed useless, as discussed in previous section. Therefore only one button was now available for starting a phone call with the intended team member, as shown in Figure 6.9a. The codification of the speech communications relied on the use of three previously trained observers. The observers hold a grid for coding the content of the enacted speech communication, the content could be (naturally not exclusively) codified as an exchange of operational information (e.g. reporting enacted actions and their outcomes) or team management related (e.g. task assignments). Figure 6.9b shows a photograph of a team member performing an experimental trial accompanied by his assigned observer.



a. b. c.

Figure 6.8 Mobile Collaborative Application revisions for the second round of experiments: a. Displayed notifications availability; b. Old Notification History Log filter mechanism and c. New Notification History Log filter mechanism.



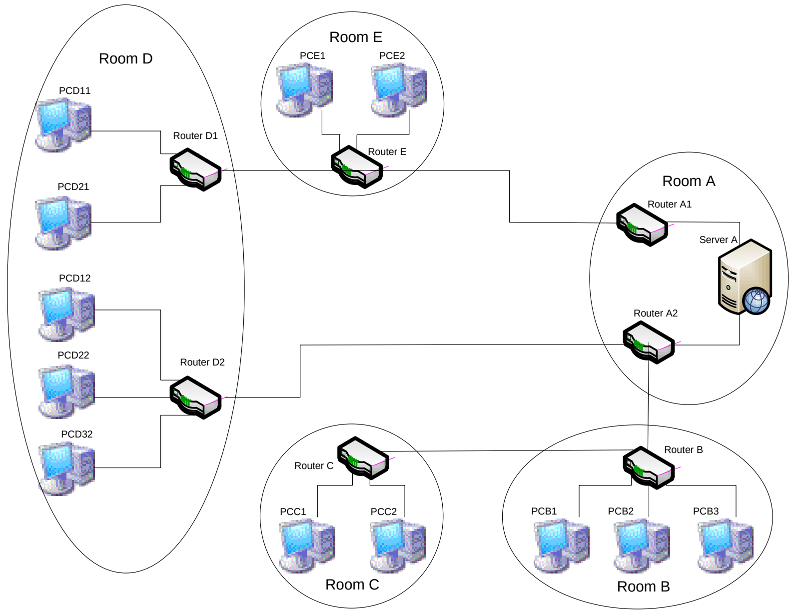
a. b.

Figure 6.9: Speech Communications content coding: a. Revised Speech Communication Support (SCS) subsystem graphical user interface and b. Photograph of an observer coding participant speech communications.

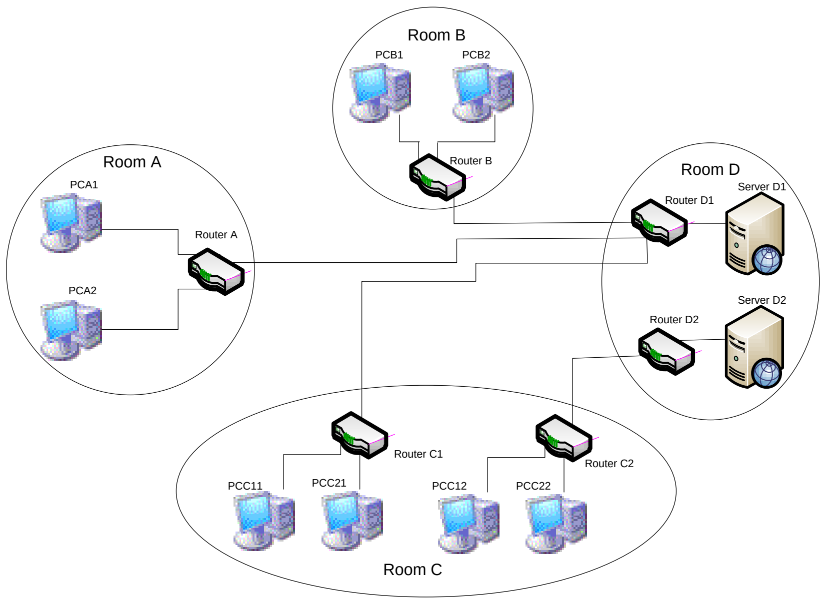
The herein described second round of experiments engaged a total of 11 additional teams totalizing 33 participants performing two new task scenarios supported on two more network architectures. The two additional network architectures loaded on the OWES are depicted on Figure 6.10a and Figure 6.10b, respectively.

To not overwhelm the present exposition the set of initial conditions and collection of events that bound task scenarios number five and six, respectively devised for network architectures number two and three are provided on Annex C.

One provide however, an overview of the each of those task scenarios in order for the reader get familiar with the issues that frame the team task.



a.



b.

Figure 6.10: Network architectures used on the second round of experiments: a. Network architecture number two and b. Network architecture number 3.

In task scenario number 5, the network connectivity problem is rooted on a broken link between the router in room A (RA1) and the server (SA). This link cannot be repaired (simulating for instance a network cable that cannot easily be replaced). Moreover, Router B is on *malfunctioning* state impairing room B and C network's equipment connectivity. The only solution to recover connectivity on RA1 is to *Connect* it to the Router RA2. This will establish naturally the room E connectivity but, room D will require that RD1 is *Restarted* in order to PCD11 and PCD21 recover their connectivity to the network. Regarding RB recovery, it is accomplished if an *Update* is performed over it. This will recover room B connectivity, but room C connectivity will only be completely established after a *Restart* is enacted toward RC.

The task scenario number 6, purports the following situation: the router D1 is preventing the connectivity of room A, router B is preventing the connectivity of room B and router C1 is irremediably damaged and cannot be replaced. To recover the network connectivity, RD1 requires to be *Replaced* which will re-establish room A connectivity, nevertheless PCA2 will require a *Restart* to achieve the *operating* state. Room B network's equipment will only recover connectivity after an *Update* action is performed on router RB. Since RC1 is unrecoverable, PCC11 and PCC21 must be *Connected* to RC2 in order to re-establish their network connectivity.

The SA questions administrated on the freeze probes questionnaires were revised and their formulation for this round of experiments is presented on Table 6.17. Given the increased level of complexity purported by the new networks architectures (network architecture number 5 and 6), the SA question number one is dynamically generated accounting only for the subset of network's equipment that are placed on the room that the team member is currently located and those that are in the path of that ones. The reasoning behind such option is of twofold, one relies on the claim that it was not reasonable to assume that participants hold awareness regarding the overall network's equipment states besides those that directly have impact on her/his current task, and second, if not formulated in this manner answering the SA question number 1 will be very disruptive of team member current focus. Figure 6.11, presents an example of the SA questionnaires delivered to team members through the freeze probes.

Also, considering the increased complexity, instead of asking team member the currently hold assignment of their team mates, SA question number 2 was replaced by asking team member where are teammates currently located. Therefore, in this second round of experiment, on had also departed from asking the fine-grain operational actions that the team member enacted and formulated SA question number 3 to ask team members their current perception of the underlying causes of the network connectivity problems.

Table 6.17: Situation Awareness questions administrated through the freeze probe questionnaires in the second round of experiments

Question ID	Question
Q1	What the state of each network equipment ?
Q2	Where are the other team members located ?
Q3	What are the devices causing the connectivity disruption (routers and servers) ?

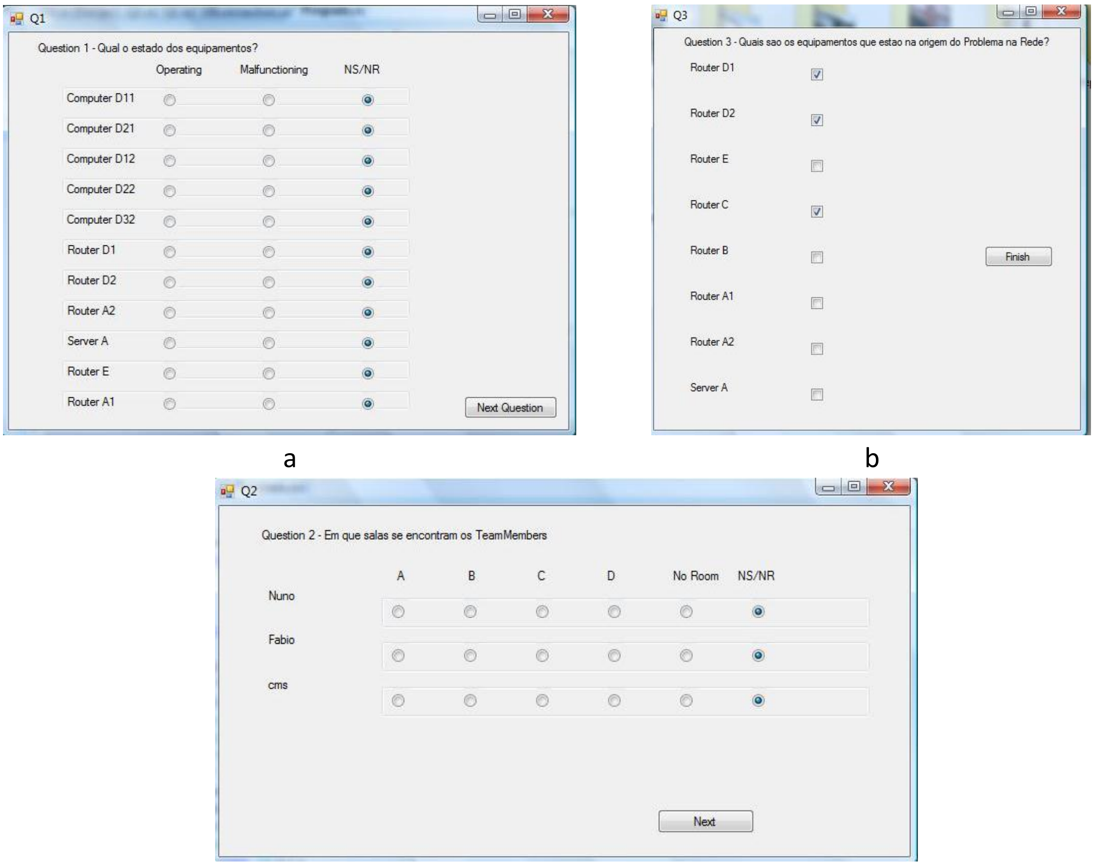


Figure 6.11: Screenshot of the freeze probe situation awareness questionnaires for the second round of experiments: a. SA question number one; b. SA question number two and c. SA question number three.

The following sub-section presents and discusses the results from the second round of experimental trials, which the apparatus was above described.

6.2.2.1 Results and Discussion

The presentation and discussion of the results achieved in this second round of experimental trials follows the same organization put forward in the section 6.2.1.1 that delivered the results from the first round of experiments. Therefore, one starts by reporting the teams task completion times and number of operational operations enacted through the OWES. Then introduce the results on SA related dependent variables, Individual SA Questionnaires Scores (IQS), Shared SA Questionnaires Scores (SQS), Individual Diagnosis Efficiency (IDE), Individual Operational Efficiency (IOE), Team Diagnosis Efficiency (TDE) and Team Operational Efficiency (TOE). It follows the results from those dependent variables respecting Team speech communications Number of Speech Communications (NCS), the Fraction of speech communications for information management (SC_IM) and the Fraction of speech communications for team management (SC_TM). And concludes the results presentation and discussion by accounting the Mobile Collaborative Application's usage, through the devised dependent variable in this realm, Report ratio (RR), Situation Monitoring screen usage percentage (SMS_U), Report screen usage percentage (RS_U), Assignments screen usage percentage (AS_U) and the Notifications History Log screen usage percentage (NHL_U).

According the values presented on Table 6.18, in this second round of experiments the teams took a longer time to accomplish the task scenarios under the “w/” experimental setting. The difference, in average of 2 minutes, holds statistical significance, which may indicate that the handling of MCA could be more demanding in more complex task scenarios. Table 6.18 also puts forward that teams perform slightly less operational actions toward task completion in the “w/o” experimental setting, although this difference does not hold for statistical significance.

Table 6.18: Average (and standard deviation) of task completion times and number of operations of teams in both experimental settings (2nd round).

	Experimental Setting		**Wilcoxon p-value**
	w/o	***w/***	
Average completion time	8,23 (1.85)	10,55 (1.90)	0,016
Average # of enacted operations	38,18 (12.05)	41,27 (9.52)	0,262

Before presenting the analysis of SA questionnaires scores on have to give notice that, given the triggering condition of the freeze probes underlying SA questionnaires administration is framed on the number of operational activities that the team collective enact, the number of freeze probes achieved by each team was different and is detailed in Table 6.19.

Table 6.19: Number of SA questionnaires freeze probe iterations achieved by the each team (2nd round).

	Experimental Setting			
	“w/o”		***“w”***	
Freeze Probes	**Teams ID**	**# of Teams**	**Teams ID**	**# of Teams**
Iteration #1	1,2,3,4,5,6,7,8,10,11	11	1,2,3,4,5,6,7,8,10,11	11
iteration #2	1,2,3,4,5,6,7,8,10,11	11	1,2,3,4,5,6,7,8,10,11	11
iteration #3	3,4,5,6,7,8,10	8	1,2,3,4,5,6,7,8,10,11	11

On Table 6.20 are presented the IQS average (and standard deviation) regarding the three SA questionnaire’s questions for each freeze probe iteration. The IQS holds similar values across experimental settings, with the exception of those related with the SA question number 2, over which higher scores were achieved in the third iteration of the freeze probes, in the “w/” experimental setting, and those concerning SA question number 3 that were lower on the “w/” experimental setting. Those differences however does not hold statistical significance since the Wilcoxon p-value is above the minimal considered threshold (*0,05*) for stating that the difference holds statistical significance.

Higher scores were consistently achieved through all freeze probe iterations, regarding SA question number 1, which addresses the awareness regarding the state of the network’s equipment that had impact on the current team member operational activity. SA question number 2 scores, which account for the awareness of the current location of each teammate had decreased as the task unfolds and scores from the SA question number 3, concerning the perception of the underlying causes of the network connectivity problem were the most lower ones, which had only experienced a slight improvement on the last (third) freeze probe iteration. This results sustain whether the experimental setting, since no statistically significant differences were found.

Table 6.20: Average (and standard deviation) of Individual Questionnaires Scores (IQS) for each iteration on each SA question (2nd round).

Exp. Setting	SA Questions								
	Q1			Q2			Q3		
	it#1 (N=33)	**it#2** (N=33)	**it#3** (N=24)	**it#1** (N=33)	**it#2** (N=33)	**it#3** (N=24)	**it#1** (N=33)	**it#2** (N=33)	**it#3** (N=24)
w/o	0,60 (0.28)	0,77 (0.24)	0,78 (0.25)	0,70 (0.28)	0,58 (0.27)	0,35 (0.23)	0,33 (0.19)	0,31 (0.19)	0,48 (0.29)
w	O,59 (0.30)	0,74 (0.23)	0,77 (0.27)	0,78 (0.31)	0,57 (0.32)	0,47 (0.31)	0,24 (0.16)	0,36 (0.20)	0,49 (0.27)
Wilcoxon p-value	-	-	-	-	-	0,15	0,05	-	-

The analysis of IQS regarding the assessment of their progression through the freeze probe iterations, relies on the earlier defined (in the previous section 6.2.1.1) Situation Awareness Improvement Ratio (SA_IR) which the value are provided on Table 6.21.

Table 6.21: Comparison of the Average (and standard deviation) of Individual Questionnaires Scores (IQS) across iterations for each SA question (2nd round).

Freeze Probe Iteration	SA Questions					
	Q1		Q2		Q3	
	w/o	***w***	***w/o***	***w***	***w/o***	***w***
it#1	0,60 (0.28)	0,59 (0.30)	0,70 (0.28)	0,78 (0.31)	0,33 (0.19)	0,24 (0.16)
it#2	0,77 (0.24)	0,74 (0.23)	0,58 (0.27)	0,57 (0.32)	0,31 (0.19)	0,36 (0.20)
SA_IR (it#2/it#1)	1,28	1,25	0,83	0,73	0,94	1,5
Wilcoxon p-value	0,0043 (N=33)	0,01 (N=33)	0,017 (N=33)	0,0029 (N=33)	- (N=33)	0,002 (N=33)
it#2	0,76 (0.25)	0,74 (0.23)	0,61 (0.25)	0,57 (0.32)	0,32 (0.21)	0,36 (0.20)
it#3	0,78 (0.25)	0,75 (0.27)	0,35 (0.23)	0,45 (0.31)	0,48 (0.29)	0,52 (0.27)
SA_IR (it#3/it#2)	1,02	1,01	0,57	0,79	1,5	1,4
Wilcoxon p-value	- (N=24)	- (N=33)	0,0046 (N=24)	0,09 (N=33)	0,026 (N=24)	0,003 (N=33)

The SA_IR shows that team members had increased their awareness regarding SA question number 1 from the first to the second freeze probe and kept a similar level of awareness, regarding the state of the network's equipment that bound their current activity at the moment of the third freeze probe, whether the experimental setting they were performing. Regarding SA question number 2, the SA_ IR supports that team members loose awareness of other teammates' location as the task scenario unfold, although that declination is less pronounced under the "w/" experimental setting. The perception of the underlying causes of the loss of network connectivity purported by the task scenario (SA question number 3) experience an improvement more noticeable under the "w/" experimental setting, although upon the third freeze probe the difference becomes minimal between the two experimental settings.

Moving the analysis to the team shared awareness supported on the Shared SA Questionnaire's Scores (SQS) presented on Table 6.22, which provide the average values devised from all participating teams.

Table 6.22: Teams average (and standard deviation) of Shared SA Questionnaires (SQS) for each iteration on each SA question (2nd round)

Exp. Setting	SA Questions								
	Q1			Q2			Q3		
	it #1	it #2	it #3	it #1	it #2	it #3	it #1	it #2	it #3
w/o	0,27 (0.36)	0,66 (0.29)	0,67 (0.21)	0,65 (0.21)	0,45 (0.19)	0,29 (0.17)	0,20 (0.11)	0,17 (0.14)	0,31 (0.23)
w/	0,25 (0.22)	0,63 (0.33)	0,60 (0.26)	0,76 (0.24)	0,46 (0.21)	0,32 (0.21)	0,13 (0.13)	0,26 (0.13)	0,32 (0.13)
Wilcoxon p-value	- (N=11)	- (N=11)	- (N=8)	0,109 (N=11)	- (N=11)	- (N=8)	0,203 (N=11)	0,139 (N=11)	- (N=8)

The SQS analysis puts forward the following trends regarding the common awareness hold by the teams. SA question number 3 exhibit the lowest scores of the three SA questions, meaning that teams hardly achieve a full collective understanding on the underlying causes of the network connectivity failure. No significant differences were found across the two experimental settings. The evolution of the SQS on SA question number 1, indicate that the team developed, as the task scenario unfold, a more shared understanding regarding the state of those network's equipment that should be commonly accounted. But such understanding reaches a plateau on the second freeze probe. The SQS on SA question number 2, follows the IQS trend, as the teamwork evolve the team loose also the consensus regarding the location of each team member. Also in both SA question number 1 and SA question number 2 no significant difference between the two experimental settings were found.

The work process tracing related dependent variables, which more readily address the operational level of teamwork are presented in Table 6.23. At the individual level, team members exhibit an increased awareness regarding the diagnosis related actions, whether the experimental setting. While regarding the operational actions toward the recovery endeavour, they appear to be more informed under the "w/" experimental setting. Although this is a promising result, a stronger statement on this difference will require further research since that with the current data it does not hold statistical significance. In this second round of experiments, both IDE and TOE appear to be more aligned with IQS, particularly to those regarding SA question number 1 (the awareness of relevant network's equipment states), which is the question that more readily tackle the issues behind the reasoning that rooted the IOE and IOE definition.

At team level both TDE and TOE presents similar results across the experimental settings. Although TOE may indicate that the operational teamwork was slightly more coordinated under the "w/" experimental setting, the difference does not hold statistical significance.

Table 6.23: Average (and standard deviation) of Individual and Team SA measures derived from work processes tracing (2nd round).

Dependent Variables		Experimental Setting		Wilcoxon p-value
		"w/o"	*"w/"*	
Individual SA (N=33)	IDE	0,71 (0.24)	0,73 (0.25)	-
	IOE	0,68 (0.41)	0,80 (0.25)	0,937
Team (holistic) SA (N=11)	TDE	0,63 (0.10)	0,66 (0.16)	-
	TOE	0,78 (0.19)	0,82 (0.12)	0,575

Given that analysis of team speech communications was viable in this second round of experiments, due the codification of the nature of their content, by three trained observers, one enrich the team level analysis rooted on the results achieved on the respective dependent variables (NCS, SC_IM and SC_TM). Table 6.24 present the results from the codifications done by the observers, and the computation of the SC_IM and SC_TM dependent variables. Their average values are summarized on Table 6.25, which supported the discussion regarding the enacted speech communications.

Keeping the trend of the first round of experiments teams exhibit fewer speech communications under the "w/" experimental setting. The difference in this case, on contrary to the first round of experiments, remains only close to statistical significance. The nature of the content of speech communications, assessed through the SC_IM and SC_TM dependent variables, reveals that under the "w/" experimental setting, although teams keep relying on speech communications for team management, operational information management is enriched by the use of the MCA, given the SC_IM lower value in that setting and that the difference is in the threshold of statistical significance.

This statement is contrasted with the MCA usage analysis supported on the respective dependent variables: Report Ratio (RR), Situation Monitoring screen usage percentage (SMS_U), Report screen usage percentage (RS_U), Assignments screen usage percentage (AS_U) and Notifications History Log screen usage percentage (NHL_U).

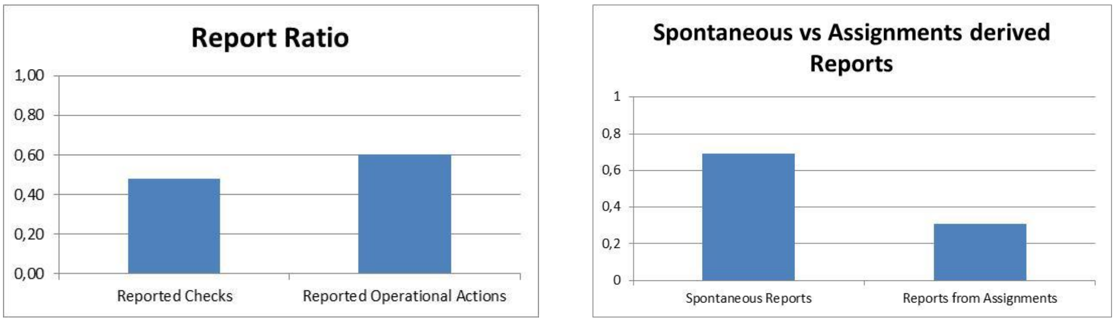
Table 6.24: Team speech communications codification and related dependent variables computation.

	w/o					w/				
Team	**Info. Manag. coding category frequency**	**Team Manag. coding category frequency**	**# of phone calls**	**SC_IM**	**SC_TM**	**Info. Manag. coding category frequency**	**Team Manag. coding category frequency**	**# of phone calls**	**SC_IM**	**SC_TM**
1	14	6	5	2,80	1,20	21	10	9	2,33	1,11
2	17	5	9	1,89	0,56	19	6	9	2,11	0,67
3	23	9	7	3,29	1,29	20	8	9	2,22	0,89
4	30	8	12	2,50	0,67	0	3	8	0,00	0,38
5	22	8	10	2,20	0,80	12	4	7	1,71	0,57
6	30	7	12	2,50	0,58	10	4	5	2,00	0,80
7	22	9	12	1,83	0,75	12	6	12	1,00	0,50
8	21	4	9	2,33	0,44	0	2	1	0,00	2,00
9	37	9	13	2,85	0,69	19	8	10	1,90	0,80
10	22	7	11	2,00	0,64	6	2	5	1,20	0,40
11	20	3	11	1,82	0,27	16	7	10	1,60	0,70
Avg	23.45	6.82	10.09	2.36	0.71	12.27	5.45	7.73	1.46	0.80
stdev	6.52	2.09	2.42	0.48	0.30	7.66	2.66	3.07	0.83	0.45

Table 6.25: Average (and standard deviation) of the team speech communication related dependent variables

Dependent Variable	**Experimental Setting**		**Wilcoxon p-value**
	w/o	***w/***	
NCS	10,09 (2.43)	7,73 (3.07)	0,007
SC_IM	2,36 (0.48)	1,46 (0.83)	0,005
SC_TM	0,72 (0.30)	0,80 (0.45)	0,789

In this second round of experiments the Report Ratio (RR) had decreased regarding the one exhibited in the first round of experiments, in now holds an average value of 0.48 (stdev = 0.18) while in first round RR was 0.71.This means that in this round of experiments teams report through the MCA roughly only half of the operationally enacted actions. Although this declination may be attributed to the increased task complexity, it still holds an interesting value regarding the documentation of the CIRM, which has it had been put forward is a valuable asset for Help Desk Teams (HDT). The nature of such reports is depicted on Figure 6.12a and Figure 6.12b. The trend of the previous round of experiments is followed regarding the type of operational actions reported, which posit that more reports are performed regarding the enacted operational actions (60%) then those regarding the performed diagnosis checks over the network's equipment (48%), as shown in the graph in Figure 6.12a. Nevertheless, the spontaneous reports are now greater (69%) than those that are directly derived from task assignments (31%), which consist in a twist regarding the first round of experiments. This is consistent with the previous speech communication analysis which yield that teams had mainly relied on speech communications for team managements (task assignments) and that operational information management appears to had been further supported by the use of the MCA.



a. b.

Figure 6.12: Mobile Collaborative Application Information Report functional feature usage: a. diagnosis versus operational activity related reports; b. spontaneous versus task assignment derived reports (2nd round).

The analysis of the reaming MCA functional features usage is performed upon the values of SMS_U (49%), RS_U (43%), AS_U (5%) and NHL_U (2%), which rooted the graph shown on Figure 6.13.

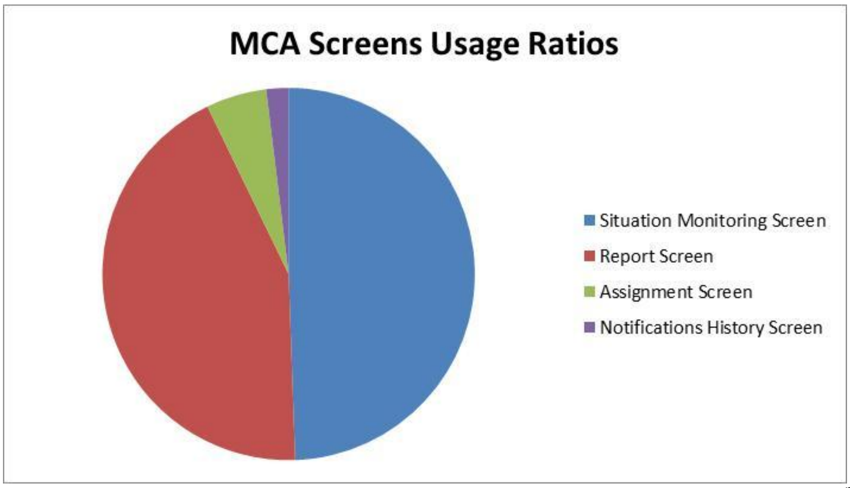


Figure 6.13: Percentage of usage of each Mobile Collaborative Application functional feature (2nd round)

According the graph on Figure 6.13, MCA functional features usage, despite the adjustment made since the first round experiments, on the notifications delivery and *Notifications History Log* filtering mechanisms, was similar to those on the first round. The graph shows that *Notification History Log* and *Assignments Management* screen were the less used ones. While in the first round of experiments speech communication analysis was impaired, due the lack of codification regarding the nature of the content of speech communications, in this round such analysis was possible and the MCA usage analysis seems to support the above discussion that states that teams relied on MCA to report operational information but still address team management (assignments) relying on speech communications.

6.3 Discussion

This chapter starts with the presentation of the developed experimental design that bounded the conducted experimental trials. The experimental design entails the followed experimental protocol and the definition of the independent, controlled and dependent variables of the experiments. The dependent variables were framed into three main categories, namely those addressing situation awareness, team speech communications and mobile collaborative application usage. The dependent variables in the realm of situation awareness category were further divided into those that address the individual and team level accounting for the shared and holistic dimensions of the team situation awareness construct. The independent variable of the experiments consisted on the availability or not of the Mobile Collaborative Applications Emulator (MCAE) on the Microworld environment, which defined the two experimental settings that the teams were submitted, for short, the “w/o” (without MCAE) and the “w/” (with MCAE) settings.

The results regarding the derived dependent variables were presented and discussed regarding two rounds of experiments. The first round was comprised by 12 experimental trials engaging respectively 12 teams, while on the second round were conducted 11 experimental trials with 11 additional teams.

Such results point that the use of the MCA have little direct impact on team situation awareness development. However one notice, that although SA dependent variables results does not support that SA development improved with the introduction of the MCA, its use appears to not either impair SA development. Nevertheless, as the complexity of team tasks increase teams took slightly longer task completion times when using the MCA. This may motivate a further account on usability aspects, that can borrow from the results that yielded that team members are more informed regarding the network equipment’s states (which was consistent across the two rounds of experiments) than on the fine grain operational activities that were enacted over them by others, even in less complex task scenarios, as the ones that were provided on the first round of experiments. Moreover, it appears that team members better hold the trace of task assignments (given the higher scores on SA question number 2 on the first round of experiments) than on the current location of their team mates (given the low scores on SA question number 2 on the second round of experiments).

The adopted use of MCA by team members was mainly toward the support of operational information reporting. This finding is supported by two results.

First, it was consistently experienced, across the two rounds of experiments, a reduction on the number of speech communications on teams endowed with MCA, although as the task complexity increase this difference between the “w/o” and “w/” experimental settings becomes smaller. Which is consistent with other research results discussed on 4.3.1, and may point for the validity of the Micrworld as an experimental instrument.

Since, it had been acknowledged that, given the richness of speech communications, capturing its contents in a more automated manner, is a challenging endeavour; the analysis of enacted communications among the team members was grounded on its codification made by three previously trained observers. Such analysis unveiled that teams rely on MCA for managing operational information and speech communications are delegated to team management.

Secondly, teams clearly choose to report operational activity through the MCA. This finding is grounded on the analysis of the MCA functional features usage, that had put forward that the report feature was the most used one. As the task complexity increase the number of spontaneous reports also increased. Reports were classified as, spontaneous, when they were volunteered made by team member without a specific assignment issued through the MCA. This finding along with the low use of the assignments management MCA's feature is consistent with the role of MCA of providing a medium for report operational information in disregard for team management support, which was still carried out through speech communications.

The unveiled MCA usage meets the requirement of documenting the CIRM operational activity, put forward in chapter 5, by the consulted help desk teams, which had pointed out that such constitute a valuable asset for post-mortem analysis and future reference.

One should notice however that as the task complexity increases the ratio between reported actions and the real enacted actions is reduced. Thus, and according the wide adoption of the report feature, future MCA design initiatives should account for an improved usability regarding operational information reporting.

7 Conclusions

This concluding chapter discusses the research work in terms of its main objectives and hypothesis.

Research Objective number one posits that, in order to evaluate the role of Mobile Collaborative Applications (MCA) in Critical Incidents Response Management (CIRM), one should devise a characteristic set of functional features by considering the specific affordances supported by mobile technology, namely, situated use and real-time information sharing. Such was made, in chapter 4.3.2, based on guidelines reported on the related literature.

The results from the conducted experiments point out that the design of MCA for assisting the operational level of teamwork in CIRM contexts should more readily attend to functional features that enhance operational information reporting, since this was the most used MCA feature by the teams in the experiments. In fact, although speech communications support was also available, the team members extensively used MCA to give account of their activities. Conversely, MCA functional features toward team management were less used, since team management was mainly accomplished through speech-based communication.

These results are consistent with current research (reviewed in section 4.3), which indicates that under stressful and time-critical contexts, speech communication constitutes the primary mean used for work coordination. Nevertheless, such research works, and in line with the inputs obtained from the consulted Help Desk Teams (HDT), that had constituted the target application domain for the practical applicability of the theoretical proposals, also emphasise the value of a medium to support the accountability of operational work in a persistent manner. Such medium, constitutes a complementary communication channel, that to some extent and under certain circumstances, may relieve the number of speech communications necessary to coordinate work; and also document CIRM instances for future reference.

A Microworld supported the evaluation of MCA in quasi-naturalistic oriented experiments. According with the related literature, current Microworld developments lack a frame of reference in order to bring it forward as a well-established experimental paradigm for evaluating collaborative applications when other existing evaluation methods reveal inappropriate given the phenomena, target application domain or promised solutions maturity constrains.

The *Research Objective number two* of devising a foundational set of Microworld building blocks was addressed in chapter 4, by comprehensively specify, in a domain independent way, the core components necessary to evaluate collaboration and collaborative applications in quasi-naturalistic apparatus.

This specification was validated through a prototype that offers a set of modular software components to promote reusability in various contexts/domains

One acknowledges that the Mobile collaborative Application Emulator (MCAE) subsystem would requires major changes to accommodate other promising MCA functional features and specific work contexts/domains.

The remaining subsystems may be reused in other contexts/domains without significant changes. The Operational Work Environment Simulator (OWES), given its state space orientation, delivered the OWES engine which may easily accommodate other operational contexts/domains. Possible changes may be mostly toward the OWES graphical user interface provided on its presentation layer, for face validity requirements regarding the specificities of the domain and research aims.

The Situation Awareness Measurement Support (SAMS) subsystem through its SA Questionnaires Management component offers a number of possibilities to inquiry experiments participants regarding a given research initiative.

The Experimental Control Manager (ECM) subsystem readily supports two main purposes. The first one, is providing the means necessary to manipulate OWES according various task scenarios directed to study teamwork regarding various phenomena of interest; while the second one is, affording the control toward each of the Microworld subsystems regarding their functional features in order to bound the independent and control variables of an experimental procedure.

The implementation of the Speech Communication Support (SCS) subsystem emulates typical phone and radio communications enacted by teams in collaborative settings. Although it has revealed short on codifying the content of such communications. This shortcoming was overcome by relying on trained observers to codify the contents of speech communication contents, since it constitute a hardly neglectable dimension of analysis of teamwork, as deeply discussed in the related literature.

Therefore, if it is intended to avoid the use of observers, given that the adoption of a Microworld as an experimental paradigm pursues a cost-/resource-effective approach to the evaluation process, improving the collection of speech communication content related data, especially with large groups, is worthy of further research.

By logging all operational activity in the context of a given task scenario, the Microworld implementation affords the analysis of teamwork at multiple granularity levels, and thus supports the scrutiny of different phenomena of interest.

The work considered Team Situation Awareness (TSA) as the main dimension of evaluation, since the literature indicates that TSA constitutes a key team asset in highly dynamic and demanding work contexts.

This goal had been addressed by *Research Objective number three*, the definition of representative TSA measures and associated measurement techniques applicable in Microworlds. As afore mentioned, the developed Microworld uses Situation Awareness

Questionnaires administrated to teams during the task, combined with activity logging, as the main mechanisms to support TSA measurement.

The TSA measures, that had constituted the experiments' dependent variables, comprising some that are grounded upon more established dimensions of the construct (individual and shared), but also extended those and, moves forward on the definition of some that addresses the holistic dimension of TSA, for which the related research literature has been pointing out as necessary but so far their practical definition has been poorly accomplished.

The selection of Help Desk Teams (HDT) as the target application domain to support the conducted experimental trials meets *Research Objective number four*. The representativeness of this domain to CIRM has been considered in chapter 5. Using a representative domain is necessary to drawn practical results from the Microworld experimental approach.

The obtained results partially support *Research Hypothesis number one*, which was formulated as: *The introduction of Mobile Collaborative Applications to support Critical Incidents Response Management will drive new ways on how teams develop Situation Awareness*.

In fact, the introduction of MCA to support teams' CIRM appears to have little impact on TSA, since TSA measures hold similar results whether teams are using or not the MCA. Although TSA is not improved by MCA, the use of MCA did not impair TSA and allowed the persistence of the enacted CI response activities, which may be relevant for training and post-hoc assessment. Moreover, it was observed that teams did not disdain the adoption of MCA, they indeed changed their operational information management behaviour, which in this aspect partially holds the devised hypothesis.

Research Hypothesis number two, posits that*: Microworlds provide a valuable experimental paradigm to develop a fine-grain understanding on how teams use Mobile Collaborative Applications in Critical Incidents Response Management.*

Indeed, the large data sets of experimental data, gathered at different granularity levels, had allowed to address the main research question: *How to achieve a fine-grain understanding of the role of mobile collaborative applications regarding team situation awareness in critical incidents response management contexts?*

This statement is supported by the ability to explore *Research Hypothesis number one* beyond causal epistemology towards a keen understanding of the factors underlying the extent that it sustains. The data resulting from the Microworld use has made possible to analyse in a very comprehensive way teamwork and the meditation role of technology in teamwork, especially regarding speech communications support and interaction with mobile collaborative applications, in operational contexts that are beyond pre-established work procedures and posit demanding work conditions, as the ones purported by CIRM.

The results of the experiments yield that, as task complexity increases, teams engage in more speech-based communication, which is consistent with prior research and seems to provide evidence about the validity of Microworlds as experimental paradigm.

Accordingly, the instrumental use of Microworlds affords to develop insights on the impact of interventions, since their earlier stages, regarding not only those technology oriented (as the introduction of collaborative work support systems, such as MCA), but also, for instance, the effects of training programs that addresses the operational level of teamwork.

Furthermore, the Microworlds approach provides environments that are safe, cost effective controlled and posit minimal constrains regarding the individuals activity. Therefore, instead of the more typical black-box evaluations, rooted on absolute measures of team performance (e.g. overall efficiency and task completion time), Microworlds may be thought to offer a white-box oriented evaluation process since they provide more fine-grained and richer understandings of teams’ performance, and thus enrich the iterative design-evaluation cycle of promised solutions.

In order to sustain this line of though, further work remains to be done on bringing together contributions from several research fields such as psychology, cognitive science, organizational psychology, and sociology, which have been identifying critical issues surrounding human individual and collective behaviour, and integrating such knowledge on software development and evaluation.

Annex A – Dissemination of Work

A.1 Publications

International journals

Sapateiro, C., N. Baloian, P. Antunes and G. Zurita (2011) "Developing a Mobile Collaborative Tool for Business Continuity Management." Journal of Universal Computer Science, 17(2), pp. 164-182. ISI 5-year impact factor: 0.788.

Book chapters

Antunes, P., C. Sapateiro, G. Zurita and N. Baloian (2011) Development of a Mobile Situation Awareness Tool Supporting Disaster Recovery of Business Operations. Supporting Real Time Decision-Making: The Role of Context in Decision Support on the Move, F. Burstein, P. Brezillon and A. Zaslavsky. Annals of Information Systems, vol. 13, pp. 337-360. US, Springer.

Sapateiro, C., Antunes, P., Zurita, G., Baloian, N., Vogt, R. (2008). "Evaluating a Mobile Emergency Response System". 14th CRIWG. Omaha, US. September 2008. Lecture Notes in Computer Science, vol. 5411, Heidelberg, Springer-Verlag.

Sapateiro, C., Antunes, P., Zurita, G., Baloian, N., Vogt, R. (2008). "Supporting Unstructured Activities in Crisis Management: A Collaboration Model and Prototype to Improve Situation Awareness". 2º International Symposium on Mobile Information Technology for Emergency Response. Bonn,Germany. May 2008. Lecture Notes in Computer Science, vol. 5424-010, Heidelberg, Springer-Verlag.

International conferences proceedings

Sapateiro, C., Ferreira, A. and Antunes, P. (2011) Using Microworlds to Study Teamwork at the Cognitive Level. VII Simpósio Brasileiro de Sistemas Colaborativos (SBSC), Paraty, Brazil, ACM Press. Full paper acceptance ratio: 39%. **Best paper award**.

Sapateiro, C., A. Ferreira and P. Antunes (2011) Evaluating the Use of Mobile Devices in Critical Incidents Response: A Microworld Approach. 20th IEEE International Conference on Collaboration Technologies and Infrastructures, Paris, France. IEEE CS Press.

Antunes , P., C. Sapateiro and J. Pino (2011) Supporting Experimental Collaborative Systems Evaluation. 15th International Conference on Computer Supported Cooperative Work in Design (CSCWD), Lausanne, Switzerland. IEEE.

Antunes , P., C. Sapateiro, J. Pino, S. Ochoa and V. Herskovic (2010) Awareness Checklist: Reviewing the Quality of Awareness Support in Mobile Collaborative Applications. Groupware: Design, Implementation, and Use. 16th CRIWG Conference on Collaboration and Technology, Maastricht, The Netherlands, vol. 6257, pp. 202-217. Heidelberg, Springer-Verlag.

Sapateiro, C., Antunes, P. (2009). " An Emergency Response Model Toward Situational Awareness Improvement ". ISCRAM. Gothenburg, Sweden. May 2009.

Sapateiro, C., Baloian, N., Antunes, P., Zurita, G. (2009). "Developing Collaborative Peer-to-peer applications on mobile devices". IEEE 13th International Conference on CSCW in Design. Santiago, Chile. April 2009.

Sapateiro, C., Antunes, P., Zurita, G., Baloian, N., Vogt, R. (2008). "Supporting Crisis Management Processes by Wirelessly Interconnected Tablet-PCs ". IEEE International Conference on e-Business Engineering. Xi'An, China. October 2008.

Sapateiro, C., Antunes, P., Zurita, G., Baloian, N., Vogt, R., Infante, P. (2008). "Assisting Teams in Emergent Work Activities". GDN. Coimbra, Portugal. June2008

Sapateiro, C., Antunes, P. (2008). "Supporting unstructured Work Activities in emergent work processes". 10th ICEIS. Barcelona, Spain. June 2008

Sapateiro, C., Antunes, P. (2008) . "Crisis Management: A collaboration model for unstructured activities", 1st Workshop on Complexity in Social Systems. Lisbon, Portugal. January 2008.

Other Related publications

Antunes , P., C. Sapateiro, G. Zurita and N. Baloian (2010) Integrating Spatial Data and Decision Models in a E-Planning Tool. Groupware: Design, Implementation, and Use. 16th CRIWG Conference on Collaboration and Technology, Maastricht, The Netherlands, vol. 6257, pp. 97-112. Heidelberg, Springer-Verlag

Sapateiro, C., Vilhena, B., Moura, P. (2008). "Improving organizational collaboration based on document tagging concept". 10th ICEIS. Barcelona, Spain. June 2008

Sapateiro, C., Grosso, S. (2010). "Capturing Distributed Contributions to an Informal Work Process: A Hospital Facility Case Study". Handbook of Research on Developments in e-Health and Telemedicine: Technological and Social Perspectives. Edited by M. Manuela Cunha, Antonio Tavares and Ricardo Simoes. Published by IGI Global.

A.2 Research Projects

(2012-14) Simulated Task Environment for Collaborative Systems Evaluation. National Project, grant by FCT (PTDC/EIA-EIA/117058/2010). Participants: FCUL.

(2010-11) Sacim - Situation Awareness in Critical Incident Management. National project, grant by FCT (PTDC/EIA-EIA/102875/2008). Participants: FCUL.

(2007-09) A-CSCW - Attentive CSCW. Financed by FCT (PTDC/EIA/67589/2006). Participants: FCUL, INESC.

A.3 Doctoral Consortiums

PhD Research Report accepted for ISCRAM'09 doctoral consortium discussion. Gothenburg, Sweden. May 2009.

PhD Research Report accepted for CRIWG'08 doctoral consortium discussion. Omaha, Nebraska - USA. September 2008.

Annex B – Experiments Participants Manual

Exercício de Estudo de Desempenho em Trabalho Colaborativo

<u>Participação</u>

A participação na experiência tem como pré-requisitos:

1. A constituição de um grupo de 3 elementos
2. A leitura atenta deste documento (<u>o mesmo deverá acompanha-lo durante a experiência</u>)
3. Assistir a uma das sessões de esclarecimento que se irão realizar
4. Um agendamento prévio da data e hora de participação do grupo
5. Assinatura do Anexo 2 deste documento.

Os agendamentos serão feitos nas sessões de esclarecimento que têm o seguinte calendário:

Data	Hora	Local
25-01-2011	15h	Sala E109 ESTSetúbal
25-01-2011	16h30	"Learning Factory" do CENI, no edifício dos serviços de acção social no campus do IPS

Cada experiência dura no máximo 1,5 hora e serão realizadas nos seguintes locais:

Data	Local
26-01-2011	"Learning Factory" do CENI, no edifício dos serviços de acção social no campus do IPS
27-01-2011	Sala F203 ESTSetúbal
28-01-2011	"Learning Factory" do CENI, no edifício dos serviços de acção social no campus do IPS

<u>Incentivos:</u>

- Todos os participantes terão potencialmente uma bonificação na avaliação das cadeiras de ES ou SO, conforme a sua origem,
- Os grupos com melhor desempenho ficam habilitados a um *prize Money* de 150€ (1º lugar – 60€, 2º e 3º Lugar - 45€), atribuído com base na avaliação da prestação do grupo (ver secção de avaliação deste documento).

Qualquer esclarecimento adicional poderá ser endereçado ao responsável da experiência:

Eng.º Cláudio Sapateiro – claudio.sapateiro@estsetubal.ips.pt

<u>Dentro do bom espírito académico agradeço desde já a participação/colaboração de forma responsável e séria.</u>

<u>O responsável da experiência reserva-se ao direito de excluir das bonificações grupos cuja participação se revele desadequada.</u>

- 1 -

Objectivo

A experiência tem como objectivo estudar o desempenho de equipas de *help desk* na resolução de problemas de uma infra-estrutura informática. Os exercícios recaem sobre uma rede simulada que apresenta problemas nalguns dos seus equipamentos (computadores, routers e servidores), limitando a conectividade de alguns dos segmentos da rede. A equipa deverá coordenar as suas acções de modo a repor em funcionamento o maior número de elementos possível (idealmente todos!). No anexo 1 encontram-se as arquitecturas das redes que serão utilizadas nos quatro exercícios que constituem a experiência (2 de treino + 2 de avaliação). Os exercícios apresentam um nível de dificuldade moderado para alunos do curso de Informática.

Aplicações Utilizadas

Os exercícios suportam-se em 4 aplicações que estarão disponíveis no PC disponibilizado a cada elemento da equipa. Serão realizados 2 exercícios de ensaio para familiarização com o interface das aplicações (que não contarão para avaliação).

Aplicação 1: Simulador da rede

A Figura 1 apresenta um screenshot do simulador da rede.

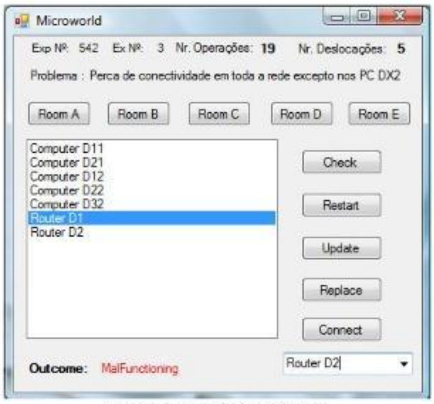


Figura 1 – Simulador de Rede

No topo do ecrã encontra-se a identificação da experiência e exercício (mera informação) e são disponibilizados (no topo superior direito) os contadores do nº de operações e deslocações já realizadas por toda a equipa.

A "label Problema" indica a descrição do problema que a rede apresenta e que a equipa deverá ser resolver. Os botões deste ecrã implementam o seguinte funcionamento:

- Ao pressionar o botão de uma sala, serão listados os equipamentos que estão situados nessa sala.
- Sobre cada equipamento é possível realizar o seguinte conjunto de operações:

1. **Check**: Verificação do estado do equipamento
2. **Restart**: Reinicia o equipamento e reporta o estado final do equipamento após a acção
3. **Update**: Realiza uma actualização do equipamento e reporta o estado final após a acção.
4. **Replace:** Substitui o equipamento e reporta o estado final após a acção.
5. **Connect:** Permite ligar o equipamento a outro (p.e. ligar o PC a outro router, ou um router a outro)

- 2 -

O estado do equipamento resultante das operações realizadas é comunicado na label "Outcome".

Notas:

1. Os equipamentos podem assumir 2 estados: ***Operating*** e ***Malfunctioning.***
2. As operações estão temporizadas da seguinte forma:
   a. Check: Gasta aprox. 0,5 seg.
   b. Restart: Gasta aprox. 1 seg.
   c. Update : Gasta aprox. 2 seg.
   d. Replace : Gasta aprox. 4 seg.
   e. Connect: Gasta aprox. 6 seg.
   f. Mudar de sala: Gasta aprox. 8 seg.
3. Como se depreende do ponto anterior é importante que se siga a ordem das operações para um melhor desempenho (além disso esta será a sequência lógica numa situação real)
4. A operação "Connect" só está disponível para alguns equipamentos.
5. As operações a realizar sobre os equipamentos dependem do perfil de competências do operador (e.g. nem todos poderão fazer um restart a um server). O perfil de competências de cada elemento da equipa será disponibilizado na aplicação de comunicação VoIP, descrita em seguida.

Aplicação 2: Comunicação Voice Over IP (VoIP)

Esta aplicação, ilustrada pela figura 2, permite a comunicação oral com os outros dois colegas de equipa, através da utilização de auscultadores e microfone (que serão fornecidos). Junto do nome de cada colega será indicado o seu perfil de competências, isto é, quais os equipamentos para os quais detém permissões para operar (se computadores e/ou routers e/ou servers). Este perfil será atribuído quando faz login na aplicação e segue o seguinte padrão:

Ordem dos logins	Numero de membro de equipa obtido	Competências
1º Elemento do grupo a efectuar login	Obtém o número 1 (TM1)	Servers, Routers e PCs
2º Elemento do grupo a efectuar login	Obtém o número 2 (TM2)	Routers e PCs
2º Elemento do grupo a efectuar login	Obtém o número 3 (TM3)	PCs

Legenda: TM – Team Member

Para efectuar a comunicação deverá clicar no botão do respectivo nome do colega com que pretende estabelecer comunicação.

Figura 2 – Módulo Comunicação VoIP

Sugestão:

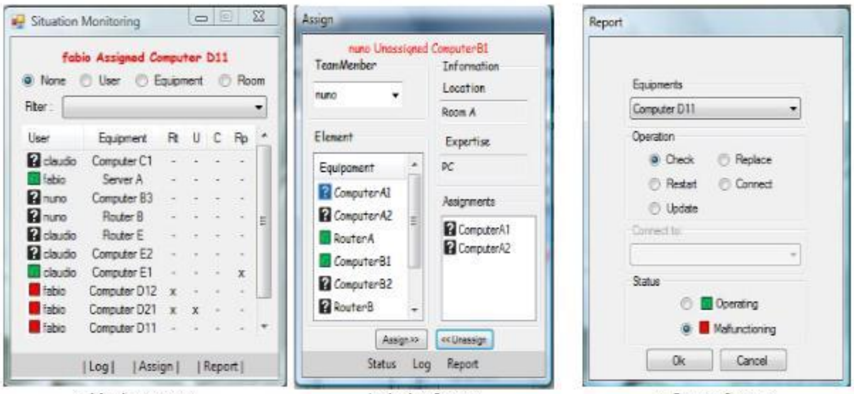
1. A duração das chamadas deverá ser curta, pois quando uma ligação está estabelecida entre dois colegas, o 3º colega da equipa fica impossibilitado de fazer contacto.

Aplicação 3: Simulador de dispositivo móvel (PDA)

Nalguns dos exercícios, será utilizada a aplicação 3 que simula um dispositivo PDA a executar uma aplicação que pretende auxiliar o trabalho da equipa na resolução do exercício. Esta poderá ser usada complementarmente à aplicação VoIP.

Esta aplicação consiste em 3 áreas principais (Figura 3 a., b. e c.):

a. ***Monitoring***: Lista dos equipamentos atribuídos aos membros da equipa, as operações (**Rt**-*Restart*, **U**-*Update*, **C**-*Connect*, **Rp**-*Replace*) já realizadas sobre os mesmos e estado actual destes (**verde**-*operating*, **vermelho**-*malfunctioning* ou **?**-*desconhecido*)
b. ***Assign***: Atribuir um equipamento (para verificação/operação) a um colega de equipa
c. ***Report***:
   I. Reportar o resultado de uma operação num equipamento resultante de uma atribuição previamente feita (p.e. por um colega): clicar no ecrã de monitorização sobre o equipamento sobre o qual vai reportar uma operação realizada; neste caso a combobox do ecrã de *Report* aparece com o equipamento automaticamente preenchido, só terá de preencher a acção realizada e o resultado da realização da mesma (o equipamentos ficou operacional ou não)
   II. Reportar o resultado de uma acção realizada voluntariamente (sem uma atribuição prévia): clicar no botão de *Report* (no rodapé do ecrã), neste caso deverá seleccionar o equipamento respectivo da combobox disponibilizada no ecrã de *Report* além da operação realizada e resultado (semelhante ao ponto anterior)
   III. Caso pretenda reportar sobre um equipamento que não lhe está atribuído (se estava atribuído a outro colega da equipa, mas acabou por ser operado por si) ser-lhe-á perguntado se quer efectuar um (Re)Assign do mesmo para si e caso o confirme poderá então reportar a operação realizada sobre acompanhada do respectivo resultado.



a. Monitor screen b. Assign Screen c. Report Screen

Figura 3 – Simulador de PDA

Como se pode observar o topo dos ecrãs pode conter uma linha escrita a vermelho com notificações relativamente à actividade que decorre na aplicação PDA. A cada click sobre a notificação será mostrada a próxima notificação até que não haja nenhuma actividade a reportar. A lista de todas as actividades na aplicação pode ser acedida através do botão de "log" (situado no rodapé do ecrã) (Figura 4).

- 4 -

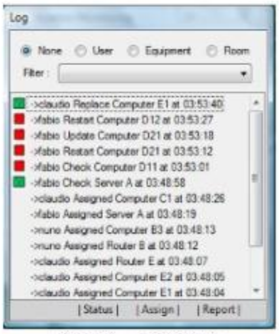


Figura 4 - Log Screen

Notas:

1. Como pode constatar nas Fig.3a e 4, o topo do ecrã tem um conjunto de radios e uma combobox que permitem filtrar as listagens presentes nesses ecrã no sentido de facilitar a consulta:
    a. A radio "User" disponibiliza na combobox os colegas de equipa, que uma vez seleccionado um, filtra a listagem de forma a que só sejam mostrados os registos relacionados com esse colega
    b. A radio "Equipment" disponibiliza na combobox os equipamentos da rede, que uma vez seleccionado um, filtra a listagem de forma a que só sejam mostrados os registos relacionados com esse equipamento
    c. A radio "Room" disponibiliza na combobox as salas da rede, que uma vez seleccionada uma, filtra a listagem de forma a que só sejam mostrados os registos relacionados com equipamentos dessa sala.
    d. A radio "None" lista todos os registos

Aplicação 4: Questionário

Periodicamente, no decorrer dos exercícios, todas as aplicações serão desactivadas e surgirá um questionário no ecrã. Cada elemento da equipa deve responder às questões com o conhecimento que detém relativamente ao estado da situação nesse preciso momento.

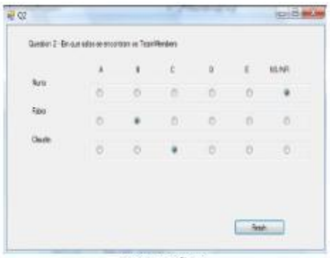
A figura 4 apresenta um screenshot ilustrativo do formato das questões. Estas são:

1. Qual o estado actual dos seguintes equipamentos?
2. Em que sala se encontram no momento os colegas de equipa?
3. Quais os equipamentos que estão a condicionar o bom funcionamento da rede?

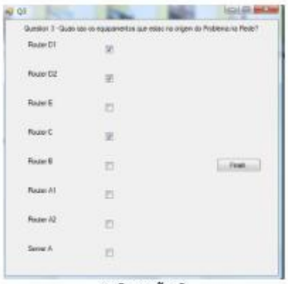
No caso de não saber responder a alguma das alíneas das questões colocadas, poderá deixar a opção default da resposta NS/NR.

- 5 -

a. Questão 1



b. Questão 2



c. Questão 3

Figura 4 - Questionários

Execução da Experiência

A experiência completa compreende as seguintes etapas:

1. Mini briefing para esclarecimento de dúvidas resultantes da leitura antecipada deste manual e da sessão de briefing prévia [máx. 5 min]
2. Resolução de um exercício de treino (com ou sem PDA – será definido no momento) [máx. 10 min]
3. Resolução de um exercício de avaliação (com ou sem PDA, igualando situação do ponto 2) [máx. 20 min]
4. Resolução de um exercício de treino (com ou sem PDA – ao contrário do ponto 2) [máx. 10 min]
5. Resolução de um exercício de avaliação (com ou sem PDA, igualando situação do ponto 6) [máx. 20 min]

Notas:

1. Ao realizar login na aplicação ser-lhe-á atribuído um Nº de membro da equipa (TM1, TM2 ou TM2). Este número ditará as respectivas competências de cada elemento da equipa como está descrito na secção deste documento relativa à aplicação VoIP.

- 6 -

Avaliação da Equipa

A participação da equipa nas experiências será avaliada segundo 3 dimensões detalhadas em baixo.

A pontuação de candidatura ao *Prize Money* é calculada com a seguinte fórmula:

Pontuação = 50% (AVG(scoreEx3,scoreEx4) + 50% debriefing

A bonificação na nota da cadeira de origem do aluno é calculada com a seguinte fórmula:

Bonificação = 50% pontuação + 50% MicroTeste

Todas as parcelas encontram-se detalhadas em baixo:

"Score" no exercício calculado em função de:

a. Eficiência = $\frac{\text{nº ideal de operações}}{\text{nº total de operações realizadas}}$

b. Eficácia = $\frac{\text{nº equipamentos operacionais}}{\text{nº equipamentos inicialmente inoperacionais}}$

c. Respostas aos questionários = $\frac{\text{nº respostas correctas}}{\text{nº de itens perguntados}}$

d. Deslocações efectuadas = $\frac{\text{nº ideal de deslocações}}{\text{nº deslocações efctuadas}}$

e. Tempo = $\frac{20-\text{tempo\_efectivamente\_gasto}}{20}$

score = a + b + c + d + e

Obs: as operações de "Check" não são contabilizadas.

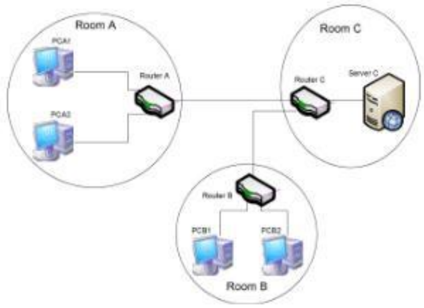
O resultado será afixado, assim que se atinja o número de 12 grupos/experiências realizadas (previsivelmente ate Março de 2011), em:

http://ltodi.est.ips.pt/csapateiro/ResultadosExp2.html

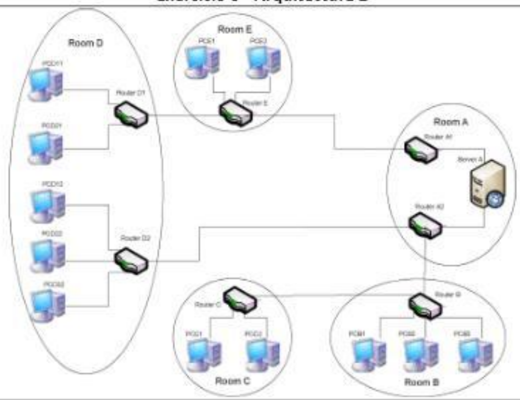
- 7 -

Anexo1 - Arquitecturas das Redes Simuladas

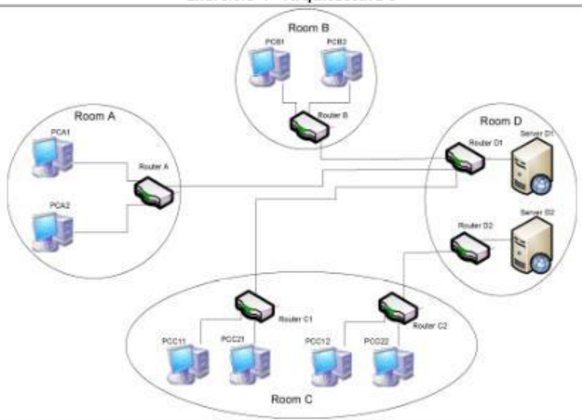
Exercício 1 e 2 (treino) – Arquitectura 1



Exercício 3 - Arquitectura 2



Exercício 4 - Arquitectura 3



- 8 -

Anexo 2 – Documento de Autorização

No contexto da minha participação na experiência "Estudo de Desempenho em Trabalho Colaborativo", da responsabilidade do Eng.º Cláudio Sapateiro, eu abaixo-assinado:

1. Declaro que tomei conhecimento e concordo com o funcionamento da experiência descrito através do documento que este anexo acompanha
2. Autorizo o uso e tratamento dos dados por mim gerados nesta experiência pelas pessoas responsáveis pela mesma;
3. Comprometo-me a ter uma postura respeitadora relativamente às pessoas envolvidas e à experiência em si;
4. Comprometo-me a não divulgar o conteúdo específico da experiência durante os dias em que esta estiver a decorrer (meses de Janeiro a Março de 2011).

Declaro ter conhecimento dos meus direitos:

a. À privacidade;
b. De poder desistir livremente da experiência.

Janeiro de 2011,

____

(o voluntário)

____

(o responsável da experiência)

Cláudio Sapateiro: claudio.sapateiro@estsetubal.ips.pt

Departamento de Sistemas e Informática da Escola Superior de Tecnologia de Setúbal

Instituto Politécnico de Setúbal

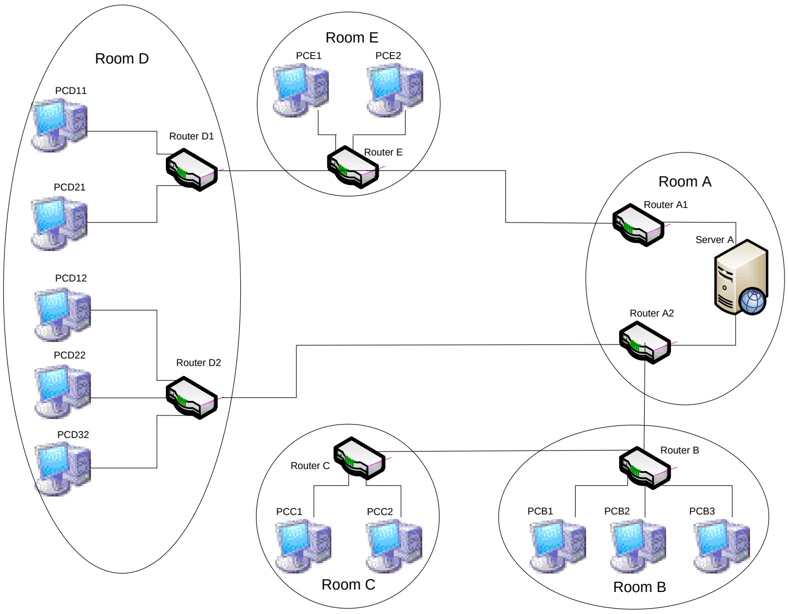
(Se o aluno desejar pode realizar uma cópia deste documento e solicitar ao responsável da experiência que a assine com vista a ficar na posse de uma cópia do documento)

- 9 -

Annex C – Second Round of Experiments Task Scenarios Characterization

C.1 Task scenario number five

Network architecture number two loaded on the Operational Work Environment Simulator (OWES).



Affordable connections between equipment on network architecture number two.

From Equipment (ID)	To Equipment (ID)
RA1 (34)	RA2 (35)
PCD21 (18)	RD2 (23)
PCD11 (17)	RD2 (23)
PDD21 (18)	RD1 (22)
PCD11 (17)	RD1 (22)
RD1 (22)	RD2 (23)

Initial conditions for the task scenario number 5 respecting the network architecture number two.

Equipment's Room	Equipment ID	Equipment Description	Equipment Label	Initial States Scenario #5
Room A	34	Router A1	RA1	*Malfunctioning*
	35	Router A2	RA2	*Operating*
	36	Server A	SA	*Operating*
Room B	30	Personal Computer B1	PCB1	*Malfunctioning*
	31	Personal Computer B2	PCB2	*Malfunctioning*
	32	Personal Computer B3	PCB3	*Malfunctioning*
	33	Router B	RB	*Malfunctioning*
Room C	27	Personal Computer C1	PCC1	*Malfunctioning*
	28	Personal Computer C2	PCC2	*Malfunctioning*
	29	Router C	RC	*Malfunctioning*
Room D	17	Personal Computer D11	PCD11	*Malfunctioning*
	18	Personal Computer D21	PCD21	*Malfunctioning*
	19	Personal Computer D12	PCD12	*Operating*
	20	Personal Computer D22	PCD22	*Operating*
	21	Personal Computer D32	PCD32	*Operating*
	22	Router D1	RD1	*Malfunctioning*
	23	Router D2	RD2	*Operating*
Room E	24	Personal Computer E1	PCE1	*Malfunctioning*
	25	Personal Computer E2	PCE2	*Malfunctioning*
	26	Router E	RE	*Malfunctioning*

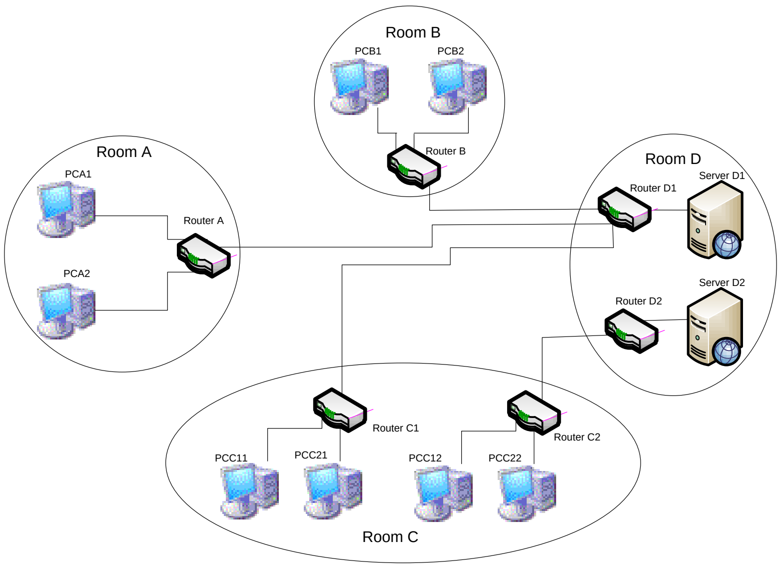
Collection of events that bound the task scenario number 5

Event ID	Targeted Equipment (ID)	Pre-conditions	Post-conditions
1	RB (33)	An *Update* action is carried out on RB.	Updates the states of equipment: PCB1, PCB2, PCB3 and RB to the *Operating* state
2	RC (29)	RB holds the *Operating* state and A *Restart* action is carried out on RC.	Updates the states of equipment: PCC1, PCC2 and RC to the *Operating* state
3	RA1 (34)	A *Connect* action is carried out over the equipment RA1 linking it to RA2	Updates the states of equipment: PCE1, PCE2, RE and RA1 to the *Operating* state
4	RD1 (22)	RA1 is connected to RA2 and A *Restart* action is carried out on RD1.	Updates the states of equipment: PCD11, PCD21 and RD1 to the *Operating* state

5	PCD11 (17)	A *Connect* action is carried out over PCD11 linking it to equipment RD2	Updates the state of PCD11 to the *Operating* state
6	PCD21 (18)	A *Connect* action is carried out over PCD21 linking it to equipment RD2	Updates the state of equipment PCD21 to the *Operating* state
7	RD1 (22)	A *Connect* action is carried out over the RD1 linking it to RD2	Updates the states of equipment: #PCD11, PCD21 and RD1 to the *Operating* state
8	PCD11 (17)	RD1 holds the *Operating* state and A *Connect* action is carried out over PCD11 linking it to equipment RD1	Updates the states of equipment PCD11 to the *Operating* state
9	PCD21 (18)	RD1 holds the *Operating* state and A *Connect* action is carried out over PCD21 linking it to equipment RD2	Updates the states of equipment PCD21 to the *Operating* state

C.2 Task scenario number six

Network architecture number three loaded on the Operational Work Environment Simulator (OWES).



Affordable connections between equipment on network architecture number three.

From Equipment (ID)	To Equipment (ID)
PCC11 (43)	RC2 (48)
PCC21 (44)	RC2 (48)
PCC11 (43)	RC1 (47)
PCC21 (44)	RC1 (47)

Initial conditions for the task scenario number 6 respecting the network architecture number three.

Equipment's Room	Equipment ID	Equipment Description	Equipment Label	Initial States Scenario #6
Room A	37	Personal Computer A1	PCA1	*Malfunctioning*
	38	Personal Computer A2	PCA2	*Malfunctioning*
	39	Router A	RA	*Malfunctioning*
Room B	40	Personal Computer B1	PCB1	*Malfunctioning*
	41	Personal Computer B2	PCB2	*Malfunctioning*
	42	Router B	RB	*Malfunctioning*
Room C	43	Personal Computer C11	PCC11	*Malfunctioning*
	44	Personal Computer C21	PCC21	*Malfunctioning*
	45	Personal Computer C12	PCC12	*Operating*
	46	Personal Computer C22	PCC22	*Operating*
	47	Router C1	RC1	*Malfunctioning*
	48	Router C2	RC2	*Operating*
Room D	49	Router D1	RD1	*Malfunctioning*
	50	Router D2	RD2	*Operating*
	51	Server D1	SD1	*Operating*
	52	Server D2	SD2	*Operating*

Collection of events that bound the task scenario number 6

Event ID	Targeted Equipment (ID)	Pre-conditions	Post-conditions
1	RD1 (49)	A *Replace* action is carried out over RD1	Updates the states of equipment: PCA1, RA and 49 to the *Operating* state
2	PC11 (43)	A *Connect* action is carried out over PC11 linking it to RC2	Updates the state of equipment PC11 to the *Operating* state

3	PC21 (44)	A *Connect* action is carried out over PC21 linking it to RC2	Updates the state of equipment PC21 to the *Operating* state
4	PCA2 (38)	RA holds the *Operating* state and A *Restart* action is carried out over PCA2	Updates the state of equipment PCA2 to the *Operating* state
5	RB (42)	RD1 holds the *Operating* state and An *Update* action is carried out over the equipment RB	Updates the state of equipment PCB1, PCB2 and RB to the *Operating* state
6	PC11 (43)	A *Connect* action is carried out over PC11 linking it to equipment RC1	Updates the state of equipment PC11 to the *Malfunctioning* state
7	PC21 (44)	A *Connect* action is carried out over PC21 linking it to RC1	Updates the state of equipment PC21 to the *Malfunctioning* state

References

Antunes, P. and C. J. Costa (2003). Perceived Value: A low-cost approach to evaluate meetingware. Lecture Notes in Computer Science. J. Favela and D. Decouchant, Springer-Verlag. **2806:** 109-125.

Antunes, P., V. Herskovic, et al. (2008). "Structuring dimensions for collaborative systems evaluation." ACM Comput. Surv. **44**(2): 1-28.

Antunes, P. and J. A. Pino (2010). A review of CRIWG research. Proceedings of the 16th international conference on Collaboration and technology. Maastricht, The Netherlands, Springer-Verlag**:** 1-15.

Araujo, R. M. d., F. M. Santoro, et al. (2004). "A conceptual framework for designing and conducting groupware evaluations." Int. J. Comput. Appl. Technol. **19**(3/4): 139-150.

Arthur, W., E. A. Day, et al. (2010). "The effect of distributed practice on immediate posttraining, and long-term performance on a complex command-and-control simulation task." Human Performance **23**(5): 428-445.

Artman, H. and C. Garbis (1998). Team communication and coordination as distributed cognition. Proceedings of 9th Conference of Cognitive Ergonomics : Cognition and cooperation**:** 151-156.

Bai, X., W.-T. Tsai, et al. (2002). Scenario-Based Modeling And Its Applications. Proceedings of the The Seventh IEEE International Workshop on Object-Oriented Real-Time Dependable Systems (WORDS 2002), IEEE Computer Society**:** 253.

Baker, K., S. Greenberg, et al. (2001). Heuristic Evaluation of Groupware Based on the Mechanics of Collaboration. Proceedings of the 8th IFIP international Conference on Engineering For Human-Computer interaction, Lecture Notes In Computer Science. **2254:** 123-140.

Baker, S. and E. Marshal (1988). Evaluating the man-machine interface - The search for data. Training, Human Decision Making and Control. J. Patrick and K. D. Duncan. Amsterdam, Elsevier**:** 79-92.

Banbury, S., S. Selcon, et al. (1998). "Being Certain About Uncertainty: How the Representation of System Reliability Affects Pilot Decision Making." Proceedings of the Human Factors and Ergonomics Society Annual Meeting **42**(1): 36-39.

Barret, R., E. Kandogan, et al. (2004). Field studies of computer system administrators: Analysis of system management tools and pratices. CSCW, Chicago, Illinois, USA.

Bartlet, F. (1932). Remembering: A study in experimental and social psychology. London, UK, Cambridge University Press.

Becker, J. and B. Niehaves (2007). "Epistemological perspectives on IS research: a framework for analysing and systematizing epistemological assumptions." Information Systems Journal **17**(2): 197-214.

Bedny, G. and D. Meister (1999). "Theory of activity and situation awareness." International Journal of Cognitive Ergonomics **3**(1): 63-72.

Benner, L. (1975). "D.E.C.I.D.E in Hazards Materials Emergencies." Fire Journal **69**(4).

Berndtsson, J. and M. Normark (1999). The coordinative functions of flight strips : air traffic control revisited. Proceedings of the international ACM SIGGROUP conference on Supporting group work, GROUP '99. New York, ACM**:** 101-110.

Berrouard, D., K. Cziner, et al. (2006). "Emergency Scenario User Perspective in Public Safety Communication Systems." ISCRAM.

Blandford, A. and B. L. W. Wong (2004). "Situation awareness in emergency medical dispatch." Int. J. Hum.-Comput. Stud. **61**(4): 421-452.

Blickensderfer, E., J. A. Cannon-Bowers, et al. (1998). Cross-training and team performance. Making decisions under stress: Implications for individual and team training. J. A. C.-B. E. Salas. Washington, DC, US, American Psychological Association**:** 299-311.

Bodker, S. (1989). A human activity approach to user interfaces. Human-Computer Iinteraction**:** 171-195.

Bolstad, C. A., H. M. Cuevas, et al. (2005). Modeling Shared Situation Awareness. 14th Conference on Behavior Representation In Modeling and Simulation (BRIMS), Los Angeles, CA.

Bolstad, C. A. and M. R. Endsley (1998). Information Dissonance, Shared mental Models and Shared Displays: An empirical Evaluation of Information Dominance Techniques. Interim Report, Wright Patterson AFB, OH: Air Force Research Laboratory.

Bolstad, C. A. and M. R. Endsley (1999). "Shared Mental Models and Shared Displays: An Empirical Evaluation of Team Performance." Proceedings of the Human Factors and Ergonomics Society Annual Meeting **43**(3): 213-217.

Bolstad, C. A. and M. R. Endsley (2000). The effect of task load and shared displays on team situation awareness. 14th Triennial Congress of the International Ergonomics Association and 44th Meeting of human Factors and Ergonomics Society.

Bowers, C. A., C. Braun, et al. (1994). Communication and Team Situational Awareness. Situational Awareness in Compex Systems. R. D. Gilson, D. J. Garland and J. M. Koonce, Embry-Riddle Aeronautical University Press**:** 305-311.

Boyd, J. R. (1996). The Essence of Winning and Losing.

Brehmer, B. (2005). "Micro-worlds and the circular relation between people and their environment." Theoretical Issues in Ergonomics Science **6**(1): 73-93.

Brehmer, B. and R. Allard (1990). Dynamic decision making. Effects of complexity and feedback delays. Distributed decision making: Cognitive models of cooperative work. J. Rasmussen, B. Brehmer and J. Leplat. Chichester, England, Wiley**:** 319-334.

Brehmer, B. and D. Dörner (1993). "Experiments with computer-simulated microworlds: Escaping both the narrow straits of the laboratory and the deep blue sea of the field study." Computers in Human Behavior **9**(2): 171-184.

Briggs, R., S. Qureshi, et al. (2004). Satisfaction Attainment Theory as a Model for Value Creation. The Thirty Seventh Annual Hawaii International Conference on Systems Sciences, IEEE Computer Society Press.

Bruinsma, G. and R. d. Hoog (2006). Exploring protocols for multidisciplinary disaster response using adaptative workflow simulation. ISCRAM.

Buchner, A. and J. Funke (1993). "Finite state automata: Dynamic task environments in problem solving research." The Quarterly Journal of Experimental Psychology **46A**(1): 83-118.

Buckley, W. F. (1967). Sociology and modern systems theory, Prentice-Hall.

Burke, C. S., K. C. Stagl, et al. (2006). "Understanding team adapatation: A conceptual analysis and model." Journal of Applied Psychology **91**(6): 1189-1207.

Burrell, G. and G. Morgan (1979). Sociological Paradigms and Organizational Analysis: Elements of the Sociology of Corporate Life, Ashgate Publishing.

Bygstad, B. (2005). "Managing the dynamics of mutual adaptation of technology and organisation in information systems development projects." Software Process: Improvement and Practice **10**(3): 341-353.

Cacciabue, P. C. and E. Hollnagel (2005). Simulation of cognition: applications, Hillsdale, NJ: LEA.

Camerer, C. and M. Weber (1992). "Recent developments in modeling preferences: Uncertainty and ambiguity." Journal of Risk and Uncertainty **5**(4): 325-370.

Cannon-Bowers, J. A. and E. Salas (1997). Teamwork competencies: the interaction of team member knowledge skills and attitudes. Workforce readiness: Competencies and assessment. O. F. O'Neil, Hillsdale, NJ: Erlbaum**:** 161-174.

Cannon-Bowers, J. A., E. Salas, et al. (1993). "Shared mental models in expert team decision making." Current issues in individual and group decision making: 221-246.

Cannon-Bowers, J. A., E. Salas, et al. (1996). "Establishing the Boundaries of a Paradigm for Decision-Making Research." Human Factors: The Journal of the Human Factors and Ergonomics Society **38**(2): 193-205.

Carlsson, S. (2006). Towards an information systems design research framework: A critical realist perspective. In Proceedings of the First International Conference on Design Science Research in Information Systems and Technology, Claremont, California.

Carroll, J. M., D. C. Neale, et al. (2003). "Notification and awareness: synchronizing task-oriented collaborative activity." International Journal of Human-Computer Studies **58**(5): 605-632.

Carter, S., J. Mankoff, et al. (2007). Momento: support for situated ubicomp experimentation. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. San Jose, California, USA, ACM: 125-134.

Cecez-Kecmanovic, D. (2001). Doing critical IS research: the question of methodology. Qualitative Research in IS: issues and trends. E. Trauth, Hershey: Idea Group Publishing: 141-162.

Chapman, T., T. Nettelbeck, et al. (2006). "Investigating the construct validity associated with microworld research: A comparison of performance under different management structures across expert and non-expert naturalistic decision-making groups." Australian Journal of Psychology **58**(1): 40-47.

Chen, W. and R. Hirschheim (2004). "A paradigmatic and methodological examination of information systems research from 1991 to 2001." Information Systems Journal **14**(3): 197-235.

Chua, W. F. (1986). "Radical developments in accounting thought." The Accounting Review **61**(4): 601-632.

Citera, M., M. D. McNeese, et al. (1995). "Fitting information systems to collaborating design teams." Journal of the American Society for Information Science **46**(7): 551-559.

Cockton, G. and A. Woolrych (2002). "Sale must end: should discount methods be cleared off HCI's shelves?" Interactions **9**(5): 13-18.

Coiera, E., R. Jayasuriya, et al. (2002). "Communication loads on clinical staff in the emergency department." Medical Journal of Australia **176**.

Cooke, N. J. and J. C. Gorman (2006). Assessment of Team Cognition. International Encyclopedia of Ergonomics and Human Factors, Second Edition - 3 Volume Set, CRC Press.

Cooke, N. J., J. C. Gorman, et al. (2008). Team Cognition. Handbook of Applied Cognition, John Wiley & Sons Ltd: 239-268.

Cooke, N. J., P. A. Kiekel, et al. (2003). "Measuring Team Knowledge: A Window to the Cognitive Underpinnings of Team Performance." Group Dynamics: Theory Research and Practice **7**: 179-199.

Cooke, N. J., E. Salas, et al. (2000). "Measuring Team Knowledge." Human Factors **42**: 151-173.

Cooke, N. J., E. Salas, et al. (2004). Advances in Measuring Team Cognition. Team Cognition: Understanding the Factors that Drive Process and Performance. E. Salas and S. M. Fiore. Washington, DC, American Psychological Association.

Crandall, B., G. Klein, et al. (2006). Thinking About Cognition. Working Minds. London, England, Cambridge: MIT Press**:** 131-148.

Crandall, B., G. A. Klein, et al. (2006). Working Minds: A Practitioner's Guide to Cognitive Task Analysis, Mit Press.

Creswell, J. (2002). Research Design: Qualitative, Quantitative, and Mixed Methods Approaches (2nd Edition), Sage Publications, Inc.

Cuevas, H. M., S. M. Fiore, et al. (2007). "Augmenting Team Cognition in Human-Automation Teams Performing in Complex Operational Enviornments." Aviation, Space, and Environmental Medicine **78**(5).

Damianos, L., L. Hirschman, et al. (1999). "Evaluation for collaborative systems." ACM Computing Surveys **31**(2es): 15.

Damianos, L., L. Hirschman, et al. (1999). "Evaluation for collaborative systems." ACM Comput. Surv. **31**(2es): 15.

Davis, A., O. Dieste, et al. (2006). Effectiveness of Requirements Elicitation Techniques: Empirical Results Derived from a Systematic Review. Requirements Engineering, 14th IEEE International Conference.

DeKeyser, V. and D. D. Woods (1990). Fixation errors: Failures to revise situation assessment in dynamic and risky situations. Systems Reliability Assessment. A. G. Colombo and A. S. Bustamante, Kluwer.

Dessalles, J. L., J. Ferber, et al. (2008). Emergence in Agent-Based Computational Social Science: Conceptual, Formal, and Diagrammatic Analysis. Intelligent Complex Adaptive Systems, IGI Global**:** 255-299.

Dorner, D. (1997). The Logic of Failure: Recognizing and Avoiding Error in Complex Situations, Perseus Books Group.

Dorner, D. and E. Pfeifer (1993). "Strategic thinking and stress." Ergonomics **36**: 1345-1360.

Dourish, P. and V. Bellotti (1992). Awareness and coordination in shared workspaces. Proceedings of the 1992 ACM conference on Computer-supported cooperative work. Toronto, Ontario, Canada, ACM**:** 107-114.

Drabek, T. and D. McEntire (2003). "Emergent phenomena and the sociology of disaster: lessons, trends and oportunities from research literature." Disaster Prevention and Management **12**: 97-112.

Driskell, J. E. and E. Salas (1991). "Group decision making under stress." Journal of Applied Psychology **76**(3): 473-478.

Dulac, N. and N. Leveson (2004). An Approach to Design for Safety in Complex Systems. International Symposium on Systems Engineering. Toulouse, France.

Dulac, N. and N. Leveson (2009). "Incorporating safety risk in early system architecture trade studies." Journal of Spacecraft and Rockets **46**(2): 430-437.

Duncan, P. C., W. B. Rouse, et al. (1999). Training teams working in complex systems: A mental model-based approach. In Human/Technology Interaction in Complex Systems. W. B. Rouse. Greenwich, JAI Press. **8:** 173-231.

Dunn, J. C., S. Lewandowsky, et al. (2002). "Dynamics of communication in emergency management." Applied Cognitive Psychology **16**(6): 719-737.

Durso, F. (1999). Situation Awareness As a Predictor of Performance in En Route Air Traffic Controllers, Oklahoma univ norman Department of psychology.

Dyer, J. L., Ed. (1984). Team Research and Team Training: A state of the art review. Human Factors Review, Santa Monica: Human Factors Society.

Dynes, R. and E. L. Quarantelli (1997). Organizational Communications and Decision Making in Crises, Disaster Research Center Report. Ohio State University.

Eby, L. T. and G. H. Dobbins (1997). "Collectivistic orientation in teams: an individual and group-level analysis." Journal of Organizational Behavior **18**(3): 275-295.

Eisenberger, R. (1993). "Decision Making in Action: Models and Methods." Journal of Behavioral Decision Making **8**: 218-229.

Endsley, M. (1995). "Toward a theory of situation awareness in dynamic systems." Human Factors **37**(1): 32-64.

Endsley, M. and D. Garland (2000). Situation Awareness Analysis and Measurement. Boca Raton, FL, CRC Press.

Endsley, M. R. (1988). Situation awareness global assessment technique (SAGAT). Proceedings of the IEEE Aerospace and Electronics Conference, NAECON 1988. .

Endsley, M. R. (1993). "A Survey of Situation Awareness Requirements in Air-to-Air Combat Fighters." The International Journal of Aviation Psychology **3**(2): 157-168.

Endsley, M. R., C. A. Bolstad, et al. (2003). Situation Awareness Oriented Design: From User's Cognitive Requirments to Creating Effective Supporting Technologies. 47th Human Factors and Ergonomics, Denver, Colorado.

Endsley, M. R., B. Bolté, et al. (2003). Alarms, Diagnosis and SA. Designing for Situation Awareness. New York, Taylor and Francis**:** 149-171.

Endsley, M. R., B. bolté, et al. (2003). Confidence and Uncertainty in SA and Decision Making. Designing for Situation Awareness. New Tork, Taylor and Francis**:** 116-131.

Endsley, M. R., B. Bolté, et al. (2003). Dealing with Complexity. Designining For Situation Awareness. M. R. Endsley, B. Bolté and D. G. Jones. New York, Taylor and Francis**:** 133-148.

Endsley, M. R., B. Bolté, et al. (2003). Determining SA Requirements. Designing for Situation Awareness:An approach to user-centered design. M. R. Endsley, B. Bolté and D. G. Jones, Taylor & Francis**:** 65-82.

Endsley, M. R., B. Bolté, et al. (2003). Evaluating Design Concepts for SA. Designing for Situation Awareness: An Approach to User-Centered Design. M. R. Endsley, B. Bolté and D. G. Jones, Taylor and Francis**:** 223-248.

Endsley, M. R., B. Bolté, et al. (2003). What is Situation Awareness. Designinig for Situation Awareness, Taylor and Francis**:** 13-30.

Endsley, M. R. and E. O. Kiris (1995). "The Out-of-the-Loop Performance Problem and Level of Control in Automation." Human Factors: The Journal of the Human Factors and Ergonomics Society **37**(2): 381-394.

Endsley, M. R. and M. M. Robertson (2000). "Situation awareness in aircraft teams." International Journal of Industrial Ergonomics **26**: 310-325.

Endsley, M. R., S. J. Selcon, et al. (1998). A comparative analysis os SAGAT and SART for evaluation of Situation Awareness. 42nd Annual Meeting of the Human Factors & Ergonomics Society, Chicago.

Entin, E. E. and D. Serfaty (1999). "Adaptive Team Coordination." Human Factors: The Journal of the Human Factors and Ergonomics Society **41**(2): 312-325.

Fallman, D. (2003). Enabling physical collaboration in industrial settings by designing for embodied interaction. Proceedings of the Latin American conference on Human-computer interaction. Rio de Janeiro, Brazil, ACM**:** 41-51.

Ferry, T. S. (1988). Modern Accident Investigation and Analysis, Wiley.

Fiore, S. M., E. Salas, et al. (2001). Group dynamics and shared mental model development. How People Evaluate Others in Organizations: Person perception and interpersonal judgment in industrial/organizational psychology. M. London. Mahwah,NJ, Erlbaum**:** 309-336.

Fiore, S. M., E. Salas, et al. (2003). "Distributed coordination space: toward a theory of distributed team process and performance." Theoretical Issues in Ergonomics Science **4**: 340-364.

Fiore, S. M., K. A. Smith-Jentsch, et al. (2010). "Towards an understanding of macrocognition in teams: developing and defining complex collaborative processes and products." Theoretical Issues in Ergonomics Science **11**(4): 250 - 271.

Fiore, S. M., C. A. P. Smith, et al. (2008). Macrocognition Research Challenges and Opportunities on the Road Ahead. Macrocognition in Teams: Theories and Methodologies. S. M. Fiore, C. A. P. Smith, M. P. Letsky and N. W. Warner, Ashgate**:** 397-416.

Fjermestad, J. and S. Hiltz (1999). "An assessment of group support systems experimental research: Methodology and results." Journal of Management Information Systems **15**(3): 7-149.

Flor, N. and E. Hutchins (1991). Analysing distributed cognition in software teams: A case study of team programming during perfective software maintenance. 4th Annual Workshop on Empirical Studies of Programmers. J. Koenemann-Belliveau, Ablex Publishing**:** 36-59.

Froehlich, J., M. Y. Chen, et al. (2007). MyExperience: a system for in situ tracing and capturing of user feedback on mobile phones. Proceedings of the 5th international conference on Mobile systems, applications and services. San Juan, Puerto Rico, ACM**:** 57-70.

Galliers, R. D. (1992). Information Systems Research: Issues, Methods and Practical Guidelines. Oxford, Blackwell Scientific Publications.

Gasson, S. (1999). "A social action model of situated information systems design." The Data Base for Advances in Information Systems **30**(2).

Gasson, S. (2004). A Framework For Behavioral Studies of Social Cognition In Information Systems. ISOneWorld. Las Vegas.

Gauducheau, N., E. Soulier, et al. (2005). Design and evaluation of activity model-based groupware: methedological issues. 14th International Workshops on Enabling TEchnologies: Infrastructure for collaborative enterprise (WETICE), IEEE Computer Society.

Gibson, J. J. (1966). The senses considered as perceptual systems, Boston: Houhton Mifflin.

Gonzalez, C., P. Vanyukov, et al. (2005). "The use of microworlds to study dynamic decision making." Computers in Human Behavior **21**(2): 273-286.

Gorman, J. C., N. J. Cooke, et al. (2005). "Coordinated Awareness of Situation by Teams (CAST): Measuring Team Situation Awareness of a Communication Glitch." Proceedings of the Human Factors and Ergonomics Society Annual Meeting **49**(3): 274-277.

Grant, T. and B. Kooter (2005). Comparing OODA & other models as Operational View C2 Architecture. Proceedings of the 10th International Command and Control Research and Technology Symposium.

Gray, W. (2002). "Simulated task environments: The Role of high-fidelity simulations, scaled worlds, synthetic environments, and laboratory tasks in basic and applied cognitive research." Cognitive Science Quarterly **205-207**.

Gray, W. D. (2002). "Simulated task environments: The role of high-fidelity simulations, scaled worlds, synthetic environments, and laboratory tasks in basic and applied cognitive research." Cognitive Science Quarterly **2**(2): 205-227.

Grudin, J. (1988). Why CSCW applications fail: problems in the design and evaluation of organizational interfaces. Proceedings of the 1988 ACM conference on Computer-supported cooperative work, Portland, Oregon, ACM.

Gualtieri, J., J. Fowlkes, et al. (1996). "Measuring individual and team knowledge structures for use in training." Training Research Journal **2**: 117-141.

Gutwin, C. and S. Greenberg (2000). The mechanics of collaboration: Developing low cost usability evaluation methods for shared workspaces. WETICE ’00, IEEE Computer Society Press.

Hale, J. (1997). "A Layered Communication Architecture for the Support of Crisis Response." Journal of Management Information Systems **14**(1): 235-255.

Hall, A. D. (1962). A methodology for systems engineering, Van Nostrand.

Hardeman, F., N. Pauwels, et al. (1998). The role of experts in decision making upon urgent countermeasures in nuclear emergency situations. European Conference on Risk analysis.

Haugeland, J. (1985). Artificial Intelligence: the very idea, Cambridge, MA:MIT Press.

Hayes, J. (2006). Safety decision making in high hazard organizations at the production/maintenance interface- A literature review.

Hazlehurst, B., C. K. McMullen, et al. (2007). "Distributed cognition in the heart room: How situation awareness arises from coordinated communications during cardiac surgery." Journal of Biomedical Informatics **40**(5): 539-551.

Heath, C. and P. Luff (1992). "Collaboration and controlCrisis management and multimedia technology in London Underground Line Control Rooms." Computer Supported Cooperative Work (CSCW) **1**(1-2): 69-94.

Heath, C. and A. Tversky (1991). "Preference and belief: Ambiguity and competence in choice under uncertainty." Journal of Risk and Uncertainty **4**(1): 5-28.

Heinrich, H. (1931). Industrial accident prevention. New York, McGraw-Hill.

Herskovic, V., J. A. Pino, et al. (2007). Evaluation Methods for Groupware Systems. CRIWG.

Hevner, A. R. (2007). "A Three Cycle View of Design Science Research." Scandinavian Journal of Information Systems **19**(2): 87-92.

Hevner, A. R., S. T. March, et al. (2004). "Design Science in Information Systems Research." MIS Quarterly **28**(1): 75-105.

Hill, J. and C. Gutwin (2003). Awareness Support in a Groupware Widget Toolkit. Proceedings of the 2003 international ACM SIGGROUP conference on Supporting Group Work, Sanibel Island, Florida, ACM Press.

Hinsz, V. B., R. S. Tindale, et al. (1997). "The emerging conceptualization of groups as information processors." Psychol Bull **121**(1): 43-64.

Hoermann, H. J. (1995). FOR-DEC. A prescriptive model for aeronautical decision-making. 21st Conference of the European Association for Aviation Psychology, Aldershot, Hampshire, UK.

Hoffman, R. R. and M. D. McNeese (2009). "A History for Macrocognition." Journal of Cognitive Engineering and Decision Making. **3**(2): 97-110.

Hoffman, R. R. and L. G. Militello (2008). Perspectives on Cognitive Task Analysis: Historical Origins and Modern Communities of Practice. New York, CRC Press.

Hollnagel, E. (1998). Cognitive Reliability and Error Analysis Method (CREAM), Elsevier Science.

Hollnagel, E. (2004). Barriers And Accident Prevention, Ashgate.

Hollnagel, E. (2005). Learning from failures: A joint cognitive systems prespective. Evaluation of Human Work. J. R. Wilson and N. Corlett, Taylor & Francis: 901-916.

Hollnagel, E., S. Pruchnicki, et al. (2008). Analysis of comair flight 5191 with the functional resonance accident model. 8th International Symposium of the Australian Aviation Psychology Association. Sidney, Australia

Hollnagel, E. and D. D. Woods (1999). "Cognitive Systems Engineering: New wine in new bottles." International Journal of Human-Computer Studies **51**(2): 339-356.

Hollnagel, E. and D. D. Woods (2005). Coping with Complexity. Joint cognitive systems: Foundations of cognitive systems engineering. Boca Raton, Florida, Taylor & Francis.

Hollnagel, E. and D. D. Woods (2005). Joint Cognitive Systems: An introduction to cognitive systems engineering, Ashgate.

Hopkins, A. (2000). Lessons from Longford: the Esso Gas Plant Explosion. Sidney, CCH Australia.

Horseley, J. and R. T. Barker (2002). "Toward a synthesis Model for Crisis Communication in Public Scetor: An initial investigation." Journal of Business and Technical Communication **16**(4): 406-440.

Hourizi, R. and P. Johnson (2003). "Towards an explanatory, predictive account of awareness." Computers & Graphics **27**(6): 859-872.

Howell, D. C. (2009). Statistical Methods for Psychology, Cengage Learning.

Howie, D. E. and K. J. Vicente (1998). "Measures of operator performance in complex, dynamic microworlds: advancing the state of the art." Ergonomics **41**(4): 485-500.

Hughes, J., V. King, et al. (1994). Moving out from the control room: ethnography in system design. Proceedings of the 1994 ACM conference on Computer supported cooperative work. Chapel Hill, North Carolina, United States, ACM**:** 429-439.

Hutchins, E. (1991). The social organization of distributed cognition, Washington, DC: American Psychological Association.

Hutchins, E. (1994). Cognition in the Wild, Cambidge: MIT Press.

Iivari, J. (2007). "A Paradigmatic Analysis of Information Systems As a Design Science." Scandinavian Journal of Information Systems **19**(2).

Janssen, J., G. Erkens, et al. (2007). "Visualization of participation: Does it contribute to successful computer-supported collaborative learning?" Comput. Educ. **49**(4): 1037-1065.

Jay F. Nunamaker, J., M. Chen, et al. (1990). "Systems development in information systems research." Journal of Management Information Systems **7**(3): 89-106.

Jhoansson, B., J. Trnka, et al. (2007). The effect of geographical information systems on a collaborative C2 task. 4th International Conference on Information Systems for Crisis Response and Management, Delft, Netherlands.

Jiang, X., J. I. Hong, et al. (2004). Ubiquitous computing for firefighters: Field studies and prototypes of large displays for incident comand. Conference on Human Factors in Computing Systems, ACM**:** 679-686.

Johnson-Laird, P. N. (1983). Mental Models, Cambridge University Press.

Johnson, B. and A. Onwuegbuzie (2004). "Mixed Methods Research: A Research Paradigm Whose Time Has Come." Journal of Educational Researcher **33**(7): 14-26.

Johnson, C. W. and C. M. Holloway (2003). "The ESA/NASA SOHO mission interruption: using the STAMP accident analysis technique for a software related 'mishap'." Software: Practice and Experience **33**(12): 1177-1198.

Johnston, J. H., J. A. Cannon-Bowers, et al. (1997). How can you turn a team of experts into an expert team? Emerging training strategies. Naturalistic Decision Making. C. E. Zsambok and G. Klein. Mahwah, NJ, Lawrence Erlbaum Associates.

Jones, D. G. and M. R. Endsley (1996). "Sources of situation awareness errors in aviation." Aviation, Space, and Environmental Medicine **67**(6): 507-512.

Jones, D. G. and D. B. Kaber (2004). Situation awareness measurement and the situation awareness global assessment technique. Handbook of Human Factors and Ergonomics Methods. N. Stanton, Hedge, Hendrick, K. Brookhuis and E. Salas. Boca Raton, USA, CRC Press**:** 42.41-42.47.

Kaber, D. B., C. M. Perry, et al. (2006). "Situation awareness implications of adaptive automation for information processing in an air traffic control-related task." International Journal of Industrial Ergonomics **36**(5): 447-462.

Kahneman, D., P. Slovic, et al. (1982). Judgment under Uncertainty: Heuristics and Biases, Cambridge University Press.

Kanno, T. and K. Futura (2006). "Resilience of emergency response systems."

Kasl, E., V. J. Marsick, et al. (1997). "Teams as Learners: A Research-Based Model of Team Learning." The Journal of Applied Behavioral Science **33**(2): 227-246.

Kassner, A., M. Baumann, et al. (2011). A Hierarchical Task Analysis of Merging onto a Freeway—Comparison of Driver's and Driver Model's Task Representation. Human Modelling in Assisted Transportation. P. C. Cacciabue, M. Hjälmdahl, A. Luedtke and C. Riccioli, Springer Milan**:** 291-298.

Kauffman, S. A. (1995). At Home in the Universe: The Search for Laws of Self-Organization and Complexity, Viking.

Kelly, C. (1999). "Simplifying Disasters: Developing a model for complex non-linear events." Australian Journal of Emergency Management **14**(1): 25-27.

Kennedy, D. M. and S. A. McComb (2010). "Merging internal and external processes: examining the mental model convergence process through team communication." Theoretical Issues in Ergonomics Science **11**(4): 340 - 358.

Keys, J. B., R. M. Fulmer, et al. (1996). "Microworlds and simuworlds: Practice fields for the learning organization." Organizational Dynamics **24**(4): 36-49.

Khoshafian, S. and M. Buckiewicz (1995). Introduction to groupware, workflow, and workgroup computing, John Wiley and Sons, Inc.

Kiekel, P. A., N. J. Cooke, et al. (2001). Automating Measurement of Team Cognition through Analysis of Communication Data, Lawrence Erlbaum Associates.

Kjeldskov, J. and C. Graham (2003). A Review of Mobile HCI Research Methods. Human-Computer Interaction with Mobile Devices and Services. L. Chittaro, Springer Berlin Heidelberg. **2795:** 317-335.

Klein, G. (1998). How people make decisions. Cambridge, Massachusetts, MIT Press.

Klein, G. (2000). Cognitive task analysis of teams. Cognitive Task Analysis. Schraagen, Chipman and Shalin, Mahwah, NJ: Lawrence Erlbaum: 417-430.

Klein, G. A., C. E. Zsambock, et al. (1993). "Team decision trainning: Five myths and a model." Military Review: 36-42.

Klein, G. L., L. Adelman, et al. (2008). Enabling Collaboration: Realizing the Collaborative Potential of Network-Enabled Command. Battle of Cognition: The Future Information-Rich Warfare and the Mind of the Commander. A. Kott, Greenwood Publishing Group: 167-193.

Klimoski, R. and S. Mohammed (1994). "Team Mental Model: Construct or Metaphor?" Journal of Management **20**(2): 403-437.

Kozlowski, S. W. J., R. P. DeShon, et al. (2004). A Psychological Fidelity Approach to Simulation-Based Training: Theory, Research, and Principles. Scaled Worlds: Development, Validation, and Applications. Aldershot, England, Ashgate: 75-99.

Kraiger, K. and L. H. Wenzel (1997). Conceptual Development and Emprical Evaluation of Measures of Shared Mental Models as Indicators of Team Effectivness. Team Performance Assessment and Measurement: Theory, Methods, and Applications. M. T. Brannick, E. Salas and C. W. Prince, Taylor & Francis: 63-84.

Kreijns, K., P. A. Kirschner, et al. (2003). "Identifying the pitfalls for social interaction in computer-supported collaborative learning environments: a review of the research." Computers in Human Behavior **19**(3): 335-353.

Landgren, J. (2005). "Supporting fire crew sensemaking enroute to incidents." International Journal of Emergency Management **2**(3): 176-188.

Landgren, J. (2006). Making Action Visible in Time-critical Work. SIGCHI Conference on Human Factors in Computing Systems, Montréal, Quebec, Canada.

Landgren, J. (2007). Designing Information Technology for Emergency Response. Department of applied Information Technology. Goteborg, Goteborg University, Sweeden. **PhD**.

Langan-Fox, J., S. Code, et al. (2000). "Team Mental Models: Techniques, Methods, and Analytic Approaches." Human Factors: The Journal of the Human Factors and Ergonomics Society **42**(2): 242-271.

LaPorte, T. R. (2006). "High Reliability Organizations: Unlikely, demanding, and at risk." Journal of Contingencies and Crisis Management **63**(4): 60-71.

Lave, J. (1988). Cognition in Pratice, Cambridge: Cambridge University Press.

Lee, M.-C. (2004). "A Testing Framework Based on Finite Automata for Object-Oriented Software Specification." Journal of Information Technology Theory and Society(1): 59-88.

Leontiev, A. N. (1974). "The Problem of Activity in Psychology." Soviet Psychology **13**(2): 4--33.

Letsky, M. P. (2008). Macrocognition in Teams: Theories and Methodologies, Ashgate.

Leveson, N. (1995). SafeWare: system safety and computers, Addison-Wesley.

Leveson, N. (2004). "A New Accident Model for Engineering Safer Systems." Safety science **42**(4): 237-270.

Leveson, N. (2008). System Safety Engineering: Back To The Future.

Leveson, N., M. Daouk, et al. (2004). "Aplying STAMP in accident analysis." MIT.

Leveson, N., N. Dulac, et al. (2009). Moving Beyond Normal Accidents and High Reliability Organizations: A Systems Approach to Safety in Complex Systems.

Leveson, N., N. Dulac, et al. (2009). "Moving Beyond Normal Accidents and High Reliability Organizations: A Systems Approach to Safety in Complex Systems." Organization Studies **30**(2-3): 227-249.

Leveson, N., N. Dulac, et al. (2006). Engineering Resilience into Safety Critical Systems. Resilience Engineering: Concepts and Percepts. E. Hollnagel, D. D. Woods and N. Leveson, Ashgate**:** 95-123.

Leveson, N. G., Ed. (1995). Safeware: System Safety and Computers, Boston: Addison-Wesley.

Levesque, L. L., J. M. Wilson, et al. (2001). "Cognitive divergence and shared mental models in software development project teams." Journal of Organizational Behavior **22**(2): 135-144.

Lloyd, W. J., M. B. Rosson, et al. (2002). Effectiveness of elicitation techniques in distributed requirements engineering. IEEE Joint International Conference on Requirements Engineering

March, S. T. and G. F. Smith (1995). "Design and natural science research on information technology." Decis. Support Syst. **15**(4): 251-266.

Markus, M. L., A. Majchrzak, et al. (2002). "A design theory for systems that support emergent knowledge processes." MIS Quaterly **26**(3): 179-212.

Mathieu, J. E., T. S. Heffner, et al. (2000). "The Influence of Shared Mental Models on Team Process and Performance." Journal of Applied Psychology **85**(2): 273-283.

Matthews, M. D., R. J. Pleban, et al. (2000). Measures of infantry situation awareness for a virtual MOUT environment. Proc. of the Human Performance, Situation Awareness and Automation: User Centered Design for the New Millenium Conference.

Maurino, D. E., J. Reasonson, et al. (1995). Beyond Aviation Human Factors: Safety in High Technology Systems. Farnham, Ashgate Publishing Limited.

McGrath, J. (1984). Groups: Interaction and performance. Englewood Cliffs, NJ, Prentice-Hall.

McGrath, J. (1995). Methodology matters: doing research in the behavioral and social sciences. Human-computer interaction: toward the year 2000. San Francisco, CA, Morgan Kaufmann**:** 152-169.

McGrath, J. E. (1976). Stress and behavior in organization. Chicago, Rand McNally.

McManus, S., E. Seville, et al. (2007). Resilience Management: A framework for Assessing and Improving the resilience of organizations.

McNeese, M. D. (1986). "Humane Intelligence: a human factors perspective for developing intelligent cockpits." IEEE Aerospace and Electronic Systems **1**(9): 6-12.

Mendonça, D. and W. A. Wallace (2007). "A Cognitive model of Improvisation in Emergency Management " IEEE Transactions on Systems, Man And Cybernetics; PartA: Systems And Humans **37**(4): 547-561.

Milis, K. and B. V. d. Walle (2007). IT for Corporate Crisis Management: Findings from a Survey in 6 different Industries on Management Attention, Intention and Actual Use. ISCRAM.

Muhren, W. J. and B. V. d. Walle (2009). Sensemaking and Information Management in Humanitarian Disaster Response: Observations from te TRIPLEX Exercise. International Conference on Information Systems for Crisis Response and Managemet (ISCRAM), Gothenburg, Sweden.

Nandhakumar, J. and M. Jones (1997). "Too close for comfort? Distance and engagement in interpretive information systems research." Information Systems Journal **7**(2): 109-131.

Nasghar, S., D. Alahakoon, et al. (2005). "A comprehensive Conceptual Model for Disaster Management." Journal of Humanitarian Assistance.

Neale, D., J. Carroll, et al. (2004). Evaluating computer-supported cooperative work: models and frameworks. Proceedings of the 2004 ACM conference on Computer supported cooperative work, Chicago, Illinois, ACM Press.

Neale, D. C., J. M. Carroll, et al. (2004). Evaluating Computer-Supported cooperative work: Models and frameworks. CSCW.

Neisser, U. (1976). Cognition and Reality: Principles and Implications of Cognitive Psychology, W H Freeman.

Neuwirth, C., J. Morris, et al. (1998). Envisioning communication: task-tailorable representations of communication in asynchronous work. CSCW '98: Proceedings of the 1998 ACM conference on Computer supported cooperative work, ACM Press.

Newell, A. and H. Simon (1972). Human Problem Solving, Prentice Hall.

Niehaves, B. (2007). "On epistemological pluralism in design science." Scandinavian Journal of Information Systems **19**(2): 93-104.

Nielsen, J. and R. Molich (1990). Heuristic evaluation of user interfaces. ACM CHI, Seatle.

Nofi, A. A. (2000). Defining and Measuring Shared Situational Awareness, Center for Naval Analyses.

Norman, D. A. (1991). Cognitive artifacts, Cambridge University Press.

Nulden, U. (2003). Investigating police patrol practice for design of IT. CHI '03 Extended Abstracts on Human Factors in Computing Systems. Ft. Lauderdale, Florida, USA, ACM**:** 820-821.

O'Brien, K. S. and D. O'Hare (2007). "Situational awareness ability and cognitive skills training in a complex real-world task." Ergonomics **50**(7): 1064-1091.

Ogata, K. (2010). Modern Control Engineering, Fifth Edition, Prentice Hall.

Omodei, M. M. and A. J. Wearing (1995). "The Fire Chief microworld generating program: An illustration of computer-simulated microworlds as an experimental paradigm for studying complex decision-making behavior." Behavior Research Methods, Instruments, and Computers **27**(3): 303-316.

Orasanu, J. and T. Connolly (1993). The reinvention of decision making. Decision making in action: Models and methods. G. A. Klein, J. Orasanu, R. Calderwood and C. E. Zsambok. Westport, CT, US, Ablex Publishing**:** 3-20.

Orasanu, J. and U. Fisher (1997). Distributed cognition in the cockpit: Linguistic control of shared problem solving. Naturalistic Decision Making. C. Zsambok and G. Klein, Hillsdale, NJ: Erlbaum**:** 343-357.

Orlikowski, W. and J. Baroudi (1991). "Studying information technology in organizations: Research approaches and assumptions." Information Systems Research **2**: 1-28.

Osinga, F. P. (2007). Science, Strategy And War: The Strategic Theory of John Boyd, Taylor & Francis Group.

Paris, C. R., E. Salas, et al. (2000). "Teamwork in multi-person systems: areview and analysis." Ergonomics **43**(8): 1052-1075.

Patterson, E. S. and J. E. Miller (2010). Macrocognition Metrics and Scenarios: Design and Evaluation for Real-World Teams, Ashgate.

Perrow, C. B. (1984). Normal Accidents, Princeton University Press.

Pettersson, M., D. Randall, et al. (2002). Ambiguities, awareness and economy: a study of emergency service work. Proceedings of the ACM conference on Computer supported cooperative work. New Orleans, Louisiana, USA, ACM**:** 286-295.

Pinelle, D. and C. Gutwin (2000). A review of groupware evaluations. Proceedings of 9th IEEE WETICE Infrastructure for Collaborative Enterprises.

Pinelle, D. and C. Gutwin (2002). Groupware walkthrough: adding context to groupware usability evaluation. Proceedings of the SIGCHI conference on Human factors in computing systems, Minneapolis, Minnesota, USA, ACM Press.

Pinsonneault, A. and K. Kraemer (1989). "The impact of technological support on groups: An assessment of the empirical research." Decision Support Systems **5**(3): 197-216.

Polson, P., C. Lewis, et al. (1992). "Cognitive walktroughs: A method for theory-based evaluation of user interfaces." International Journal of Man-Machine Studies **36**(5): 741-773.

Purao, S., C. Y. Baldwin, et al. (2008). The Sciences of Design: Observations on an Emerging Field, Harvard Business School Finance Working Paper No. 09-056.

Putman, L. (1983). The interpretive perspective: An alternative to functionalism. Communication and organizations, an interpretive approach. L. Putman and M. E. Pacanowsky. Bevery Hills, Sage**:** 31-54.

Qureshi, Z. H. (2007). A review of accident modelling approaches for complex socio-technical systems. Proceedings of the twelfth Australian workshop on Safety critical systems and software and safety-related programmable systems - Volume 86. Adelaide, Australia, Australian Computer Society, Inc.**:** 47-59.

Rasmussen, J. (1983). "Skills, Rules and Knowledge: signals, signs and symbols and other distinctions in human performance models." IEEE Transactions on Systems, Man And Cybernetics **13**(3).

Rasmussen, J. (1997). "Risk Management in a Dynamic Society: A Modelling Problem." Safety Science **27**(2/3): 183-213.

Rasmussen, J., A. M. Pejtersen, et al. (1994). Cognitive systems engineering, Wiley.

Reason, J. (1990). Human Error, Cambridge University Press.

Reason, J. T. (1997). Managing the risks of organizational accidents, Aldershot: Ashgate.

Reason, J. T., J. Carthey, et al. (2001). "Diagnosing “vulnerable system syndrome”: an essential prerequisite to effective risk management." Quality in Health Care **10**(suppl 2): ii21-ii25.

Rentsch, J. R. and R. J. Hall (1994). Members of great teams think alike: A model of team effectiveness and schema similarity among team members. Advances in Interdisciplinary Studies of Work Teams. M. Byerlein and D. Johnson.

Roberts, K. H. (1990). "Managing high reliability organizations." California Management Review **32**(4): 101-114.

Rolo, G. and D. DÍaz-cabrera (2005). "Decision-making processes evaluation using two methodologies: field and simulation techniques." Theoretical Issues in Ergonomics Science **6**(1): 35-48.

Rosen, M. A., S. M. Fiore, et al. (2008). "Tightly Coupling Cognition: Understanding how communication and awareness drive cognition in teams." C2 Journal **2**(1): 1-30.

Rosenthal, U., A. Boin, et al. (2001). Managing Crises: Threats, Dilemmas, Opportunities., Springield.

Salas, E., C. S. Burke, et al. (2001). "Understanding Command and Control Teams Operating in Complex Environments." Inf. Knowl. Syst. Manag. **2**(4): 311-323.

Salas, E., T. L. Dickinson, et al. (1992). Toward an understanding of team performance and training, Nordwood, NJ.

Salas, E., S. Fiore, et al. (2010). "Emerging multi-disciplinary theoretical perspectives in team cognition: an overview." Theoretical Issues in Ergonomics Science **11**(4): 245-249.

Salas, E., C. Prince, et al. (1995). "Situation awareness in team performance." Human Factors **37**(1): 123-126.

Salas, E., D. E. Sims, et al. (2005). "Is there a “Big Five” in Teamwork?" Small Group Research **36**(5): 555-599.

Salas, E., K. C. Stagl, et al. (2007). Fostering team effectiveness in organizations: Toward as integrative theoretical framework of team performance. Modeling complex systems: Motivation, cognition and social processes, Nebraska Symposium on Motivation, Nebraska, Lincoln, NE: University of Nebraska Press.

Salmon, P. M., N. A. Staton, et al. (2008). "What really is going on? A review of situation awareness models for individuals and teams." Theoretical Issues in Ergonomics Science **9**(4): 297-323.

Sauer, J., D. Burkolter, et al. (2008). "The effects of heuristic rule training on operator performance in a simulated process control environment." Ergonomics **51**(7): 953-967.

Sauer, J., D. G. Wastell, et al. (2000). "A conceptual framework for designing micro-worlds for complex work domains: a case study of the Cabin Air Management System." Computers in Human Behavior **16**(1): 45-58.

Scholtz, J. and M. P. Steves (2004). A framework for real-world software system evaluations. Proceedings of the 2004 ACM conference on Computer supported cooperative work. Chicago, Illinois, USA, ACM: 600-603.

Schraagen, J. M. and J. v. d. Ven (2008). "Improving decision making in crisis response through critical thinking support." Journal of Cognitive Engineering and Decision Making. **2**: 311-327.

Serfaty, D., E. E. Entin, et al. (1998). Team coordination training. Making decisions under stress: Implications for individual and team training. J. A. C.-B. E. Salas. Washington, DC, US, American Psychological Association: 221-245.

Shrivastava, P. (1992). Bhopal: Anatomy of a Crisis, SAGE Publications.

Shu, Y. and K. Furuta (2005). "An inference method of team situation awareness based on mutual awareness." Cognition, Technology & Work **7**: 272–287.

Simon, H. (1996). The Sciences of the Artificial, The MIT Press.

Sinnreich, R. H. (2008). Variables and Constants: How the Battle Command of Tomorrow Will Differ (or Not) from Today’s. Battle of Cognition: The Future Information-Rich Warfare and the Mind of the Commander. A. Kott, Greenwood Publishing Group: 10-36.

Smith, J. A. (1995). Semi-Structured Interview and Qualitative Analysis, Londom: Sage.

Smith, K. and P. A. Hancock (1995). "Situation awareness is adaptive, externelly directed consciousness." Human Factors **37**: 137-148.

Stanton, N. A., P. R. G. Chambers, et al. (2001). "Situational awareness and safety." Safety science **39**(3): 189-204.

Stanton, N. A., R. Stewart, et al. (2006). "Distributed situation awareness in dynamic systems: theoretical development and application of an ergonomics methodology." Ergonomics **49**(12-13): 1288-1311.

Stanton, N. A., R. Stewart, et al. (2006). "Distributed situation awareness in dynamic systems: theoretical develoment and application of an ergonomics methodology." Ergonomics **49**(12-13): 1288-1311.

Sterman, J. (1994). "Learning in and about complex systems." System Dynamics Review **10**(2-3): 291-330.

Steves, M. P., E. Morse, et al. (2001). A comparison of usage evaluation and inspection methods for assessing groupware usability. International ACM SIGGROUP Conference on Supporting Group Work, Boulder, Colorado, USA, ACM.

Stewart, D. and P. Shamdasani (1990). Focus Groups : Theory and Practice, SAGE Publications.

Stiermerling, O. and A. Cremers (1999). "The use of cooperation scenarios in the design and evaluation of a CSCW system." IEEE Transact on Software Engineering **24**(2): 1171-1181.

Stout, R. J., J. A. Cannon-Bowers, et al. (1999). "Planning, Shared Mental Models, and Coordinated Performance: An Empirical Link Is Established." Human Factors: The Journal of the Human Factors and Ergonomics Society **41**(1): 61-71.

Suchman, L. (1987). "Plans and Situated Actions: The problem of human-machine communication." Cambridge University Press.

Tamhane, A. C. and D. D. Dunlop (2000). Statistics and data analysis: from elementary to intermediate, Prentice Hall.

Taylor, M. R., M. R. endsley, et al. (1996). Situational Awareness workshop report. Applied aviatin psychology: Achivement, change and challenge. B. J. Hayward and A. R. Lowe. Aldershot, UK, Ashgate Publishing**:** 447-454.

Taylor, R. M. (1990). Situational Awareness Rating Technique (SART): the development of a tool for aircrew systems design. Situational Awareness in Aerospace Operations (AGARD-CP-478).

TeamSpeak. "http://www.teamspeak.com/."

Thompson, J. D. (1967). Organizations in Action: Social Science Bases of Administrative Theory, Transaction Pub.

Trauth, E. M. (2001). The Choice of Qualitative Methods in IS Research. Qualitative Research in IS: Issues and Trends, IGI Global**:** 1-19.

Tung, L. and E. Turban (1998). "A proposed research framework for distributed group support systems." Decision Support Systems **23**(2): 175-188.

Tuomisto, A. (1999). Autonomy of Group Work: The Realm of Work Groups and Information Systems. Department of Computer Science. Turku, Finland, University of Turku. **Licenciate**.

Turoff, M. (1991). "Computer-mediated communication requirements for group support." Journal of Organizational Computing **1**(1): 85-113.

Turoff, M., M. Chumer, et al. (2004). "Assuring Homeland Security: Continuous Monitoring, Control & Assurance of Emergency Preparedness." Journal of Information Technology Theory and Application **3**(3): 1-24.

Turoff, M., M. Chumer, et al. (2004). "The Design of a Dynamic Emergency Response Management Information System (DERMIS)." Journal of Information Technology Theory and Application **4**(3).

Turvey, M. T. and R. E. Shawn (1995). Toward an ecological physics and physical psychology. The science of the mind: 2001 and Beyond. R. L. Solso and S. W. Massaro, New york: Oxford University Press**:** 144-169.

Tversky, A. and D. Kahneman (1974). "Judgment under Uncertainty: Heuristics and Biases." Science **185**(4157): 1124-1131.

Urban, J. M., S. A. Bowers, et al. (1995). The importance of team architecture in understanding team processes. Advances in Interdisciplinary Studies in Work Teams. M. Beyerlein. Greenwich, JAI Press. **2:** 205-228.

Vaughan, D. (1996). The Challenger Launch Decision: Risky Technology, Culture and Deviance at NASA, University of Chicago Press.

Vicente, K. (1999). Cognitive Work Analysis : Toward Safe, Productive, and Healthy Computer-Based Work, CRC Press.

Vidulich, M., C. Dominguez, et al. (1994). Situation Awareness: Papers and Annotated Bibliography. Interim report., Armsrtrong Lab Wright-Patterson AFB OH Crew Systems Directorate.

Wagenaar, W. A., J. Groeneweg, et al. (1994). "Promoting safety in the oil industry. The Ergonomics Society Lecture Presented at the Ergonomics Society Annual Conference, Edinburgh, 13-16 April 1993." Ergonomics **37**(12): 1999-2013.

Wagner, F. (2006). Modeling Software With Finite State Machines: Practical Approach, CRC PressINC.

Wagner, F., T. Wagner, et al. (2004). Closing the gap between software modelling and code. Engineering of Computer-Based Systems, 2004. Proceedings. 11th IEEE International Conference and Workshop on the.

Weaver, J., C. Bowers, et al. (1995). "Networked simulations: New paradigms for team performance research." Behavior Research Methods, Instruments, & Computers **27**(1): 12-24.

Wegner, D. M. (1986). Transactive memory: A contemporary analysis of the group mind. Theories of group behavior. B. Mullen and G. R. Goethals. New York, Springer-Verlag**:** 185-208.

Weick, K. E. (1987). "Organizational culture as a source of high reliability." California Management Review **29**(2): 112-127.

Weick, K. E. (1988). "Enacted sensemaking in crisis situations." Journal of Management Studies **25**(4): 305-318.

Weick, K. E. and D. Meader (1992). Sensemaking and Group Support Systems, Macmillan.

Weick, K. E., K. M. Sutcliffe, et al. (1999). Organizing for High Reliability: Processes of Collective Mindfulness. Research in Organizational Behavior. I. R. Sutton and M. B. Staw, Elsevier Science/JAI Press. **21:** 81-123.

Wellens, A. R. and D. Ergener (1988). "The C.I.T.I.E.S. game: a computer-based situation assessment task for studying distributed decision making." Simul. Gaming **19**(3): 304-327.

Whittaker, S. and B. Amento (2003). Seeing what you are hearing: Co-ordinating responses to trouble reports in network troubleshooting European Conference on Computer Supported Cooperative Work. K. Kuutti, E. H. Karsten, G. Fitzpatrick, P. Dourish and K. Schmidt. Helsinki, Finland, Springer**:** 219-238.

Wickens, C. D. (1984). Engineering psychology and human performance, Merrill.

Wiener, N. (1961). Cybernetics: Or, Control and Communication in the Animal and the Machine, Mit Press.

Wilson, K. A., E. Salas, et al. (2007). "Errors in the Heat of Battle: Taking a Closer Look at Shared Cognition Breakdowns Through Teamwork." Human Factors **49**(2): 243-256.

Woo, D. M. and K. J. Vicente (2003). "Sociotechnical systems, risk management, and public health: comparing the North Battleford and Walkerton outbreaks." Reliability Engineering & System Safety **80**(3): 253-269.

Woods, D. D. (2006). Essential Characteristics of Resilience. Resilience Engeneering: Concepts and Percepts. E. Hollnagel, D. D. woods and N. Levenson, Ashgate**:** 21-34.

Woods, D. D. and E. M. Roth (1986). Models of cognitive behavior in nuclear power plant personnel. Tech. Rep. NUREG-CR-4532. Washington DC, Nuclear Regulatory Commission.

Wybo, J.-L. and M. Latiers (2006). "Exploring complex emergency situations' dynamic: theoretical, epistemological and methodological proposals." Int. J. Emergency Management **3**(1): 40-51.

Xiao, T., W. Broxham, et al. (2010). Two Human-Centered Approaches to Health Informatics: Cognitive Systems Engineering and Usability. E-Health. H. Takeda, Springer Berlin Heidelberg. **335:** 15-26.

Xiao, Y., W. A. Hunter, et al. (1996). "Task Complexity in Emergency Medical Care and Its Implications for Team Coordination." Human Factors: The Journal of the Human Factors and Ergonomics Society **38**(4): 636-645.

Xiao, Y., C. F. Mackenzie, et al. (1998). "Team Coordination and Breakdowns in a Real-Life Stressful Environment." Proceedings of the Human Factors and Ergonomics Society Annual Meeting **42**(3): 186-190.

Zsambok, C. E., L. R. Beach, et al. (1992). A literature review of analytical and naturalistic decision making. Fairborn, OH, Klein Associates Inc. Technical Report.