

Assessing Shared Strategic Understanding

by

Peter Berggren



Linköping Studies in Arts and Science No. 677
Faculty of Arts and Sciences

Department of Computer and Information Science
Linköping University
SE-581 83 Linköping, Sweden

Linköping 2016

Linköping Studies in Arts and Science • No. 677

At the Faculty of Arts and Sciences at Linköping University, research and doctoral studies are carried out within broad problem areas. Research is organized in interdisciplinary research environments and doctoral studies mainly in graduate schools. Jointly, they publish the series Linköping Studies in Arts and Science. This thesis comes from the Division for Human-Centered Systems at the Department of Computer and Information Science.

ISBN 978-91-7685-786-1
ISSN 0282-9800
Printed by: LiU-Tryck 2016

Copyright © 2016 Peter Berggren
Department of Computer and Information Science 2016

Cite as:

Berggren, P. (2016). *Assessing Shared Strategic Understanding*. Linköping Studies in Arts and Science, Dissertation No. 677. Linköping, Sweden: Linköping University Electronic Press.

URL: <http://urn.kb.se/resolve?urn=urn:nbn:se:liu:diva-126346>

Cover illustration © Eric Palmquist/Bildupphovsrätt 2016

Tous pour un, un pour tous

Les trois mousquetaires (1844) by Alexandre Dumas

På ärans och hjältarnas språk:

En för alla, och alla för en.

Abstract

This thesis describes the development of an instrument for assessing shared understanding in teams. The purpose was to develop an instrument that would be *usable, understandable, objective, flexible* and *self-explanatory*. Teams working in naturalistic settings are expected to have a shared understanding of common goals and how to achieve these. The problem investigated in this thesis is that current techniques and instruments for assessing shared understanding in teams generally suffer from one or more of the following drawbacks, namely that they are expensive, difficult to use, time-consuming, requiring expertise, and are often based on subjective perceptions. Departing from existing theory in team cognition techniques and theories, the research questions posed in this thesis are: 1) How can shared understanding be measured without the disadvantages of existing methods? 2) How can shared understanding be assessed without the bias of self-ratings and/or assessments by experts/observers? 3) Can team performance be better understood by the outcomes of an instrument that measures shared understanding?

These research questions are answered through six studies that are presented in this thesis. Over the six studies an instrument was iterated and subsequently developed, called *the shared priorities instrument*. When using this instrument, team members are instructed to generate items and rank these in order of importance. By comparing these rank orders from different participants, a team measure of shared understanding can be calculated. The advantages of this instrument compared to earlier measures are that it is less expensive, easier to use, less time-consuming, does not require subject matter expertise, and that the instrument is distanced from subjective perceptions. Furthermore, the final study provides results where outcomes from the shared priorities instrument correlate with performance, supporting earlier research connecting shared understanding in teams with team performance. A structural equation model, a result of the final study, shows that the instrument is both valid and reliable.

Acknowledgement

The work in this thesis has been carried out at Linköping University and the Swedish Defence Research Agency. The contribution of my friends, colleagues, and former colleagues are gratefully acknowledged. In particular, thanks to:

Professor Jan Andersson and Professor Erland Svensson, my initial supervisors, who got me started on this scientific journey. Their focus on empirical studies and analysis has always been an inspiration to me. My appreciation also extends to Dr. Maud Angelborg-Thanderz who was part of the initial ensemble and who taught me the value of networks and respect for expertise.

Dr. Björn JE Johansson, my friend and supervisor, who has kept me on my path towards the dissertation and who also made the second half of the thesis project enjoyable and fun.

Professor Nils Dahlbäck, my supervisor, for his encouraging support and theoretical advice.

Professor Henrik Artman, my supervisor for a brief time and challenger of the final manuscript.

To my friends Dr. Staffan Nählinder and Dr. Johanna Nählinder, Dr. Henrik Danielsson, and to my colleague Professor Tom Ziemke for reading and commenting on the text. To my friends Dr. Martin Castor and Dr. Joakim Dahlman who provided encouragement and support while being co-PhD students, and who, together with the persons mentioned above, pushed me to finalise this thesis. To Dr Jens Alfredson for the productive collaboration. To Dr. Rego Granlund for support in running experiments in the C3Fire microworld.

To the students who helped in running experiments (in chronological order): Andreas Kjellin, Jonathan Svensson, Therese Allvar, Fredrik Höglund, Sandra Jonsson, Erik Prytz, Anna Tullberg, Sara Berglund, and Nicoletta Baroutsi.

I am grateful for the help and support from the Swedish Armed Force's personnel and the students who participated in the studies.

To my friends outside the scientific community for letting me have a normal life without methodological concerns or questioning my theoretical assumptions.

To my family for all love and care; my mother and father, Göran and Marianne Berggren, my brother Jimmy, and my sister Sonja with families.

This thesis is dedicated to my children Isak, Love, Siri, and Wera with love.

I want to end this acknowledgement with a thank you to Karolina who has stood by my side and supported me in finishing this, sometimes seemingly endless, journey.

Peter Berggren

Linköping March 2016

Table of Contents

Foreword	1
Aim of the work	2
Delimitation	3
A note on the title	3
Progress of the work	4
Chapter 1. Introduction	7
Why teams?	9
Useful for what domains or which target groups?	9
The problem	10
Purpose	11
Research questions	12
Instrument developmental process	13
Outline of the thesis	15
Chapter 2. Theoretical background	17
Central concepts	17
Teams	18
Team cognition	19
Directions in team cognition	19
Shared understanding	23
Summary	25
Synthesis of theoretical approaches to team cognition	25
Measuring	28
Level of analysis.....	28
Measuring team cognition and shared understanding	29

Approaches to measuring.....	33
Empirical studies	35
Chapter 3. Study 1	37
Measurement of Overlap	38
Measurement of Calibration	39
Method	39
Participants.....	39
Design.....	40
Material.....	40
Procedure.....	42
Scoring.....	42
Results	44
Overlap and calibration	44
<i>Overlap and Calibration in relation to team performance.</i>	46
Discussion	48
Acknowledgement and publication	50
Chapter 4. Study 2	51
Purpose.....	52
The Swedish Air Force Combat Simulation Centre	52
Teams and fourships	53
Method	54
Participants.....	54
Design.....	54
Material.....	55
Apparatus/Platform.....	55
Scenario.....	55
Procedure.....	56
Scoring.....	56
Results	56
Discussion	58
Conclusion	59
Acknowledgement and publication	60
Chapter 5. Study 3	61

Microworlds	61
C3Fire	62
Purpose	63
Method	63
Participants.....	63
Design.....	63
Material.....	64
Apparatus/Platform.....	64
Scenario.....	65
Procedure.....	66
Scoring.....	66
Results	67
Discussion	70
Summary.....	72
Acknowledgement and publication	73
Chapter 6. Study 4	75
Purpose	76
Background	76
Method	76
Participants.....	76
Design.....	77
Material.....	77
Apparatus/Platform.....	77
Scenario.....	78
Procedure.....	78
Scoring.....	78
Results	78
Shared understanding.....	79
Shared situation awareness.....	79
Team performance.....	79
Feedback on the instrument.....	80
Discussion	80
Conclusion	80
Summary.....	81

Acknowledgement and publication	81
Chapter 7. Study 5	83
Distributed Assessment of Team Mutual Awareness (DATMA)	84
The crew awareness rating scale (CARS)	84
Purpose	85
Method	85
Delimitation	85
Participants	85
Design	86
Dependent Measures	87
Material	87
Apparatus/Platform	89
Scenario	89
Procedure	90
Results	91
Discussion	95
Summary	95
Acknowledgement and publication	96
Chapter 8. Study 6	97
Purpose	97
Method	98
Delimitation	98
Participants	98
Design	98
Material	100
Apparatus	102
Scenario	102
Procedure	103
Analysis	103
Results	103
Is it possible to distinguish between trained and non-trained teams using the shared priorities measure?	104
Does the shared priorities measure correlate with performance?	106

Does the shared priorities measure correlate with other established measures of mutual awareness and shared situation awareness?	106
Modelling the data	109
Discussion	110
Summary	112
Acknowledgement and publication	113
Chapter 9. Discussion.....	115
Findings.....	115
Study 1 – Operation Flashpoint	115
Study 2 – the Swedish Air Force Combat Simulation Centre.....	116
Study 3 – C3Fire1.....	116
Study 4 – the Tank Crew Training Facility.....	117
Study 5 – C3Fire2.....	118
Study 6 – C3Fire3.....	118
The research questions	119
General discussion	121
Theoretical implications	122
Shared strategic understanding.....	125
Ecological validity.....	125
Instrument rationale	126
Criticism	127
Future work	128
Chapter 10. Conclusions	129
Contribution.....	131
Concluding remarks.....	132
Chapter 11. References	133
Chapter 12. Appendix 1	149
Instructions on using the shared priorities instrument	149

Foreword

I have researched military and civil crisis response problems from a Human Factors perspective most of my working life. The work of this thesis started in 2001 while I was employed at the Swedish Defence Research Agency (FOI). At the Swedish Defence Research Agency most of the work concerned research questions regarding the Swedish Armed Forces, but I have been engaged in several domains where operators¹ are involved in some kind of operations or activity; aviation (Dekker et al., 2001; Magnusson & Berggren, 2002; Zon et al., 2004; Nählinder et al., 2005), command and control (Svensson et al., 2006; Berggren et al., 2009; Baber et al., 2010; Essens et al., 2010; Svensson et al., 2010; Thunholm et al., 2011; Banko et al., 2013; Paris et al., 2014), assessment (Nählinder et al., 2004; Castor et al., 2003), fire-fighting (Lindgren et al., 2007), and oxygen saturation (Andersson et al., 2002b)

The main part of the work within this thesis has been carried out with the needs of the Swedish Armed Forces in mind. During the late 1990s and until 2007 a large development project was concerned with creating the next generation of command and control systems with concepts such as NEC (Network-Enabled Capability; Alberts, Huber, & Moffat, 2010; Alberts & Hayes, 2003), NCW (Network-Centric Warfare; Alberts, Garstka, & Stein, 1999), and the Swedish equivalent NBF (Nätverksbaserat försvar). The NBF-studies have been described in several reports (Berggren, 2004a; 2004b; 2005; 2006). Substantial interest was focused on system interoperability in coalition forces (such as C4ISR – Command, Control, Communications, Computers, Intelligence, Surveillance and Reconnaissance – and network-oriented defence). Several multinational studies were carried out at the Swedish Armed Force’s developmental environment in Enköping, Sweden (JCDEC – Joint Concept Development and Experimentation

¹ The term operator indicates the user of a system. This could be the soldier, fighter pilot, staff officer, etc.

Centre, in Swedish FMKE – Försvarsmaktens centrum för Konceptutveckling och Experimentverksamhet). The project was split into four fronts – method, organisation, personnel, and technology. One of the issues that arose was how to assess a command team's shared understanding of the common operational picture. Another development project with similar needs was the Command & Control Warfare Simulator (LKS²) project where electronic warfare could be simulated and where operators could test and try different approaches, consequences, and outcomes of electronic warfare (Hammervik et al., 2007; Hammervik et al., 2006; Castor et al., 2008; Berggren et al., 2008). In parallel with the technical development of the simulator, assessment methods for performance, situation awareness, shared understanding of the common operational picture, and decision making were evaluated and developed. The need for an assessment method to evaluate to what extent a team of operators had a shared understanding of the situation was reoccurring in several studies (Rencrantz et al., 2005; Rencrantz et al., 2007) and was considered as important for the Swedish Armed Forces.

The military and civil crisis response domains are applied settings closely linked to the naturalistic situations characterised by Klein et al. (1993). Coming from a cognitive science/cognitive psychology background I have seen the use of knowledge achieved in structured laboratory experiments, while also having worked hard on building knowledge through field studies and experiments. Often the research issues in the domains where I have worked have concerned teams, and as I have seen the use for increased knowledge about teams I have worked to incorporate knowledge from the field of team cognition into these work domains, while at the same time trying to resolve applied problems through research. I have tried to study the problems in a structured manner, using the tools I am trained in handling (such as experimental design, statistical analysis, methods and instruments for assessing different concepts, skills, and functions, while trying to explain a phenomenon through a theoretical perspective). The research has mainly been funded by the Swedish Armed Forces research and technology programme, which also means that I have tried to look at the results from a user/benefit-perspective for the Swedish Armed Forces.

Aim of the work

A basic assumption of this thesis aligns with the words written by Mohammed and Dumville (2001, p. 89) “The general thesis of the shared mental model literature is that team effectiveness will improve if team members have an adequate shared understanding of the task, team, equipment, and situation”. Of interest in this thesis is to assess this shared understanding,

² In Swedish LedningsKrigföringsSimulator

focusing on shared strategic understanding. Shared strategic understanding relates to the desired future state/s that the team is planning for. The researchers in the field of team cognition have come up with plenty of measurement techniques. Coming from a background where employing the methods to applied situations is an important aspect for the funder, many of these techniques seemed less useful. Problems are for example, cost, time consumption, and difficulties in applying measures in the field. Consequently I wanted to develop an instrument that would be usable, understandable, objective, and flexible, while not requiring experts (or scientists).

Delimitation

Of course, as the research was mainly funded by the Swedish Armed Forces research and technology programme this frames the research somewhat: the programme demanded a focus and results that benefit the Armed Forces, on a personal level you think about the ethical and moral concerns and implications your efforts will have, and you have to argue for your project in competition with other research projects. Another early delimitation was the team focus of this thesis. While the field of team cognition was growing, so was the awareness among stakeholders for projects funded by the Swedish Armed Forces. The field of team cognition was expanding, and many of the implications were of interest for the Swedish Armed Forces (for example, macrocognition in teams, team decision making, teamwork, etc.). I chose to focus on small teams with an applied benefit. The rationale for this is multifaceted: many teams are small, allowing for direct interaction between team members. Other reasons are the pragmatic reasons which include the requirements formulated by the Armed Forces, and the possibility to accomplish experiments that were realisable.

This thesis touches upon areas such as social psychology, sociotechnical systems, cognitive systems engineering, and distributed cognition, while the main aim is applied team cognition in terms of assessment techniques. Accordingly, the theoretical foundation will mainly be rooted in team cognition.

A note on the title

The title *Assessing shared strategic understanding* is a description of the work done and a reflection of the progress. The use of the word *strategic* included in the title is a result of the findings from the latter studies, a consequence of the empirical findings indicating that the instrument is focusing on the team goals that are further ahead, and not what is happening in the current situation.

The use of different levels (strategic, tactical, and operational) for command and control is mainly to distinguish between different time scales (cf. Trnka, 2009; Lagerlöf & Pallin, 1999). The

concept strategic is used both in the military and crisis response domains. In contrast to operational and tactical, strategic is aimed at the longest time scales. Operational and tactical levels have different meanings in the military and civilian domain. Tactical level is in the military setting the lowest level where execution takes place. For crisis response this is the operational level. Operational level for the military setting concerns coordination and handling of a specific response operation. For crisis response this is called operational level.

Progress of the work

With the purpose of developing an instrument for assessing shared understanding, I was using an approach where each study has been based on the previous study, and where each step provided a stepping-stone for the next move forward (see Figure 1).

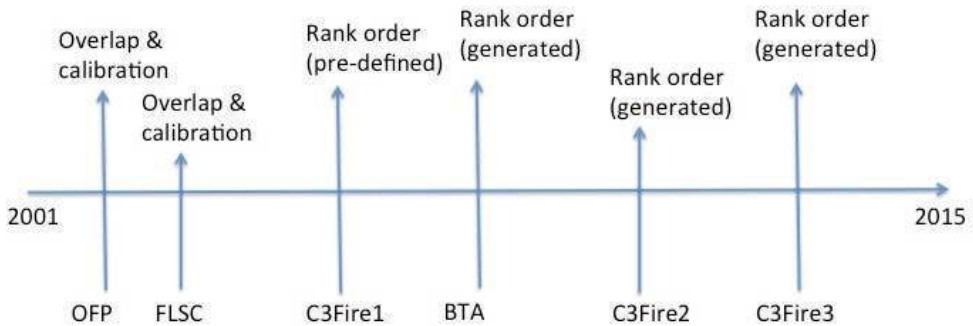


Figure 1. Timeline presenting the order of the different studies.

There are some methodical changes over the different studies. These are presented in Table 1 below. Study 1, 3, 5, and 6 included students as participants, whereas study 2 and 4 involved professionals (fighter pilots and tank commanders). One of the studies (study 3) required subject matter experts, and study 1 and 2 demanded considerable time for the personnel administering the data collection to be able to understand the instrument.

Table 1. Comparison of the different instrument versions.

Factor to consider	Study 1	Study 2	Study 3	Study 4	Study 5	Study 6
Instrument type	Overlap & calibration	Overlap & calibration	Predetermined items	Self-generated items	Self-generated items	Self-generated items
Preparation time	Minutes	Minutes	Hours	Minutes	Minutes	Minutes
Need SME	No	No	Needed for several hours	No	No	No
Requires domain expertise	No	No	Yes	No	No	No
Time to learn for administrator	A couple of hours to understand the concepts, the measure, and the scoring	A couple of hours to understand the concepts, the measure, and the scoring	15-30 minutes	15-30 minutes	15-30 minutes	15-30 minutes
Experimental platform	Operation Flashpoint	FLSC	C3Fire	BTA	C3Fire	C3Fire

The versions of the instrument used in study 3 required a subject matter expert to prepare the items that the lists were composed of. The development of the instrument has gone from more difficult and time consuming to use and prepare to quite quick and easy to deploy. Several different platforms have been used for the different studies (see Table 2). All of them are different simulation platforms, ranging from a PC-game, microworlds³, and different training platforms for professionals.

Table 2. Different platforms used and number of participating teams for the different studies

	Study 1	Study 2	Study 3	Study 4	Study 5	Study 6	Total
Participants	Students	Air Force Pilots	Students	Tank officers	Students	Students	
N	120	64 (8)*	72	9	36	36	337 (281)
N team	40	8 x 2	24	3	18	12	113
Team size	3	4	3	3	2	3	
Design	Split plot	Within- groups	Within- groups	Within- groups	Within- groups	Split plot	

* Eight pilots participated in eight scenarios.

The development has been planned in some respects (structured data collections, iterations between studies with students and studies with professionals from applied settings, and the aim of reducing complexity and difficulty in the use of the instrument), while other aspects have had to depend on external factors (for example, access to military participants, financing of military projects with an interest in the problem, and organisational changes). As is seen in this chapter, the scientific and developmental progress is based on earlier findings and development. When a dead-end was reached, the rationale for the back-tracking and re-planning of a next step was based on what was known at that stage. Details are presented in the respective chapters.

³ A microworld is a computer simulation of a system that is complex, dynamic and opaque (Brehmer & Dörner 1993). It can be used as a research environment.

Chapter 1. Introduction

Being a leader (or project manager, coach, commander, CEO), how do you know that your team has a shared understanding⁴ about where you (as a team) are heading? A good leader might very well be able to detect signals that indicate to what extent a team has a shared understanding about the team's goals. Leaders who do not possess such an ability would still like to know to what extent their team has a shared understanding. No matter if a leader is better or worse at picking up the signals, s/he needs to confirm and validate that the signals are correctly perceived. One way that researchers have studied the sharedness within a team is by turning to the shared mental model concept⁵ (Johnson et al., 2007; Klimoski & Mohammed, 1994; Cannon-Bowers et al., 1993). While the literature agrees that there is a relation between shared mental models and performance, many methods and techniques available to assess this sharedness are cumbersome, difficult to use and/or take a lot of time (Wildman et al., 2014). A popular approach to assess different aspects of teams has been to aggregate responses or assessments of individual team members to achieve a team level measure (Rousseau, 1985). There are several concerns connected to this approach, which also has benefits. One such concern is what happens when individual level measures are aggregated, for example, is team performance the sum of the individual members' performance? In complex settings where performance assessments might be based on the individual team members' perceived performance, the sum of all individual assessments does not have to reflect team cognition in a correct manner. Another problems is that data are collected at one level, aggregated to say something about another level, while conclusions are drawn about yet another organisational level (for example, asking students questions, aggregating results to classroom level, while drawing conclusions about the school system).

⁴ Shared understanding will be discussed later in the following chapter.

⁵ There exist differences in definition between the concept shared mental models and team mental models and there are researchers who use the two concepts interchangeably. The differences and similarities, both conceptually and analytically, are addressed in the next chapter.

In this thesis I move forward to develop an instrument that can capture shared understanding in teams. A common approach is to ask team members questions and by comparing the answers from these responses it will be possible to provide an account of the shared understanding within the team. In this thesis I will advance and expand earlier approaches to develop an instrument that reduces the disadvantages and weaknesses of recognised methods and techniques.

The way teams perform, and the underlying factors shaping that performance is an area that has gained increased interest over the last decades, especially since the introduction of networked computing that enables teams to work jointly on problems, both as co-located and distributed teams (Heath & Luff, 1992; Artman & Waern, 1998; Artman, 2000; Hutchins & Klausen, 1996; Jones & Roelofsma, 2000; Hollnagel, 1998). Teams, and what makes them effective, have become a major concern for practitioners and researchers. In many cases, findings from group research have been applied on teams (Jones & Roelofsma, 2000; Fullager & Egleston, 2008), but there is also a vast body of research on teams that has emerged (see for example Wildman et al., 2014; Mathieu et al., 2008 for a comprehensive overview). A large part of this research has focused on investigating the importance of team awareness and/or sharedness or simply team cognition, especially in research focused on teams in dynamic control tasks. This is motivated by the idea that teams, just as individuals, must be able to sense, understand and act in order to remain in control of a situation (perform within an acceptable performance envelope, Hollnagel, 1998; Hollnagel and Woods, 2005). What mostly is studied (measured) is the *sense and understand* part, in many cases based on the concept of situation awareness. Situation awareness, as defined by Endsley (1995a), is a commonly used construct that, if put rather simplistically, is based on the premise that a person who has a good understanding of a situation has a good basis for handling the same situation (this is not necessarily true, but many studies take as a starting point this line of argumentation). The situation awareness concept, has since then been expanded to teams (Saner et al, 2009). However, this is not unproblematic. Team awareness or team sharedness consists of more than a mere understanding of what is going on in the surrounding world. Rather, it involves the team members' understanding of the different roles in the team, the assumptions of others' knowledge of other team members' understanding of the situation and even interpersonal relationships. As pointed out by Salas and Cannon-Bowers (2001), the whats and hows of assessing team performance and awareness or other aspects of teamwork are diverse and lack consensus.

An important assumption in this thesis is that understanding on the team level is a product of the interactions in the team during work (Cooke et al., 2013; van der Haar et al., 2015) in relation to

context, and as a consequence of team formation (Tuckman, 1965). The actual outcome in terms of shared understanding is thus a consequence of events within the team and in the situation, as well as the maturity of the team. Thus, while an understanding of on-going events is important, it is equally important to have a strong shared understanding of desired goal states at a strategic time horizon.

Why teams?

Teams are, in most organisations, the basic building blocks for achieving more complicated tasks and goals. When organisations are dealing with difficult and complex tasks, teams are the chosen approach (Salas et al., 2008). This becomes more noticeable as work changes because of technology, globalisation, and organisational restructuring (Zajac et al., 2014).

Useful for what domains or which target groups?

The field of team cognition is connected to several research fields, such as social psychology (Aronson et al., 2013; Fiske & Taylor, 2013), macrocognition (Cacciabue & Hollnagel, 1995; Klein et al., 2003; Patterson & Miller, 2010), naturalistic decision making (Klein et al., 1993; Klein, 2009; 1999), human factors (Fitts, 1947), and cognitive systems engineering (Hollnagel & Woods, 1983; Hollnagel & Woods, 2005; Woods & Hollnagel, 2006). Also, distributed cognition (Hutchins, 1995) is closely related to the applied field of team cognition, as is the research field of sociotechnical systems (Emery & Trist, 1960; Stahl, 2010; Duff et al., 2014) and command and control (Andriole & Halpin, 1986; Stanton et al., 2008; Prytz et al., 2010). The domains considered as target domains for this thesis are where work has to be coordinated among multiple actors, in dynamic and complex situations with multiple and changing goals. These settings were defined by the naturalistic decision making researchers (Orasanu & Connolly, 1993), and have since then been used to characterise the dynamic environments that are described within this line of research. The eight factors are: ill-structured problems, uncertain dynamic environments, shifting, ill-defined, or competing goals, action/feedback loops, time stress, high stakes, multiple players, and organisational goals and norms. As Klein and Wright (2016) also stress, experience is another central aspect that was put into focus with the naturalistic decision making (NDM) paradigm. Using the NDM framework on other cognitive phenomena, macrocognition was established (Klein et al., 2000). The macrocognitive framework is suited to study and understand cognitive processes that affect performance of natural tasks (ibid).
Macrocognition is defined:

“as the study of cognitive processes affecting people such as firefighters, pilots, nurses, and others who had to wrestle with difficult dilemmas in complex settings under time pressure and uncertainty” (Klein & Wrigt, 2016, p. 2)

Macro-cognition is concerned with naturalistic decision making, sensemaking/situation assessment, planning, adaptation/replanning, problem detection, and coordination (Klein et al., 2003). These functions are central for the teamwork carried out in both military command and control and in crisis response (Essens et al., 2005; Essens et al., 2010; MacDermott, 2009; Uhr, 2014). In these situations, the common operational picture is often used as a tool for planning and coordinating actions, and also for maintaining awareness among team members about what is happening. A frequent problem is that the common operational picture is confused with the shared situational awareness, i.e. practitioners and system designers tend to make a one-to-one mapping between what is presented in the information system and the understanding that individuals or members of a team using that equipment have of a situation. Another fallacy is the assumption that all members of a team will interpret the same information in the same way. However, *any team or group of people that need to have a shared understanding of their task and goals would benefit from an instrument that could indicate their level of sharedness*. No matter if it is a command staff in theatre (i.e., in the field), an ad-hoc crisis response team during flooding, or an emergency medical dispatch team trying to handle a large accident. To coordinate and synchronise actions they need to have a shared understanding of where they are heading and how to get there. This shared *strategic* understanding would help plan and coordinate team behaviour and actions to move closer towards the strategic goal/s.

The problem

Most methods and techniques for assessing shared understanding that are available today come with some disadvantages: expensive, difficult to use, time-consuming, and often demanding expertise (cf, Wildman et al., 2014). To be able to apply an instrument during operations/exercises it needs to be useful, easy to deploy and analyse, provide a result quickly, and say something about the situation.

To develop an instrument that is psychometrically sound there are some considerations to take into account. Meister (1985) mentions *effectiveness, ease of use, cost, range, flexibility, validity, reliability, and objectivity* as aspects to consider when selecting methods to utilise. These are all concerns that need to be addressed. *Effectiveness* is about the extent to which the method achieves its purpose. *Ease of use* is how straightforward and demanding the method is to carry out. *Cost* involves both financial cost, but also things that are indirectly related to cost (personnel needed, equipment, time, etc.). *Range* concerns the quantity of events, behaviours or activities that can be analysed

and measured by the method. *Flexibility* has to do with different types of contexts the methods can be used in. *Validity* is if the method measures what it is intended to measure. *Reliability* is if the method provides identical/similar results if applied to the same phenomena. *Objectivity* relates to the extent that the method is unaffected by and independent of the researcher, meaning that the measurement and method are true even outside of an individual's subjective biases, feelings, and interpretations.

Purpose

The purpose of this thesis is to develop an instrument for measurement of shared understanding that would be *usable*, *understandable*, *objective*, and *flexible*, while not requiring experts, or even scientists, to manage. These ambitions are expanded below:

- *Usable* so that the instrument is easy enough to operate and employ. That is, the need for equipment and/or experts/administrative personnel is limited. Apart from decreasing complexity (of more persons involved) this would also reduce costs. Partly because the need for expensive recording/data collection/analysis equipment could be minimised, partly since experts/personnel cost. Costs related to experts and/or equipment for assessing teams might very well be a sensible investment in order to get useful results, but these costs will always be compared with the effect money would have if it were spent on operations instead.
- *Understandable* so that most people could use it (if trained) and also easy to understand so that the results are comprehensible for the intended audience (the teams). Also, understandable in the sense that data are collected and presented in a comprehensible manner.
- *Objective*⁶ in the sense that the instrument is not relying on the subjective assessments of self-ratings through team member introspection; "To what extent does my team have a shared understanding?" Also objective in the sense that the outcome of the results is reliable and valid. In addition, the results of the instrument should be independent of the researcher.
- *Flexible*, meaning that it should be possible to use the instrument in different situations and domains, both in the laboratory and in the field.
- Lastly, the instrument should not require subject matter experts to understand the context. Neither to use or to collect data, nor to interpret the outcome. It should, if possible, be *self-explanatory*.

⁶ Objective in the sense that Meister (1985) is using the word.

A second purpose is to understand to what degree a refined instrument might assist in understanding team cognition. This would provide a deeper understanding of team cognition using the shared priorities instrument.

Research questions

The main research questions are:

1. How can shared understanding be measured without the disadvantages of existing methods?

Motivation: Today's methods and techniques are difficult, time-consuming, often demand subject matter experts, require extensive understanding of the environments/context, and are expensive to use (cf. Wildman et al., 2014). These downsides will be elaborated on in the theoretical background chapter.

2. How can shared understanding be assessed without the bias of self-ratings and/or assessments by experts/observers?

Motivation: Subjective ratings are biased by the individual's view and conceptions about a phenomenon or the situation. Another dimension of subjectivity might be the expert's interpretations, if an expert is needed to rate behaviours or to interpret statements and behaviours of participants (cf. Meister, 1985). Hence, an instrument that can avoid these biases is preferred.

3. Can team performance be better understood by the outcomes of an instrument that measures shared understanding?

Motivation: Relating a measure to the outcome is important as this can help us understand to what extent a measured concept explains the outcome/performance. In turn that can help in development of training programmes, when it is known how important a concept is in contributing to success.

The ambitions of the instrument, stated in the purpose, are linked to the research questions in Table 3 below.

Table 3. Relating the research questions to the ambitions with an instrument.

	RQ1	RQ2	RQ3
Usable	X	-	-
Understandable	X	-	-
Objective	X	X	-
Flexible	X	-	-
Self-explanatory	-	X	X

Research question 1 concerns how an instrument can be easy to operate (usable), comprehensible (understandable), not relying on subjective assessments of a phenomenon (objective), and is usable in different situations (flexible). Research question 2 concerns objectivity in the sense of valid and reliable, and that the instrument is self-explanatory in the sense that experts are required to interpret the results. Research question 3 concerns if an instrument is to be self-explanatory a correlation between the outcome of the instrument and team performance is required, i.e., that the outcome of the instrument explains in part the variance in team performance linking the shared understanding in a team to how well that team performs.

Instrument developmental process

Several approaches to tackle the problem of developing an instrument for assessing something could be used. Here, a quantitative approach has been chosen to facilitate generalisation (Yin, 1994).

A common approach to method design is provided by Johnson et al. (2007) explaining how they develop a questionnaire for assessing sharedness of team-related knowledge. The authors describe the development steps and the conceptual framework for factors associated with shared mental models. The procedure includes six steps: development of the initial instrument, content validation, descriptive statistical analysis, exploratory factor analysis, conceptualisation of factor analysis, and confirmatory factor analysis.

Another approach is described by Teo (2013) who, in order to develop and validate an assessment scale, undertakes three phases: item generation, pilot test, and validation. Teo generated items by reviewing the literature and empirical studies to identify indicators of digital nativity (his area of interest). Then a focus group discussed the initial set of items to reduce the number of items, to formulate statements that could be understood by future informants and which allowed for an informant to agree or disagree with. The next step was phase two, the pilot test, where the aim was to test and refine the items proposed for the scale, which was done with participants who were representative of intended future users of the scale. The third phase was the validation of the scale, where Teo tested the model that he had achieved during the pilot test, with new participants.

A slightly different approach was taken by Essens et al. (2005), where the researchers developed an instrument for assessing command team effectiveness. The authors review a large number of team effectiveness and team performance models from which they develop a model that fits with military command teams. The different components of the model are then operationalised using criteria such as: fit a command team environment, with an established significant contribution to team effectiveness, with minimal overlap between items, and possible to measure. The instrument captures subjective judgments. This command team effectiveness instrument is comprised of 122 items. This instrument was later verified (Essens et al., 2010) in a follow-up study. The follow-up study resulted in a reduced number of items in the instrument (now 32) and the instrument being used in studies of command team effectiveness (Hof et al., 2010; Allberg et al., 2014; Thunholm et al., 2014).

With inspiration from Teo (2013) the work in this thesis has applied an iterative procedure with the following steps in order to develop the instrument:

1. Identify demands on the instrument and what frame the instrument is intended to be used within.
2. Build initial instrument.
3. Test instrument.
4. Change the instrument based on outcome of step 3. Carry out step 3 again if changes are considerable.
5. Validate the instrument.

The whole point of developing, testing, iterating, and testing the instrument is to come up with a valid and reliable tool that can help in the pursuit of better understanding the world and how different concepts relate to one another.

Outline of the thesis

Firstly, an introduction where the problem is presented, and where the purpose and the research questions are stated. This section is followed by a theoretical background where relevant literature is reviewed framing the problem. The next section is a research progress chapter explaining the development of the thesis project. Next, the six empirical studies are presented, showing how the results from one study become the input and starting point for the next study. The studies are presented chronologically. The following section is the discussion where the results are viewed in the light of the literature. At the end are the conclusions and references.

Chapter 2. Theoretical background

This chapter aims at providing a theoretical foundation for the empirical studies presented later in this thesis. Firstly, the notion of team is introduced. After that team cognition is presented. Team cognition is the overarching concept under which other concepts are included, such as shared mental models, shared situation awareness, and shared understanding. One section is dedicated to method development to explain how measurement methods can be developed and validated.

Central concepts

The concepts that are central to this thesis are: teams, team cognition, and shared understanding. These will be briefly presented here, and more thoroughly discussed below. Team cognition is “the cognitive activity that occurs at a team level” (Cooke et al., 2009, p. 158). In this thesis team cognition is the wider theoretical framework which other concepts are linked to. Team performance is the outcome that is used as a standard that other concepts have an impact on, in a positive or negative way. Shared understanding is commonly assumed to be a pre-requisite for coordinating behaviour among the members of a team. Shared understanding indicates that team members who have a shared understanding will be able to coordinate their behaviours (Smart et al., 2009).

In this work there are some additional concepts used to describe measurement techniques: method, measurement, instrument, and assessment. The method is the means to answer the research question (Meister, 1985). Measurement is the process of assigning a number to a characteristic of an event, behaviour or an object, which can be compared with other events, behaviours or objects (Field & Hole, 2003). An instrument is a technique to measure a phenomenon. For example, a thermometer is an instrument that measures temperature. The scale

can be Kelvin, Celsius, or Fahrenheit. Assessment is a judgment by, in the case of behavioural research, an informant (for example the participant, an expert, an observer, or the researcher).

Teams

A team consists of members, who all take certain roles and undertake specialised tasks aimed at reaching a common goal (Orasanu & Salas, 1993; Mathieu et al., 2008). Further, team members' actions mostly take place within the same time-frame and are interrelated and interdependent (Brannick & Prince, 1997). This thesis concerns small teams involving 2-10 members performing in complex settings. Teams are, according to a definition by Salas et al. (1992), "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, and who have a limited life span of membership" (ibid. p. 4). This can be seen in contrast to groups where the roles are not specified, and where the group usually does not have a limited life span.

Definitions of groups, in contrast to teams, are often focused on the social relationships (Donelson, 2006) and that the members of the group define themselves as a group (Benson, 2000). In addition, a group should be recognised as a group by others, they share beliefs, values and norms about areas of common interest. The members identify with each other, and they come together to work on common tasks and for agreed purposes (ibid.). Hence, the main differences are that in teams members have appointed roles with specialised tasks, whereas groups are focused on the social relationships.

Size is known to affect team and group processes and behaviours (Wheelan, 2009). Wheelan shows how group size affects productivity where smaller groups (3 to 8 members) were significantly more productive and more developed than groups with 9 members or more. Šumanski et al. (2007) show that interpersonal relationships are of better quality in small teams compared to large teams. Slater (1958) found that members of five-person groups were most satisfied compared to larger and smaller groups. This was when groups were to discuss human relations problems, yet satisfaction is closely related to the ability to perform, and the task – of course – puts demand on size in terms of team members.

Team members can be co-located or distributed. Teams can be distributed both geographically and temporally. For the purpose of this thesis, focus is on co-located teams collaborating within the same time-frame.

Team cognition

The advances in the field of team cognition were a natural development of applied research rooted in cognitive psychology with questions such as “Do individually oriented cognitive theories and methods apply to teams? Are teams somehow different from a collection of individuals? How can we measure, assess, and design for team cognition?” (Cooke et al., 2007b, p. 239). A broad definition of team cognition is the cognitive activity that occurs within a team (Cooke et al., 2009). Team cognition is about the study and understanding of teams (Salas et al., 2010; Salas & Fiore, 2004; Cooke et al., 2007a). This concerns how team members manage information, communicate, coordinate actions and collaborate, while also looking at how this can be assessed, measured, and modelled. Much attention has been given to aspects of team performance and effectiveness. This is coming from applied problems such as military command and control, sport teams, project teams, teams of first responders, and more (Fransen et al., 2013; Granåsen & Andersson, 2015; Hof et al., 2010; Mathieu et al., 2008; McKendrick et al., 2014; Onag & Tepeci, 2014; Swezey & Salas, 1992; Wijngaards et al., 2006).

Directions in team cognition

Gorman and Cooke (2011) make a distinction between team cognition as shared cognition and team cognition as interactive team cognition. The shared cognition perspective is viewed “as represented by relatively stable, emergent team level knowledge structures that exist within team members’ heads and combine to represent the team” (Wildman et al., 2014, p. 913), whereas the interactive team cognition perspective is viewed “as the dynamic cognitive processes that occur within the team as represented by the interactions between the members of the team” (ibid. p. 913). How these two perspectives relate to team cognition is illustrated in Figure 2 below.

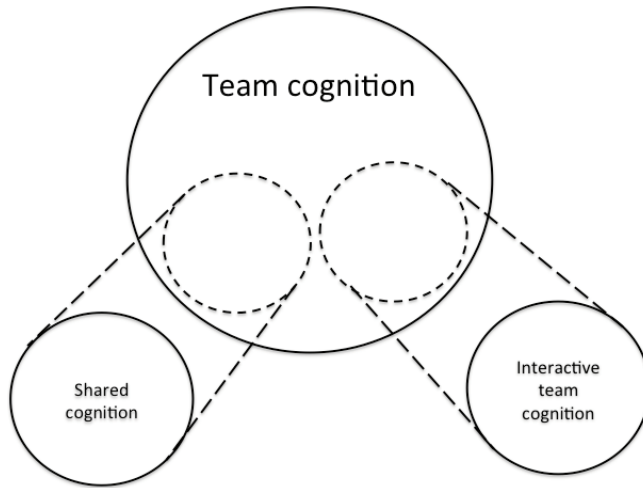


Figure 2. Illustration of how the shared cognition perspective and interactive team cognition perspective are related to team cognition.

Shared cognition

The shared cognition perspective views team cognition as shared knowledge representations (DeChurch & Mesmer-Magnus, 2010b; Banks, & Millward-Purvis, 2009; Cannon-Bowers & Salas, 2001). Shared knowledge can be defined as either complementary or overlapping (Cooke et al., 2003; Cooke et al., 2000; Saner et al., 2009). Complementary is when knowledge is distributed among team members, so that member A knows piece I, member B knows piece II, member C knows piece III, and so on. For example, member A knows how many people are in a building that is on fire. Member B knows which exits are accessible for evacuation. Member C is handling communications with the fire-fighters. Overlapping is when the knowledge that team members have is similar. For example, all members know where the assembly point is if the team needs to leave the theatre of operations and regroup. The shared cognition view of team cognition is rooted in Input-Process-Output (IPO)-models (Gorman & Cooke (2011). These IPO-models are widely used in team research (Mathieu et al., 2008, Hackman, 1987; Essens et al., 2005). An IPO-model (see example below in Figure 3) consists of three main parts (input factors, processes, and outputs), and sometimes a feedback loop is included. The model presents how the input variables (for example contextual factors, task description, the organisational structure, the environment, etc.) can be seen as a starting point, providing a basis for the processes. In team cognition, the processes can be task or team oriented (cf. Essens et al, 2005). These processes affect the outcomes, which also can be task and team related. In some models a feedback loop indicates that both the input factors and processes can be affected by the outputs.

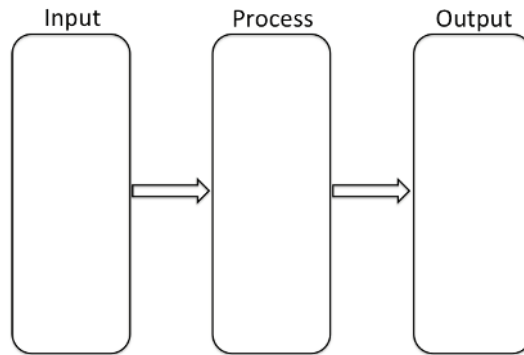


Figure 3. Illustration of generic IPO-model.

Under the shared cognition perspective there are two concepts in particular that are of interest for the purpose of this thesis: shared (or team) mental models and shared situation awareness.

Mental models are organised knowledge structures that allow individuals to interact with their surroundings (Gentner & Stevens, 1983). Accordingly, shared mental models (SMM) concern how the organised knowledge structures are shared in a team. Maynard and Gilson (2014, p. 8) say that “SMMs represent the overlap or convergence among members’ mental representations regarding various aspects of their team and task”. Mathieu et al. (2000) make a distinction between shared mental models regarding task or team. Task-related features the situation (equipment, task, strategies, etc.) whereas team-related features team aspects (team interaction, team members’ knowledge, etc.). Teams with similar mental models imply that team members would work toward a common goal with a shared vision of how their team will achieve that. Shared mental models (both team and task) correlated positively to team process and team performance (ibid.). Mohammed et al. (2010) provide a definition of team mental models (TMM): “TMMs are team members’ shared, organised understanding and mental representation of knowledge about key elements of the team’s relevant environment“ (p. 879).

Team situation awareness originates from the concept of situation awareness(Endsley, 1995a), which relates to the individual’s ability to identify important elements in the situation, understand what these elements mean, and draw conclusions on what is going to happen in the near future. Endsley’s model is rooted in an IPO model. Both shared and team situation awareness (Salas et al., 1995; Endsley & Jones, 2001; Salmon et al, 2008) relate to how the individual team member’s situation awareness and team processes support the team in reaching the team goals. Team situation awareness and shared situation awareness have been studied in a variety of domains: aviation (Endsley,1995b; Hauland, 2008; Jeannot et al., 2003), command and control (Gorman et

al., 2006; Salmon et al., 2006; Artman 1999), emergency management (Danielsson & Larsson, 2014), Breathing Apparatus rescue teams (Fogel et al., 2004), and the cyber arena (Jajodia et al., 2010; Cooke et al., 2013).

The information-processing view of situation awareness can be criticised since it includes cognitive phenomena that are not entirely understood (for example, attention and decision making, cf. Endsley, 1995a, p. 41) and the process of achieving situation awareness seems relatively static even though it concerns highly dynamic situations. Situation awareness is traditionally concerned with the present situation and that the projection of future states mainly concerns states that are seconds to minutes away. There are historic reasons for this, since the early research on situation awareness looked at fighter pilots (Fracker, 1991; Endsley, 1995b; Endsley & Jones, 1997) who are handling threats that are seconds to minutes away, and who most of the time have fuel for roughly an hour of combat. Hence, situation awareness is about dealing with the current and immediate, developing situation. For a team in a different domain the shared goal might be quite some time away. For example, with a crisis management team dealing with a wildfire the shared final goal might be hours, days or even weeks away (Brehmer & Svenmarck, 1995).

Interactive team cognition

Within the interactive team cognition perspective, team cognition is viewed “as the enactment of team processes or interactions” (Wildman et al., 2014, p. 913). The other view on team cognition, interactive team cognition, focuses on the interaction among team members (Gorman & Cooke, 2011). This is a perspective where interaction within the team is team cognition, which is directly observable and correlated to team effectiveness. Team interaction is shaped by changes in the situation. This approach is a reaction to the problems that the shared cognition approach has had trying to link shared knowledge to team effectiveness using mediating IPO-frameworks Salas et al. (2008) say that the team cognition approach generally views teams as information-processing agents where encoding, storing, and retrieval of information are applicable on the team level. Cooke et al. (2008) on the other hand say that ‘between the head’ (cf. interactive team cognition) is rooted in ecological psychology, where focus is on the dynamics between person and environment. They also state that shared cognition is often based on aggregated individual-level inputs, whereas interactive team cognition is looking at team-level inputs. On the team level, communication is seen as a central process (ibid.).

Relation between the two perspectives

Gorman and Cooke explain how the shared cognition view and the interactive team cognition view are related (Gorman & Cooke, 2011, p. 305):

“Therefore, understanding and agency to act as a team must be dynamically assembled through interactions, where the cognitive content of those interactions, as they relate to shared knowledge, may be incidental to the exigencies of the situation at hand. Pursuing the interactive argument, the ontogeny of shared cognition is incidental to team interaction, where interaction is a cognitive substrate upon which shared knowledge may be modified ...”

and continue with

“Thus, although interaction processes do not mediate between shared knowledge and team effectiveness, as in the IPO framework, interaction processes may modify shared knowledge as it relates to team performance.”

This means that understanding and intent is dynamically constructed through the team interactions, while the cognitive content of the interactions might be an effect of the urgent needs due to the situation. The maturity of the shared cognition is a minor consequence of team interaction as the interactions might alter shared knowledge. This indirect influence can be seen as a mediating effect.

The shared cognition and interactive team cognition perspectives both talk about the processes. Shared cognition uses IPO-models, where the processes are something that mediates the Inputs to reach the Outcomes, while the interactive team cognition approach mainly concerns the process in terms of interactions among team members. Most IPO oriented research collects data in terms of “snapshots” of the I, P, and O in the form of query-based assessments or similar techniques. Interactive team cognition on the other hand mostly looks to qualitative aspects of the actual interaction within the team (capturing the dynamic aspects of the processes), and pays little attention to the actual outcome of the interaction, nor the pre-conditions in terms of inputs.

Shared understanding

A concept from outside the team cognition research field is shared understanding. The use of the concept has increased over the last couple of years. It has been/is being used in a diverse set of domains, ranging from jazz improvisation (Schober & Spiro, 2014) to business and IT systems (Jentsch & Beimborn, 2014; Briggs, 2014), development (Corvera Charaf et al., 2013; Steen et al., 2014), from health care (Bates et al., 2014) to teacher education (Lane et al., 2014), from workgroups (Bittner & Leimeister, 2014; Oppl & Stary, 2013) to military coalitions (Smart et al., 2009). Several researchers take a linguistic approach to studying shared understanding (Schober &

Spiro, 2014; Jentsch & Beimborn, 2014; Corvera Charaf et al., 2013). Bates et al. (2014) tested an instrument to measure shared clinical understanding for paediatric cardiac intensive care unit handoffs. Oppl and Stary (2013) studied how cooperative work processes were used to assess the development of a common understanding.

Shared understanding is defined by Smart et al. (2009) as “the ability of multiple agents to exploit common bodies of causal knowledge for the purposes of accomplishing common (or shared) goals” (p. 2). This definition indicates that members (agents) who have a shared understanding will be able to coordinate their behaviours to reach the shared goals. Bittner et al. (2014) define *shared understanding* as “the degree to which people concur on the value of properties, the interpretation of concepts, and the mental models of cause and effect with respect to an object of understanding” (p. 5). Shared understanding is described as a multidimensional construct. Kleinsmann and Valkenburg (2008) define *shared understanding* as “a similarity in the individual perceptions of actors about either how the design content is conceptualised (content) or how the transactive memory system works (process)” (p. 371). Transactive memory (Wegner, 1987) combines the shared awareness about who knows what with knowledge held by the individuals.

Another field focusing on shared understanding is research on strategic consensus (Hambrick, 1981; Gonzalez-Benito et al., 2010; Kellermanns et al., 2011), which is concerned with studying shared understanding in top-level management teams where the goals are set for the future. This field mainly concerns business research, and looks at business performance as the main outcome. However, the concept strategic consensus is of interest as it is defined by Kellermanns (2005, p. 721):

“Strategic consensus is the shared understanding of strategic priorities among managers at the top, middle, and/or operating levels of the organization”

and on page 720:

“the premise that strategic consensus enhances organizational performance by improving coordination and cooperation within the organization.”

This, instead of focus on group processes, is directed at degree of agreement among managers within an organisation.

To conclude, several research fields have investigated and used the term shared understanding. Some studies draw conclusions about shared understanding without measuring the concept, possibly because the term is not always defined and is used in everyday language. When the concept has been studied it has been done using different measures and approaches. This

diversity in research methods, research fields, and applied settings signals that the concept could benefit from an integration of the different uses into a more coherent utilisation.

Summary

The concepts team mental models, shared situation awareness, and strategic consensus, are closely related to shared understanding, both theoretically and conceptually. There are some differences that make them less suitable for the purpose of this thesis. Team mental models concern knowledge structures among team members. These knowledge structures might relate to knowledge that team members have about solving the task or about the roles. The problem here is that the team mental models concept is not capturing the dynamics that are necessary to reflect the development and change in the teams. This is seen in the interactive team cognition approach (for example studying communication), while these approaches often require several days to process and analyse. Hence, methods that take hours to process cannot deliver a result that is relevant in the moment. Shared situation awareness focuses on the immediate, developing situation, whereas a team's shared goal might be distant in time. Strategic consensus is focused on degree of agreement rather than on the processes. What is sought for in the work of this thesis is a measure that is sensitive to dynamic change in the team on a collective level.

Synthesis of theoretical approaches to team cognition

Looking at the different theoretical concepts described above, there are some aspects worth mentioning as they have an impact on this thesis. The team cognition paradigm has several interesting contributions. It provides a framework for team research to compare findings and generate hypotheses about team related concepts. From the shared cognition perspective, a lot of focus has been on near static aspects such as shared knowledge representation, i.e., shared mental models (a team member's expert skills do not change very quickly). In addition, the shared situation awareness assesses aspects that are changing quickly, namely how the current situation is perceived and interpreted, while also predicting how the next moment or situation is evolving. The interactive team cognition perspective studies dynamically changing aspects, namely the interaction between team members. How the concepts shared mental models (SMM), shared situation awareness (SSA), and interactive team cognition (ITC) are related to each other when plotting them regarding dynamics on one axis and time-frame on another axis is seen in Figure 4.

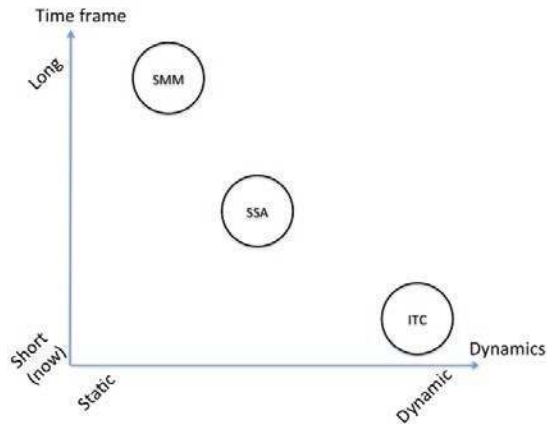


Figure 4. Plot of the concepts shared mental models (SMM), shared situation awareness (SSA), and interactive team cognition (ITC).

The plot visualises how the concepts are separated from each other, both regarding how sensitive the concepts are to change (dynamics), and also to what time-frame the concepts are handling. For example, knowledge representations about skills, knowledge and abilities do not change from moment to moment, whereas the situation evolves affecting the shared situation awareness and the awareness of which tasks to prioritise. The interaction within a team reflects the change of the situation at an even quicker pace.

Shared understanding is focused on how team members have an ability to use shared knowledge to reach shared goals. Combining the team cognition perspective with shared understanding focuses on how a team can reach shared goals (which are semi-static) that are dependent on a dynamically changing situation. An illustration is presented in Figure 5 below.

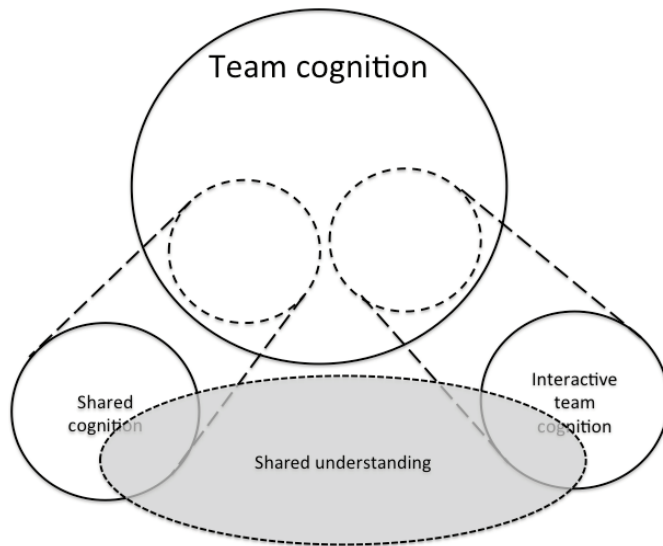


Figure 5. Illustration of how shared understanding can be combined with the team cognition perspective.

In this thesis the definition of shared understanding by Smart et al. (2009) will be used:

“the ability of multiple agents to exploit common bodies of causal knowledge for the purposes of accomplishing common (or shared) goals” (p. 2)

This definition relates the multiple agents’ shared understanding to reach the common goals. Multiple agents can be seen as a team (two or more individuals striving towards (a) shared goal/s). To exploit common bodies of causal knowledge relates to coordinating actions to reach the shared goals. These goals can be considered to be on a strategic timeline. Hence, the shared understanding this thesis is aimed at can be called shared strategic understanding. Maintaining this ability, to have a shared understanding, helps a team to accomplish its common goals. Teams who can maintain a shared understanding will be better at coordinating their actions to reach the shared goals than teams who have problems maintaining a shared understanding. This aligns with the claim made by Mohammed and Dumville (2001), that a shared mental model is correlated with team performance. It is also clear that the methods and techniques for assessing shared understanding are still quite immature and need to be developed to be useful outside of scientific research, but also to gain a higher validity and be less time-consuming (Wildman et al., 2014; Salas & Cannon-Bowers, 2001).

Measuring

Meister (1985) presents different methods for measuring the world (or the part of the world that is of interest). He makes a distinction between subjective and objective methods. Subjective methods are the ones where the result can be affected by the creator of the method, as “the observer is the measuring instrument in subjective techniques” (p.331, *ibid*). This makes it necessary to determine accuracy validity and reliability when using subjective methods. Objective methods are, according to Meister (1985), methods lacking interpretation by the data collector when recording the data. The interpretation, which comes later, is not a crucial part of the measurement process. Åsberg, Hummerdal, and Dekker (2011) conclude that in the field of ergonomics (a research field that has had a significant impact on team cognition, both in terms of theories and methods) people are studying people. There is the self-interpretation in persons being studied, and there are “the human researchers themselves, who are, of course, constituted in a particular context that offers a particular set of constructs and methods and techniques” (p. 414, *ibid*). Nonetheless, clarifying what interpretations the researcher brings to the table and reducing the amount of interpretation reduces the risk of biases, making interpretation less coloured by subjectivity.

Level of analysis

Level of analysis in team research ranges from micro to macro (sometimes even meso level). This means from individual to organisational level over dyads, teams, and teams of teams. Regarding team research, data collection is often performed at the level of individual team members (Cooke et al., 2000; Langan-Fox et al., 2000). For example, assessing shared situation awareness using individual assessments (cf. Saner et al., 2009), which are aggregated up to a team level so that outcomes can be discussed on a team level. At other times the focus is on the team level directly (Cooke et al., 2004; Gorman et al., 2006).

Issues related to level in organisational research concern both level of measurement and level of analysis (Rousseau, 1985). There is the level of reference (in organisational research called *focal unit*). Level of measurement concerns the component that is being directly measured, for example psycho-physiological data are on the individual level, team size is on a team level). Level of analysis is the component level that is concerned for hypothesis testing and statistical analysis. To further confuse a reader, the level of analysis is not necessarily the level that is trying to be generalised to. One of the main methodological issues in multi-level research is the use of aggregated data, that is, how combined/merged data on one level can represent attributes or phenomena on a higher level. There are statistical concerns regarding how combined data change

variance and covariance, which in turn affects correlations. This might add meaning or error to the conclusions. Other problems are misspecification, aggregation bias, cross-level fallacies, and contextual fallacies. Misspecification is when an observed relationship is credited to a level other than the observed level. Aggregation bias is the extent to which a relationship is a product of the aggregation method (how data are combined). Cross-level fallacies are false generalisations from individual data to higher levels. Contextual fallacies are when the effect physical or social settings have on relationships between variables are not specified, for example how group norms affect behaviours of an individual.

Measuring team cognition and shared understanding

Different concepts have been measured differently. Below a summary of representative methods used to study team cognitive concepts is presented (see Table 4):

Table 4. Summary of different methods for measuring team cognitive concepts.

Concept	Method	Source
Shared cognition	Knowledge elicitation	Cooke et al. (2007b)
	Shared/team mental models	Cannon-Bowers et al. (1993)
		DeChurch & Mesmer-Magnus (2010a)
		Johnson et al. (2007)
		Matteson (2015)
		Maynard & Gilson (2014)
Langan-Fox et al. (2000)		
Shared situation awareness	SAGAT (situation awareness global assessment technique)	Saner et al. (2009)
	CARS	McGuinness & Foy (2000)
	Subjective assessment	

Concept	Method	Source
	Communication analysis	Nonose et al. (2010)
	Eye-movement	Hauland (2008)
Interactive team cognition	Communication analysis	Gorman & Cooke (2011)
	Video analysis	Jensen, 2009
	Observer protocols	Cooke et al. (2002)
Shared understanding	Questionnaire	Aubé et al. (2014)
		Bowman & Ambrosini (1997)
	Dialogue acts	Kim et al. (2013)
	Ranking procedure	Winner et al. (2007)
	Conversation analysis	Kwon et al. (2014)

To capture shared cognition there are several methods and techniques. In particular, methods for measuring shared mental models have gained a lot of attention (Langan-Fox et al., 2000; Mohammed et al., 2000; Wildman et al., 2014). For example, Mathieu et al. (2000) captured shared mental models letting each member in a two-person team assess their individual mental models of the task and of the team. They then assessed the sharedness by calculating convergence between the two team members using a computer algorithm to calculate the QAP⁷ correlation between the two matrices. This was done for both the team and the task mental models. They showed that the team convergence and task convergence measures correlated to team processes, and in turn to team performance. Their method for assessing the mental models was difficult. Firstly, the dimensions/factors were identified, either by using subject matter experts and/or reviewing the literature in relation to the task and situation that the participants were handling. Secondly, the participants individually responded to a matrix where all attributes were presented on both axes (see Figure 6 below) using a 9-graded scale, ranging from -4 –

⁷ The quadratic assignment procedure (QAP) correlation (Hubert & Schultz, 1976).

negatively related to 4 – positively related. The next step was to calculate mental model convergence between the two team members, which was done using the UCINET⁸ computer program to achieve a QAP correlation.

amount of information							
quality of information							
coordination of action							
roles							
liking							
team spirit							
cooperation							
	amount of information	quality of information	coordination of action	roles	liking	team spirit	cooperation

Figure 6. Matrix for responding to the team mental model (attributes taken from Mathieu et al., 2000).

Van den Bossche et al. (2006) used a self-scoring instrument to measure mutually shared cognition using two questions “At this moment, this team has a common understanding of the task we have to handle” and “At this moment, this team has a common understanding of how to deal with the task”. This approach requires considerably less time to prepare, while it has the problems of self-reported perceptions.

The examples above from Mathieu et al. (2000) and Van den Bossche et al. (2006) indicate some of the difficulties associated with assessing shared mental models: time-consuming preparation, need for subject matter experts, self-reported perceptions, complicated analysis procedures, need for specific computer programs to handle complex datasets such as response matrices. The patterns seen in these two examples are common when measuring shared mental models (DeChurch & Mesmer-Magnus, 2010a; Johnson et al., 2007; Saetrevik & Eid, 2013; Smith-Jentsch et al., 2005; Avnet & Weigel, 2012).

Many methods suggested to measure situation awareness are subjective, others are based on performance (implicit), or based on measures that are assumed to directly affect situation awareness (explicit) (Endsley, 1995a; Endsley & Garland, 2000). Subjective measures are

⁸ A network analysis program UCINET (Borgatti et al., 1992).

measures where the respondent is asked to evaluate her/his situation awareness by responding to direct questions (self-reporting). Implicit measures assumes that situation awareness correlates to performance, and improved situation awareness will lead to improved performance. Explicit measures of situation awareness are where, for example, the operator is asked about information concerning the current situation, on one or several of Endsley's (1995a) three levels (cf. SAGAT, Endsley, 1995b). These different types of method categories also occur when shared situation awareness is studied (Gorman et al., 2006).

Interactive team cognition is assessed using real-time observations of behaviour, video recordings, communication transcripts, and self reported perception of processes (Wildman et al., 2014). These methods have in common that they require a lot of time to be invested in analysis. The exception is self-reported perceptions of processes. However, as instruments for automatic analysis of video and communication recordings are becoming more and more efficient, the obstacle might be less of a problem in the future.

Shared understanding has been measured using many different methods and techniques. Aubé et al. (2014) measured perceived shared understanding using eight items. Perceived shared understanding measures to what extent an individual experiences for example *the same performance standards*. The scale items referred to agreement among team members. This approach has the disadvantages of subjective measures, i.e., they are influenced by the respondent's interpretation of the question. Kim et al. (2013) used dialogue acts to capture strength of agreement in a real-world conversation dataset to assess planning meetings. Dialogue acts represent the role of utterances. These are then tagged as questions, statements, suggestions, etc. This method used automated communication analysis and trained statistical machine learning algorithms. It does not capture and analyse what is not verbalised. Winner et al. (2007) assessed shared interpretation of the commander's intent using a course of action ranking procedure where participants' rankings were compared with a normative rank. This procedure demands a *correctly* ranked list to compare with, i.e., that there is some true course of action or behaviour. This is not always the case in complex natural settings. Bates et al. (2014) used the Patient Knowledge Assessment Tool to measure shared clinical understanding for paediatric cardiac intensive care unit patients. It is a multiple item questionnaire which is analysed to calculate the degree of agreement between assessors. The semantic alignment process by which stakeholders achieve shared understanding was studied using a linguistic approach (Corvera Charaf et al., 2013). Communication analysis was used to achieve knowledge of shared understanding in development teams. A communication analysis approach requires transcription and coding.

Approaches to measuring

The methods and techniques mentioned in Table 4 are categorised according to what extent they can reflect how dynamic the situation is, how much time is required to use the method (for preparation or analyses after data have been collected), and the degree of subjectivity that the method is associated with (see Table 5). Subjectivity here includes, apart from the degree that the result can be affected by the creator of the method, also the extent to which the respondent's interpretation is considered or can be validated against a natural world phenomenon. For example, "To what extent does your team have a shared situation awareness?" is difficult to connect to a natural world phenomenon, in contrast to "How many enemy aircraft are your team fighting?"

Table 5. Categorisation of different methods according to how dynamic they are, how much time is required, and the degree of subjectivity.

Method	Dynamic	Time	Degree of subjectivity
Knowledge elicitation	Static	Time demanding (hours to prepare)	Objective
SSM	Static	Time demanding (hours to prepare)	Objective
SAGAT	Quasi dynamic*	Time demanding (hours to prepare)	Objective
CARS	Quasi dynamic*	Requires little time (minutes)	Subjective (respondent's perception)
SSA–communication analysis	Dynamic	Time demanding (hours to analyse)	Subjective (analyser's chosen categories)
Eye-movement**	Dynamic	Time demanding (hours to analyse)	Objective
ITC–communication	Dynamic	Time demanding (hours/days to analyse)	Subjective (analyser's chosen categories)

Method	Dynamic	Time	Degree of subjectivity
ITC–video	Dynamic	Time demanding (hours/days to analyse)	Subjective (analyser’s chosen categories)
ITC–observer protocols	Dynamic	Requires little time	Subjective (observer’s perception)
SU–questionnaire	Static	Requires little time	Quasi objective***
SU–ranking procedure	Static	Requires little time	Quasi objective***
SU–conversation analysis	Dynamic	Time demanding (hours to analyse)	Subjective (analyser’s chosen categories)

* Quasi dynamic denotes that the measure can be used repeatedly to reflect change.

** The method only measures what the subject is looking at and for how long. From a situation awareness perspective it does not say anything about how the subject interprets the situation, nor how the situation is predicted to unfold. From a team perspective, it is possible to analyse when team members are looking at the same area, or when one member is looking at an area before other team members look at the same area.

*** Quasi objective means that the questions are used the same way, and not reformulated between questionnaire occasions.

For the purpose of this thesis, there are some influences which are worthwhile considering. The shared cognition perspective has some advantages, namely, there are methods that have been validated and are correlated to performance. The drawback is that concepts that are measured by shared cognition methods are rather static, and do not reflect changes due to a developing situation. Another disadvantage is that several methods require hours of preparation in order to be used in very specific situation.

The situations that the instrument developed in this thesis is aimed at being used in are developing dynamically due to situational circumstances, changing goals, and organisational constraints. As has been discussed in earlier sections of this thesis, existing techniques and methods have different limitations associated with them. Some of these limitations are due to

theoretical considerations (cf. the focus on knowledge representations for shared mental models) while others have methodological limitations (cf. CARS and subjective – the respondent's perception).

Empirical studies

In the following chapters the empirical studies for investigating the development of an instrument for assessing shared understanding are presented. In the early studies this was explored using the concepts overlap and calibration. The subsequent studies examined rank ordered lists as a means for assessing the shared understanding in teams. These rank order approaches were refined over several iterations and resulted in an instrument that measures shared strategic understanding. For a more comprehensive overview see the section Progress of the work in the Foreword chapter.

Chapter 3. Study 1: Overlap and calibration in Operation Flashpoint

Working with the armed forces and the military domain, a need for measuring teams shared understanding regarding what was important was recognised. This need from an applied domain was not met by the current assessment methods. Therefore, the start of this thesis project was founded on the need for an instrument to measure the extent to which a team shared awareness, or in the words of the early 2000s “had a shared mental model”. The initial study was carried out in 2001 after a suggested development of an instrument to capture the sharedness of a team. The suggested instrument was reported in Andersson et al. (2002a). To capture the shared mental models within a team the concepts overlap and calibration were explored. These two concepts came from three theoretical lines: shared mental models (as they were interpreted around 2000: Canon-Bowers & Salas, 2001; Langan-Fox et al., 2000; Langan-Fox et al., 2001; Mathieu et al., 2000), common ground (Clark, 1992; 1996), and transactive memory (Wegner, 1987). The present study involved teams of students, where friend teams were playing against non-friend teams.

There were several theoretical concepts that pushed this research forward, namely team and shared mental models (Canon-Bowers & Salas, 2001), common ground (Clark, 1996), and the reduced cue effectiveness (RCE) hypothesis (Andersson & Rönnerberg, 1997).

Shared mental models relate to the degree to which team members hold similar cognitive representations of a situation. “Team mental models are team members’ shared, organised understanding and mental representations of knowledge about key elements of the team’s relevant environment” (Mohammed & Dumville, 2001, p. 90).

Common ground (Clark, 1992; 1996) argues that all actions work as cues for the team (verbal and nonverbal). If cues are interpreted correctly, adaptation is easier to achieve. Thus, if one person can predict actions of others, s/he can adapt to the situation. That is, if a common ground exists (team members interpret cues in the same way), teams possess an ability to adapt. It can be argued that friends have common ground, that is, they interact regularly, creating mutual knowledge, beliefs and suppositions.

The reduced cue effectiveness hypothesis (RCE) states a) a friendship effect, (i.e., friend groups worked more effectively in terms of memory performances), and b) friends' cues work better than non-friends' cues (Andersson & Rönnerberg, 1997). Thus, if two individuals work together and communicate to achieve goals, the receiver needs to understand the information that the sender provides at a specific moment. As such, the group benefits from a mutual understanding of cues. The purpose of this study was to study how team members' degree of agreement with each other's behaviour and performance affects team performance. Two different constructs were developed to measure degree of agreement between team members. Based on the friendship effect, shared mental models, common ground, and the RCE hypothesis, we expected teams would perform better if they had a high degree of overlap and calibration. The specific purpose of our study was therefore to pin-point how overlap and calibration were related to team performance. Two aspects were studied, namely performance and behaviour. We asked the participants to predict how each one of the team members would perform and behave.

Degree of agreement can be captured in different ways. Two measurements were therefore developed. One is called overlap and one is called calibration.

Measurement of Overlap

Overlap relates to the team members' representations of the team members' knowledge and abilities, and the extent to which these representations are in agreement with each other (Andersson et al., 2002a).

If person A predicts that person B will be outstanding today and if person B predicts that person B will be outstanding today, it reveals that person A and person B overlap in their predictions about person B.

If person B predicts that he will be outstanding today, and person A thinks that person B will make that prediction, they have high "overlap" regarding their predictions about person B's performance.

This illustration reveals that person A and person B are in agreement about person B. Overlap is therefore a measure of how both of them understand a phenomenon, i.e., they agree upon the fact that person B will be outstanding today.

Measurement of Calibration

Calibration also refers to the expectations/experiences that the team members have of each other, but focuses more on an enhanced degree of understanding (Andersson et al., 2002a).

If person A predicts that person B will be outstanding today and if person B predicts that person A will predict that person B will be outstanding today, it reveals that person A and person B are calibrated to some extent. Furthermore, if person B predicts that person B will be outstanding today and if person A predicts that person B will predict that person B will be outstanding today, it reveals that they are even more calibrated.

The example above illustrates that person A and person B are calibrated. Calibration can be scored in a similar way as overlap (see result section for a detailed scoring procedure). Like overlap, the concept of calibration is intra-team degree of agreement. The calibration score is, however, based on the degree of agreement between how person A understands a phenomenon and how Person B understands how person A understands the same phenomenon. Thus, even if person A and B have different opinions about a certain phenomenon (low overlap) they can be calibrated since they have an understanding of how the other team member thinks. For example, if person A thinks that person B will be outstanding today, but at the same time, thinks that B will predict that s/he (B) will under perform today *and* person B thinks that s/he will under perform today, than person A and B are low on overlap but high on calibration.

Method

Participants

120 male students volunteered to participate in this study. Only males were used as participants in order to eliminate potential variance due to gender. There were two groups, friends and non-friends. All participants were chosen from among volunteer students who, when applying to participate, stated whom they were friends with. After that the experimenter formed teams consisting of either friends or non-friends. Each group consisted of 60 people. Groups were matched on age and computer game experience. The Friend groups' mean age was 25.7 years (SD = 4.7) and their subjectively rated gaming experience was 3.9 (SD = 1.8) on a 7 point scale (ranging from 1 to 7, where 7 indicated most experience). Friends rated that they knew each other (mean 5.0) on a 7 point scale where 1 = "I do not know the other" and 7 = "I know the

other very well". The Non-Friend groups' mean age was 26.6 years (SD = 5.1) and rated gaming experience was 3.6 (SD = 2.0) on the 7 point scale. The non-friends rated that they knew each other as a 1.2 on the 7 point scale.

Design

The study was a split-plot design (2 x 2). The first factor referred to Relationship (friends versus non-friends). The second factor referred to Question type (performance or behaviour).

Dependent variables

Several dependent variables were used. Scores on overlap and calibration were calculated for each team as were the proportion of wins and team ratio. Proportion of wins was the number of times the team had won divided by number of rounds during the experiment session. This was a team level measure. The performance team ratio score was based on number of times a participant defeated an opponent minus number of times a participant was defeated during the experiment session. Negative team ratio scores indicated that a participant was defeated more times than he defeated someone else. The team ratio measure was calculated by summing each team members' individual performance ratio scores.

Independent variables

Relationship and Question type were used as independent measures. In addition, categories based on scores on overlap and calibration, that is high and low overlap and high and low calibration, were also used.

Material

Questionnaires

Three questionnaires were used: Background Questions, Questions Before a Mission, and Questions After a Mission. The Background Questions mainly focused on the quality of the relationship between team members (i.e. how long they had known each other) and PC-gaming experience. The purpose of the Before the Mission and After the Mission Questionnaires was to measure consensus, calibration and overlap. All participants rated their own and their team members' performances and how easy it was to predict how a team member would behave/behaved on a scale from 1 to 7, where 1 = less well and 7 = really well (see Appendix 1 for details). They also rated whether they thought that their team would perform better or worse than the opposing team.

Observation sampling

Using a scorecard, the experimenter recorded who defeated whom, and also how many times each team won or lost during a session.

Apparatus/Platform

Six stationary PCs (Pentium III 800 MHz, 256 MB Ram, with a GeForce 32 MB Graphic card, and 21" screens, Microsoft Windows 98) that satisfied the requirements of the game were used. A LAN-network was used. Each team was located in a separate room. The participants were sitting in a triangle facing inwards with a computer in front of each participant (see Figure 7). Operation Flashpoint from Bohemia Interactive Studio (2001) served as a virtual environment for the teams to compete in.



Figure 7. Example of setup for one of the participating teams.

Scenario

The mission was designed as to impose a high tempo on the teams so that they were forced to take action. The instructions for the mission included a short background story and an objective for the team. The background story described how a crisis had arisen on a small island in the Mediterranean Sea. The objective was to search a small village and fight any hostile action. The

Friend team and the Non-Friend team received identical instructions and were told they were to fight team vs team.

Procedure

All six participants for one complete test session arrived at the test facility at the same time. They were assigned into a Friend team or a Non-Friend team, each team consisting of three persons. The people in the “Friend team” acknowledged they were friends, whereas the “Non-Friend team” did not. The Non-Friend teams were informed that the other team was comprised of friends and vice versa. The participants were given information about the purpose of the study and offered coffee. Next, participants completed the Background Questions. Thereafter, the two teams were separated. The Friend team was located in one room and the Non-Friend team was located in another room. After that, all participants individually played a “training mission” to learn how to interact with the virtual environment (i.e. how to move in the virtual world and how to use the characters’ tools). Thereafter, the experimenter carried out a briefing about the mission. The briefing ended with the participants answering the Before the Mission Questions individually. The experiment session (i.e., “mission”) started immediately after the questionnaire was completed. As the mission progressed, the experimenter observed and recorded who eliminated whom, using the observation sampling scorecard. As soon as one round was finished (all team members in one team were eliminated) the mission was restarted. This went on for 60 minutes, so that the number of rounds played depended on teams’ performances. After the experiment session was finished, the participants individually answered the After the Mission Questions. The participants were then debriefed, thanked, and received a cinema ticket for their participation. The entire procedure took approximately 120 minutes.

Scoring

The cells in Table 6 illustrate how the measures of overlap and calibration were calculated (This example is taken from the Before a Mission Questions). The questions in the questionnaire after a mission were identical except for changes in word tense.

Table 6. Example of participant ratings for scoring of overlap and calibration.

Answering participant	How will you perform during the scenario?	How will the other person perform during the scenario?	How do you think that the other person thinks you will perform?	How do you think that the other person thinks of his/her performance?
A	1	6	3	2
B	3	2	4	4

A matrix consists of all possible couple relations within each team with three members (i.e., couple A–B shown in Table 6, but also couple A–C, and couple B–C). The grey cells are used for scoring when A is the target. In Table 6, example above, only two persons are used for simplicity.

Overlap

Based on the data shown in the white cells of Table 6, A thinks that B will perform quite well (6); whereas B thinks that s/he will perform relatively badly (3). Thus, there is a 3 point difference in the estimates (6 minus 3). Therefore, when B is the target and A is the reference: $B_A = |6 - 3| = 3$. The absolute value was used, that is minus signs were removed. A higher value (larger difference) indicated a larger deviation from a shared picture suggesting a smaller overlap. The dependent measure for *Overlap* was the total sum of overlap between all team members. For example, A's difference score for overlap $A_{BC} = (A_B + A_C)$ + B's difference score for overlap $B_{AC} = (B_A + B_C)$ + C's difference score for overlap $C_{AB} = (C_A + C_B)$. That is, $A_{BC} + B_{AC} + C_{AB} = \text{Team}_{(ABC)} \text{ overlap}$.

The measures of overlap were initially analysed as dependent measures. Second, the distribution of teams on this measure was divided into two categories: High and Low overlap teams. These two team categories were used as an independent variable. Teams categorised as having high overlap had low difference scores and vice versa. This distinction is important for the continuing presentation of the results.

Calibration

The grey cells in Table 6 illustrate how calibration was computed. What B thinks about A's performance (2); compared with what A thinks that B believes that A rates his/her performance as(3); but also what B thinks of how A rates his/her performance (4); compared to how A rates his/her performance (1). Thus, $A_B = |2 - 3| + |4 - 1| = 1 + 3 = 4$ (differences in opinion, i.e., deviation score). The same method is used to calculate A_C . The calibration score for person A is therefore $A_{BC} = A_B + A_C$. Hence, a deviation score for calibration (A_B) is based on 4 answers. In contrast, overlap (A_B) is based on 2 answers. The dependent measure for *Calibration* was the total sum of calibration between all team members. For example, A's difference score for calibration $A_{BC} = (A_B + A_C) + B$'s difference score for calibration $B_{AC} = (B_A + B_C) + C$'s difference score for calibration $C_{AB} = (C_A + C_B)$. That is, $A_{BC} + B_{AC} + C_{AB} = \text{Team}_{(ABC)} \text{ calibration}$.

Results

Comparing the groups in the study on matching factors revealed no differences, see Table 7.

Table 7. Differences between friends and non-friends on matching factors.

	Friends (n, SD)	Non-friends (n, SD)	t-values
Age	M=25.7 (60, 5.14)	M=26.6 (60, 4.69)	.98 (ns.)
First person games	M=3.88 (60, 1.81)	M=3.60 (60, 1.96)	-.82 (ns.)
Network games	M=1.41 (60, 1.20)	M=1.34 (60, 0.98)	-.37 (ns.)

On the team level of analysis, the relationship groups were compared using team ratio and proportion of wins (see Table 9)

Overlap and calibration

In order to better understand the overlap and calibration variables, descriptive analyses were conducted and are presented in table 8.

Table 8. Scores for both overlap and calibration (mean, SD, and n) for Performance and Behaviour distributed over relationship.

Relationship and type of mutuality	Anticipated performance	Anticipated behaviour	Rated performance	Rated behaviour
Friends, overlap	M=5.42 (2.43), n=19	M=10.70 (4.37), n=20	M=5.28 (3.10), n=20	M=10.65 (4.60), n=20
Non-friends, overlap	X	X	M=7.79 (3.10), n=19	M=11.00 (5.00), n=19
Friends, calibration	M=9.89 (3.67), n=19	M=22.05 (5.38), n=20	M=9.88 (4.87), n=20	M=21.10 (8.57), n=20
Non-friends, calibration	X	X	M=13.89 (4.88), n=19	M=17.53 (5.53), n=19

The cells in table 8, for anticipated performance and anticipated behaviour are intentionally blank for non-friends since non-friends could not have an opinion regarding the other team members before the session. A lower mean indicates a higher degree of overlap. The data in table 8 show raw scores. When performing analyses, the N for within subject designs varied due to some participants not responding to a particular question.

A 2 x 2 within participant ANOVA of Session (before/after) x Question Type (performance/behaviour) was completed. This analysis was conducted only for the friend teams since the non-friend teams' "Before" cells were missing. There was a main effect of Question Type, $F_{(1,18)} = 68.34, p < .001$. No other effects were significant. Thus, friends didn't change their opinions about their team members. It was possible to differentiate between the two question types. The higher mean scores on behaviour, compared to performance, indicated lower overlap for ratings of behaviour.

Second, a 2 x 2 split plot ANOVA using ratings of performance and behaviour after the session was computed for Relationship x Question type. This analysis revealed, again, a main effect of Question Type $F_{(1,37)} = 22.98, p < .001$. Again, it was possible to differentiate between the two question types. In terms of overlap, performance generated lower deviations than behaviour. No other effects were significant.

In order to gain an understanding of the calibration variable, a descriptive analysis was conducted (see table 8). Two 2 x 2 ANOVAS were computed. The first was a 2 x 2 within-participant ANOVA of Session (before/after) x Question Type (performance/behaviour) using data from Friends only. There was a main effect of question type ($F_{(1,18)}=108.64, p<.001$). No other effects were significant. The second (split-plot) ANOVA (for rated performance and behaviour after the session) was computed for Relationship x Question type. There was a main effect of question type $F_{(1,37)}=28.59, p<.000$, and an interaction effect $F_{(1,37)}=7.47, p<.01$. The ANOVA's revealed that Friend teams were stable over time and that participants were more closely calibrated on performance questions than on behaviour questions. The interaction revealed that friends were more calibrated on performance questions than non-friends and that non-friends were more calibrated on behaviour questions than were friends.

Overlap and Calibration in relation to team performance.

The next step was to examine how team performance was related to mutual understanding when overlap and calibration were transformed into categories (see scoring section in Method section for details). However, analyses on team ratio will be presented initially since team ratio was used as the dependent measure.

Relationship and team performance measures. The effect of friendship was analysed using an ANOVA (friends x non-friends) for each dependent measure. The results are shown in Table 9. The results regarding friendship revealed a clear pattern: friends performed better than non-friends. The correlations between the dependent measures (team ratio & proportion of wins) were significant ($r = .73, p<.000$). Team ratio was chosen as the dependent measure of performance because it was believed to be a more sensitive measure than proportion of wins. Individually based performance measures were also calculated (they correlated significantly with the team based performance measures and are therefore not discussed any further).

Table 9. Mean and F values for relationship conditions on two different dependent variables.

Dependent variable	Friends	Non-friends	ANOVA
Proportion of wins	M=.68	M=.32	$F_{(1,38)}=16.11 p<.001$
Team Ratio	M=3,50	M=-5,45	$F_{(1,38)}=17.64 p<.001$

Friend teams performed better compared to non-friend teams in this non-professional task (see Table 9). Thus, the results indicate an effect of friendship on performance and that team performance measures correlated strongly.

Overlap. The high overlap teams were teams with low deviation scores; low overlap teams were teams with high deviation scores. This distribution can be seen in Table 10 below. All of the following analyses are based on the post session data only.

Table 10. Distribution of Team Ratio for overlap and calibration (M, SD, and n).

Relationship and type of mutuality	Performance		Behaviour	
	High	Low	High	Low
Friends, overlap	M=3.93 (7.6), n=14	M=2.50 (5.6), n=6	M=3.75 (6.1), n=8	M=3.67 (6.8), n=9
Non-friends, overlap	M=-6.33 (7.1), n=6	M=-5.23 (6.7), n=13	M=-7.50 (6.7), n=10	M=-2.75 (6.4), n=8
Friends, calibration	M=5.25 (7.4), n=12	M=2.67 (5.4), n=6	M=1.50 (7.8), n=8	M=6.33 (6.2), n=9
Non-friends, calibration	M=-5.83 (7.5), n=6	M=-5.25 (6.8), n=12	M=-8.00 (6.5), n=10	M=-2.86 (6.9), n=7

Negative team ratio scores indicate that a participant was defeated more times than he defeated someone else. It is important to note that the same participants or teams do not necessarily fill the same cells for different conditions. For example, one team could be categorised as High overlap for performance, and Low overlap for Behaviour. Therefore, two separate 2 (Friendship: friend/non-friend) x 2 (Category: high/low, for either performance or behaviour) between participant ANOVAs were computed.

A main effect of Relationship was obtained in both ANOVAs. Friends performed better than non-friends $F_{(1,35)}=13.79, p<.01$ (performance ratings), $F_{(1,31)}=15.93, p<.001$ (behaviour ratings). Neither the effects for high overlap versus low overlap, nor the interaction effects were

significant. However, the mean values in Table 10 suggest that friends with high overlap on performance questions outperformed all the other conditions. These mean differences are not statistically significant.

Calibration. To further investigate calibration, we divided teams into high and low calibration groups in the same way we divided teams into high and low overlap groups. This distribution can be seen in Table 10. As was the case for the overlap variable, the same participants or teams did not necessarily fill the same cells for different conditions. The first 2 x 2 (on performance ratings) between-participant ANOVA (performance) for Relationship (friend/non-friend) by calibration category (high/low), using team ratio as dependent variable, revealed a main effect of Relationship, $F_{(1,32)}=15.02$, $p<.001$, (friends mean=3.96, SD=1.73; non-friends mean=-5.54, SD=1.73). No other effects were significant. Thus, friends' were better than non-friends.

The second 2 x 2 between-participant ANOVA (on behaviour ratings) for Relationship (friend/non-friend) by calibration category (high/low), using team ratio as dependent variable, revealed a main effect of Relationship, $F_{(1,30)}=15.69$, $p<.001$, (friends mean=3.92, SD=1.66; non-friends mean=-5.43, SD=1.68). There was also a main effect of Category, $F_{(1,30)}=4.47$, $p<.05$ (high mean=-3.25, SD=1.62; low mean=1.74; SD=1.72). The interaction effect was non-significant. In contrast to the earlier findings, the effect of category of calibration indicated that low calibration behaviour teams performed better than highly calibrated teams.

Discussion

In the present study the concepts overlap and calibration were related to team performance. Both concepts concern degree of understanding among team members. The results showed no significant differences between friend teams and non-friend teams for overlap and calibration. Teams who were categorised into high overlap did not differ in performance from teams that were categorised into low overlap. The same was true for calibration – no difference in performance between high and low calibration teams.

It was evident that friend teams outperformed non-friend teams. Even if a team of friends outperformed a non-friend team it was not possible to capture this friendship effect with the proposed measures. Thus, it was not possible given the operational definitions used for mutual understanding, proposed here, to disentangle to what extent different degrees of mutual understanding (e.g. overlap and calibration) were related to team performance.

Taken together, the data suggest that degree of mutual understanding might be able to explain some of the variance in team performance. This is not clear since the results are not univocal.

One plausible explanation is that the teams studied in this experiment were not experts on the studied task. It was therefore difficult to predict how a member was to behave or to predict how a member would perform. Another plausible explanation is of course that degree of mutual understanding is unrelated to the performance outcome. None of these two explanations can, however, be verified by this experiment.

Three minor findings were also obtained in the experiment:

1. The overlap and the calibration measures for performance ratings were higher (i.e., lower deviation scores) than the overlap and calibration for behaviour ratings, indicating that it was easier for team members to rate each other's performance than to rate each other's behaviour. This was true for one-sided and double-sided measures of mutual understanding.
2. Friend teams' ratings of performance and behaviour did not change between the before and the after occasions. Thus, friends didn't change their opinions about their team mates due to participating in the study. This was true for both one-sided and double-sided measures of mutual understanding.
3. Different types of dependent measures for the team were strongly correlated. The analyses revealed that the two measures presented in this chapter correlated (Team ratio and Proportion of wins). Thus, even if the performance outcome dependent measure was varied (different types of team based measures were used) the similar pattern of not straightforward answers was obtained.

The lack of association between mutual understanding and team performance might reflect that the participants were non-professionals. They were not trained for the task they undertook. If participants had a better understanding of the requirements for a single participant to carry out this type of task, they might have been better able to take each other's perspectives. Professionals performing in their own domain might have been able to complete the ratings more accurately. The next step was decided to be that the degree of mutual understanding in professional teams should be studied.

The findings did not reveal whether the overlap and calibration measures could assess shared mental models in student teams on an artificial task in a simulated environment. The respondents perceived the questions as difficult to understand and difficult to respond to. However, the teams that consisted of friends outperformed the non-friend teams. Not differing in any other aspect than being friends, the friend teams possessed some kind of shared understanding just from

being friends that was not possible to capture with the instrument. This mutuality was consistent with Clark's (1992; 1996) theory on common ground and how it developed between people over time.

Even though friends outperformed non-friends, which was reflected in terms of performance, the instrument did not capture any difference between the two groups. Yet, the instrument had not been able to detect this. One explanation for this would be that the students were not familiar enough with each other (regarding the task behaviour), nor were they very familiar with the actual task. An assumption made at the time was that the teams did not understand their roles and tasks, or take it seriously enough.

This lack of evidence for shared awareness, although friend teams outperformed non-friend teams, led to the suggestion that the instrument should be tested in other conditions.

Acknowledgement and publication

The experimental design was planned together with Jan Andersson, and the construction of the overlap and calibration instrument was a team effort (cf. Andersson et al., 2002a). The data for this study were collected in 2001/02 with the help of Andreas Kjellin. Analyses were conducted together with Jan Andersson.

Material from this study has been published as follows:

Andersson, J., Berggren, P., Castor, M., Magnusson, S., & Svensson, E. (2002a).

Instrumentutveckling för mätning av gruppens samverkanspotential (FOI-R--0429--SE). Linköping: FOI.

Berggren, P., & Andersson, J. (2003). Team Performance in a Simulated Low-Cost Military Environment. In M. J. Cook (Ed.), *Human Factors of Decision Making in Complex Systems*. Hilton Dunblane Hydro Hotel.

Chapter 4. Study 2: Overlap and calibration at the Swedish Air Force Combat Simulation Centre

The results from the first study showed a difference in performance between teams who knew each other beforehand (friends) compared to teams where the team members did not know each other beforehand (non-friends). There were no conclusions regarding the instrument for overlap and calibration. Teams with friends did not have a higher degree of overlap nor were they more calibrated than non-friend teams. The question was then if the participants, who were students, were too unfamiliar with the task that they had carried out. One conclusion from the prior study was that it might take professionals acting in their familiar domain to be able to respond to the instrument. Hence, the same instrument was tested in a different environment – the Swedish Air Force Combat Simulation Centre (Flygvapnets Luftstrids Simulerings Centrum, FLSC)⁹, this time with trained and experienced pilots who knew each other well from working and flying together, and were also well acquainted with the simulator.

The first study tried to assess sharedness within teams in terms of overlap and calibration, without success. However, there was something there; teams composed of friends did perform better. The conclusion was that either the concepts of overlap and calibration did not measure this “sharedness”, or the questions were too difficult for the participants to answer. This could be

⁹ <http://www.foi.se/en/Our-Knowledge/Aeronautics/Swedish-Air-Force-Combat-Simulation-Centre-FLSC/Simulator-facility/> (visited September 8, 2015).

because the participants lacked experience of working together, of a professional domain, and possibly also being less experienced at assessing their own ability. Turning towards a domain of highly trained professionals – fighter pilots – was the next step. These are people who are highly trained as operators of an advanced technical platform, and used to working together.

Purpose

The purpose of this study was to test the overlap and calibration instrument with fighter pilots in an air-to-air combat simulation facility. This relates to the overall research questions of the thesis, namely, making an attempt to measure shared understanding without the downsides of existing methods, moving away from self-ratings and expert opinions, and trying to relate the measure to a dependent measure that is useful – in this case team performance. The simulation facility allowed for both control of scenario and for collection of performance measures. Example of performance measures are: number of missiles fired, number of enemy aircraft defeated, number of times own team was defeated, etc. (cf. Castor, 2009).

The research questions in this study were:

- 1) Could overlap and/or calibration assess sharedness?
- 2) Could overlap and/or calibration predict team?

The Swedish Air Force Combat Simulation Centre

Since this study was carried out at the Swedish Air Force Combat Simulation Centre a brief description and the possibilities it provides will be presented below. FLSC is a simulation facility for many-to-many manned air combat, stating on the webpage that:

“Eight pilots can, for example, fly four against four, or six against two, in a scenario that includes a large number of other rule-governed simulated units such as other aircraft, air defence systems or ships. All pilot stations are based on general models which means that, apart from JAS 39 Gripen, it is possible to fly other aircraft types at the same time in the same scenario. The facility is also capable of so-called distributed simulation in which the facility is networked with others via a secure computer link.”

The facility included eight pilot stations and also positions for fighter controllers. Training was focused on Beyond Visual Range (BVR) fighter training. Beyond Visual Range is described in Castor (2009, p. 21):

“Beyond Visual Range (BVR) means a scenario where the enemy is engaged before they can be seen visually. This is possible due to the performance of the sensors available to the formation, either their own sensors (primarily the aircraft radars), and the sensors and information available to the fighter controller. Political implications of downed aircraft and the resulting Rules of Engagement (RoE) might force a

Within Visual Range (WVR) engagement, but for the missions used in the SBA radar study, BVR engagements were in focus. Regardless of engagement scenario the pure tactical goal would be to shoot down as many as possible of the enemy aircraft, while not getting shot down yourself.”

The FLSC facility meets several research criteria (see for example Field & Hole, 2003; Meister, 1985; Cook & Campbell, 1979), such as providing a replicable environment, where the system recordings allow for structured collection of behaviour data and system outputs.

Teams and fourships

In this study, the pilots were working in teams of four with the addition of a fighter controller. As pointed out by Castor (2009), environment, technical system and the job task are all part of the work context that is characterised by being very dynamic and with high temporal speed. Only rarely do fighter pilots operate alone, only coordinating with a fighter controller. Hence, the basic building block for any air operation is a team, be it a fourship or a twoship formation.

“The fourship or the twoship formations are the basic building blocks used in a mission. Only rarely does a pilot operate by himself. In addition to the pilots, a fighter controller is an important fifth or third member of the group. The fighter controllers sit in a bunker underground or in a flying command and control aircraft, and have more powerful long-range sensors than the fighter aircraft sensors at their disposal.” (Castor, 2009, p. 25)

The fourship formation, used in this study, thus has a common goal, interacts dynamically, and is working interdependently. Hence, a fourship team meets the team definition (cf. Salas et al. 1992). The interdependence of the team can be discussed in terms of input interdependence (cf. Mathieu et al., 2001). This refers to what extent teams must share inputs such as environmental constraints, equipment, people, facilities, and information related to collective goal accomplishment. Castor addresses this in relation to the fourship formation (2009, p. 26):

“The weapons of the fourship formation can be considered one such shared resource. Every JAS 39 Gripen has four main pylons underneath the aircraft, and thus can carry, for example, four medium range missiles, apart from two short range missiles on the wing-tips. A fourship formation can thus carry a maximum of 16 medium range missiles and every missile launch decreases the strength of the unit rather considerably as the combat strength of one aircraft can be considered to decrease by 25% after each one of the missiles has been fired. The decision to launch is therefore one of the main decision points during a BVR engagement.”

Method

Participants

Eight male pilots took part in the study as a part of their regular training exercise at the training facility at the Swedish Air Force Combat Simulation Centre. The participants were informed about the research project. They were familiar with the simulated aircraft and the simulator facility. All pilots volunteered to participate. They all had experience from the training facility and were familiar with the equipment. Their descriptive statistics can be seen in Table 11.

Table 11. Descriptive statistics for participants.

Factor	Mean	SD
Age (yrs)	29.0	3.30
Flight (hrs)	803.1	557.8
Simulator (hrs)	117.5	44.6
Experience of facility (hrs)	19.4	9.5

Also, on a question regarding to what extent they knew each other (on a scale from 1 to 7, where 1 = not at all, and 7 = very well) the mean was 5.13 (SD = 1.73).

All pilots stayed in the same team (fourship formation) throughout the experiment. Each team were defenders and attackers, respectively, the same number of times (four). The fighter controller always worked with the same team. However, the fighter controllers were not taking part as participants in the study.

Design

A within-groups design was used in this study. Overlap and calibration were used, first as dependent measures to explore if pre- and post assessments would differ. Secondly, overlap and calibration were categorised into groups of high and low overlap and calibration and used as independent measures, using the proportion of returning aircraft as dependent measures. Aircraft that did not return home but were not shot down were counted as “not lost” (technical and simulator system problems of any kind). Returning aircraft is argued to be a suitable team performance measure as the pilots should not be shot down (it takes a long time and costs a great

deal of money to train fighter pilots), it reduces the possibility to reuse the aircraft during the next combat (and any Air Force only has a finite number of aircraft and pilots). Alas, since the pilots are trained to win, yet stay alive, this measure was chosen. All dependent measures are on a team level. This means that all scores are pooled over all four members in a team.

Automatic computer performances recordings: The computer logged all the pilot's actions. A large number of dependent measures were possible, but only a few were to be used for the present purposes. The measures of interest for this paper are presented in the Design section, above.

Which team was attacker and which was defender was balanced over the eight scenario sessions.

Material

Three different questionnaires were used. One was completed before the first simulated mission (background information: years of experience, team familiarity, and role in the group). Another questionnaire was completed before each mission regarding anticipated need for support, performance, and behaviour predictions (Before Questionnaire). One questionnaire was completed after each mission to collect post ratings of need for support, performance, and behaviour (After Questionnaire). Both the Before and After questionnaires were used to calculate overlap and calibration.

Apparatus/Platform

The simulator included eight pilot stations, an experiment organisation area and one room for flight control management. Each one of the pilot stations included a cockpit with displays and instruments. Head-up displays and out-of-cockpit view were projected on a large screen in front of each pilot station (approximately 120"). All pilots wore a headset allowing communication within each group of 4 pilots and the fighter controller. Moveable walls divided the stations.

Scenario

The scenarios used in this experiment was a task often used in the regular training for Swedish Air Force pilots. The scenario was a BVR combat scenario with the objective to either attack or defend an object. The task was to defend an object for one team (Team 1) and the task was to attack the same object for the other team (Team 2). The pilots were familiar with this kind of scenario and task. The scenario included an in-flight phase, an attack phase, a second wave phase, and an out-flight phase.

Procedure

The procedure will only describe experiment relevant information. All pilots and fighter controllers were gathered in a lecture room and briefed about the routines for the day. They were informed about the research project and asked to participate. The chief trainer then informed the whole group about the training sessions and the schedule for the day. After that, each team of four pilots and fighter controller had a briefing to discuss the task – that is, to plan the mission. Then the BEFORE questionnaire was answered individually by all participating pilots. The next step was that the participants went to the simulator stations and prepared for the scenario to start. After the chief trainer had cleared with each pilot that they were ready the scenario was started and the teams carried out their tasks. As soon as a participant had finished the session he received the AFTER questionnaire, that is, he had been shoot down, or the session was ended. The pilot was still seated in the cockpit and could not see the God's eye perspective¹⁰. When all pilots had completed the AFTER questionnaire, the chief trainer discussed the session with the purpose of increasing the training effects. After this discussion the chief trainer briefed about the next mission. A session took approximately 30 minutes to complete (instruction, planning, preparing simulator, questionnaires, mission completion, and training discussion). The training day was interrupted by two coffee breaks and one lunch break. A total of 8 scenarios were completed.

Scoring

The scoring of overlap and calibration was performed as described in study 1.

Results

In order to better understand the overlap and calibration variables, descriptive analyses were conducted and are presented in Table 12.

¹⁰ A point of view assuming knowledge only God would have. In this case, a view of the scenario where both sides' operational picture is presented in the same view. That is, both sides' aircraft and missiles are visible, often including data such as airspeed, direction, radar coverage, etc.

Table 12. Scores for both overlap and calibration (mean, SD, and n) for Support, Behaviour, and performance.

	Overlap	Calibration
Anticipated support	M=16.75 (5.16), n=16	M=35.31 (8.77), n=16
Rated support	M=18.36 (7.86), n=14	M=37.64 (12.86), n=14
Anticipated behaviour	M=18.00 (4.80), n=14	M=34.07 (7.73), n=14
Rated behaviour	M=15.92 (7.28), n=12	M=31.33 (11.75), n=12
Anticipated performance	M=1.56 (2.68), n=16	
Rated performance	M=1.75 (2.67), n=16	

Repeated measure ANOVAs (for Support and Behaviour) were performed to test if pre-and post assessments of overlap and calibration differed. In addition overlap pre-post test of Performance was tested. None of the analyses were significant, that is no difference was detected using repeated measure ANOVAs.

The next step was to assign the teams into groups of high or low overlap and calibration. High meant that a team had lower deviation in rating scores from each other, that is, a highly overlapped or calibrated team had a more similar view regarding, for example, need of support. Low indicated that that a team had a less similar view. The split point of whether a team was high or low was chosen so that half of the teams fitted into the high category, and half of the teams fitted into the low category. The distribution can be seen in Table 13.

After that ten one-way ANOVAs were performed to compare if high or low (overlap or calibration) could explain the proportion of returning aircraft.

Table 13. Distribution of proportion returning aircraft for overlap and calibration (mean, SD, and n).

	Variable	High	Low
Overlap, Anticipated	Support	M=.66 (.23), n=8	M=.81 (.22), n=8
	Behaviour	M=.64 (.20), n=7	M=.75 (.25), n=7
	Performance	M=.68 (.21), n=10	M=.83 (.26), n=6
Overlap, Post-rated	Support	M=.71 (.27), n=7	M=.75 (.20), n=7
	Behaviour	M=.79 (.19), n=6	M=.67 (.26), n=6
	Performance	M=.78 (.23), n=9	M=.68 (.24), n=7
Calibration, Anticipated	Support	M=.69 (.22), n=8	M=.78 (.50), n=8
	Behaviour	M=.58 (.13), n=6	M=.78 (.25), n=8
Calibration, Post-rated	Support	M=.71 (.27), n=7	M=.75 (.20), n=7
	Behaviour	M=.79 (.19), n=6	M=.67 (.26), n=6

None of the ten one-way ANOVAs were significant, that is, there was no difference in the proportion of returning aircraft that could be explained by whether a team had a high or low deviation in how support, behaviour, or performance was viewed.

Discussion

The purpose of this study was to test the overlap and calibration instrument with fighter pilots in an air-to-air combat simulation facility. However, neither overlap, nor calibration, for either support or behaviour related questions, or overlap for performance questions, indicated that the instrument could distinguish between pre and post assessments when a team had worked together to carry out a team task. These results verified the concerns that came with the previous study – the instrument did not capture shared understanding in teams. It could be because the questions did not tap into aspects that are related to the team’s outcome, or that it was too difficult for a respondent to take someone else’s perspective. Even though professionals had participated who were doing their ordinary jobs with colleagues they knew well, the outcome from the overlap and calibration assessments did not provide any insights into shared understanding in teams.

The two research questions for this study were:

- 1) Could overlap and/or calibration assess sharedness?
- 2) Could overlap and/or calibration predict team performance?

The answer to both of these questions is “no”. Regarding question 1 – could overlap and/or calibration assess sharedness? The study did not manage to assess sharedness using the concepts of overlap and calibration. This was the second study not able to confirm that the concepts could assess shared understanding, using this approach. Regarding question 2 – could overlap and/or calibration predict team performance in a professional environment? Although the FLSC simulation platform provided good measures of performance, it was not possible to say anything about team performance from the concepts of overlap and calibration.

An instrument focusing on vague aspects such as need for support, behaviour, and performance in general as a basis for team agreement did not provide helpful results. There could be several reasons for that. Teams who are in agreement about something in general do not necessarily have to perform better. That is, a team can be in agreement about something without this having a logical relation to performance. For example, all team members can agree about the team not having the skills necessary for performing the task. Hence, the teams are not “good” at doing the task about agreement. What a team agrees on should have something to do with the task and higher goals of the team’s existence, something that unifies the team in reaching shared and valued goals.

Conclusion

The purpose of this study was reached, without us getting any closer towards having an instrument that could measure shared understanding. We had learnt something about shared understanding – that the overlap/calibration instrument was not a feasible approach – this was less of a development than we had hoped for. It would have been so nice to propose/develop an instrument that would have worked instantly as intended. It was still valuable to measure shared understanding in a non-difficult fashion, that demanded less understanding about the operational environment, no subject matter experts, and which should be inexpensive to use. In addition, it was still desirable to get away from biases of self-ratings or experts/observers.

There were no results that indicated that overlap and calibration were a feasible path. Again, the proposed instrument had proven to be difficult to understand and answer for the respondents.

Acknowledgement and publication

The experimental design was planned together with Jan Andersson. The data for this study were collected in 2002 with the help of Jonathan Svensson. Analyses were conducted together with Jan Andersson.

Material from this study has been published as follows:

Andersson, J., & Berggren, P. (2003). Team performance: The relation between shared mental model measures and team performance. In M. J. Cook (Ed.), *Human Factors of Decision Making in Complex Systems*. Hilton Dunblane Hydro Hotel.

Chapter 5. Study 3: C3Fire 1 – Ranked items

In study 2, where overlap and calibration were studied at the Swedish Air Force Combat Simulation Centre, teams of fighter pilots performed in simulated air-to-air fights and responded to the overlap and calibration instrument. Again, the instrument offered inconclusive results. The conclusion was that the answers were too difficult to understand, even for professionals performing in a well-known setting. To assess shared awareness using the proposed overlap and calibration instrument was abandoned. After reconsidering the approach tested in the previous two studies (study 1 and 2), a new direction was attempted. Having carried out two experiments with the overlap/calibration tactic without success a new approach was called for. The problem had not been that the respondents were non-professionals outside a professional setting (cf. students in study 1), the instrument failed with trained fighter pilots in a tactical training facility (cf. study 2). Our main hypothesis was that the instrument was too difficult to comprehend for the respondents and the questions were too general. Another issue was that the overlap/calibration tactic had conceptually been difficult to score and analyse. The next step was then to revise how to assess the shared understanding within the team, to move from difficult-to-answer general questions towards easier-to-answer specific questions.

Microworlds

In order to have a controllable, yet relevant experimental environment, microworlds can be used. A microworld is, according to Svenmarck (1998), a “simplified task where the complexity can be manipulated and yet maintain the essential characteristics of the real-world task of interest” (p.7). Brehmer and Dörner (1993) characterise microworlds as dynamic, complex, and opaque. From the decision making perspective a microworld is a dynamic situation that calls for teamwork such

as collaborative efforts, communication, and coordination. Microworlds have been used for several decades and typically present the participants in a study with a situation that is changing dynamically in a non-transparent fashion (Brehmer, 1992; Brehmer & Dörner, 1993; Gonzalez et al., 2005; Howie & Vicente, 1998; Jobidon et al., 2006; Omodei et al., 2004). At the same time, all conditions in the microworld can be controlled by the researcher, and a number of quantifiable performance measures can be retrieved. A microworld thus offers a setting that is experienced as dynamic and complex by the participants in the study, while providing a controllable and tractable tool for the researcher. There exist several microworld platforms, for example: Lohausen (Dörner et al., 1983), Moro (Dörner et al., 1986), Newfire (Løvborg & Brehmer, 1991), D3Fire (Svenmarck & Brehmer, 1991), MS Antwerpen (Strohschneider & Gerdes, 2004), NeoCities (McNeese et al., 2005), and C3Fire (Granlund et al., 2001). Often microworlds have been used to study different aspects of teamwork, decision-making, and command and control (cf. Cooke & Gorman, 2013; Elliott et al., 2007; Fullagar & Egleston, 2008; Howie & Vicente, 1998; Sapateiro et al., 2010; Granlund et al., 2001; Johansson et al., 2003; Kylesten & Nählinder, 2011; Lindgren & Smith, 2006; Svenmarck, 1998; Gustavsson & Bäckman, 2006).

C3Fire

The C3Fire microworld environment is mainly used in command, control and communication research and in training of team decision making (Granlund, 2003). In C3Fire it is possible to record everything that is happening and create scenarios that develop according to scripts with time-initiated events. The C3Fire microworld (ibid.) was developed to allow for studies of command and control in a simulation mimicking forest fires. C3Fire can be used by either an individual or a group of people collaborating with the goal of extinguishing a simulated forest fire (Granlund et al., 2001; Granlund, 2002). The collaboration can be supported by different means of communication and it is possible to configure the simulation so that a dependency is created between the different members. The participants in a C3Fire run typically take the roles of chiefs of a number of vehicles, such as fire engines, water trucks or fuel trucks. They must then coordinate their efforts in order to fight the forest fire as efficiently as possible. The fire engines must fight the forest fire, but they depend on the water trucks to provide them with water for extinguishing the fire. Both fire engines and water trucks depend on the fuel trucks to keep moving. The interactions with the simulated, subordinate units are conducted through a geographical information system, which is an interactive electronic map (see Figure 9). C3Fire is highly configurable, making it possible for the experiment leader to design the flow of information from units in the simulation to different roles in the fire-fighting organisation (the experiment participants), as well as information flow between the participants in the experiment.

Purpose

The purpose with the following study was to change the design of the questions for assessing shared understanding. Earlier studies had proven that the previous questions were too difficult to answer and that the questions were too general.

The research question in this study was to test a new approach for assessing shared understanding. Can rank ordered lists be used to assess shared understanding?

Method

For this study we returned to using student participants to test out a new approach. Since the number of professionals is limited, this approach was used to test the new approach prior to testing it with professionals. The experiment was designed as a within-group design team experiment. In order to increase experimental control the C3Fire microworld was used.

Participants

Students from Linköping University were recruited. Seventy-two participants working in three-person teams with a mean age of 23.7 years (std dev = 3.7), where 38% were female. Regarding friendship, the participants' average rating is 4.2 (std dev = 1.9) on a 7-graded scale. The question was "how well do you know team member A and B respectively?" (1 = not at all, 7 = very well).

Design

Two experimental conditions (full view vs limited view) were used. In the full view all team members could see everything that was happening in the operational picture (the map view in C3Fire), whereas in the limited view each team member could see around their own vehicles with a line of sight of 5 x 5 cells around the vehicle. The order of the conditions was balanced over experimental runs. In addition to the full view vs limited view condition data were collected on three occasions for every scenario to see if the shared understanding developed over the course of the scenario.

There were several dependent variables used. Simulation score was calculated from the number of cells that had been affected by the fire; depending on what the cell contained a different score was given (see scoring section below). In addition, shared understanding as degree of agreement, shared situation awareness (SSA), and team performance were also used as a dependent measure. All measures were used on a team level.

Shared understanding was a measure to assess similarities within the team on a ranking task where the participants were to rank order three lists with four, five, and six items respectively. A

subject matter expert had been involved in developing the lists so that the items should reflect the situation when the scenario was paused. Three lists of different length were used to see if the number of items in the list would affect the outcome.

Shared situation awareness was measured by calculating the mean of the individual team members' perceived shared situation awareness. Team performance was measured by calculating the mean of the individual team members' perceived team performance. Alas, both shared situation awareness and team performance were aggregated measures.

Material

Three questionnaires were used. 1) A questionnaire for background information, including questions about expectations and experiences (PreGQ – Pre Game Questionnaire). 2) A questionnaire that was presented after 7, 14, and 20 minutes of the scenario (DGQ – During Game Questionnaire). This included questions about SSA, team performance, and shared understanding as degree of agreement. 3) A post session questionnaire (PostGQ – Post Game Questionnaire), including questions about outcome, collaboration, etc. The questionnaires were in Swedish.

Apparatus/Platform

To provide a complex environment with a clear team task the C3Fire microworld environment (Granlund, 2002; Granlund 2003; Granlund, Johansson, & Persson, 2001) was used. Four desktop computers with 20" screens were used. One computer and screen for each participant in the team and one computer and screen for the experiment leader to host the simulation. Figure 8 shows the physical setup for the experiment.

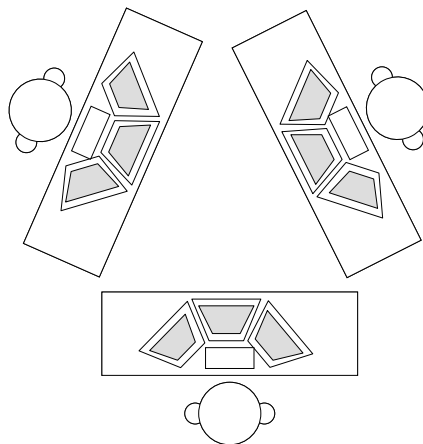


Figure 8. Physical setup of experiment equipment.

Scenario

Four equivalent fire-fighting scenarios developed for the C3Fire environment were used. The initial values of the 4 scenarios were labelled i, ii, iii, and iv. A map with a 40x40 grid was used. See Figure 9 for example of map. The scenario layout was the same except that the scenarios had been mirrored and rotated to maintain the same scenario difficulty while deceiving the participants into believing they were different scenarios.

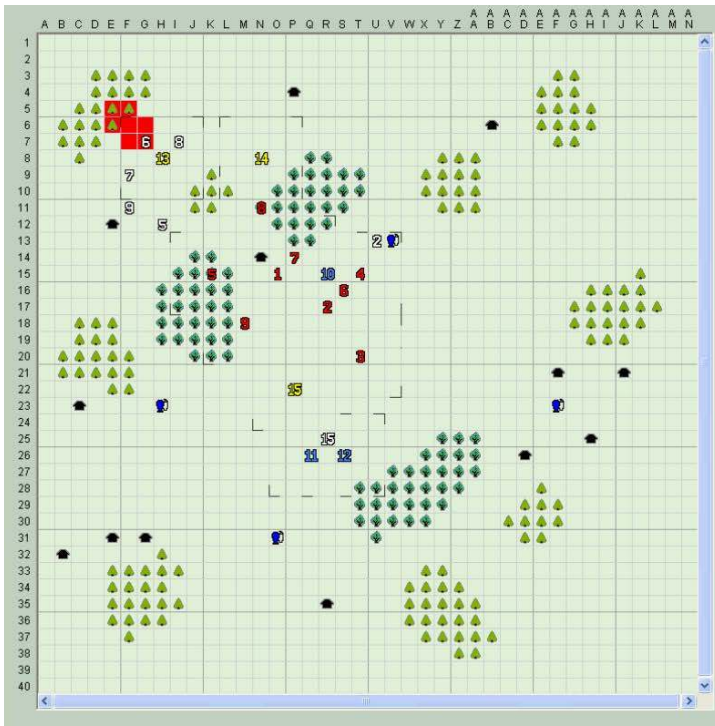


Figure 9. Screen dump of map view from C3Fire.

Each participant controlled five vehicles. Each team member had a different set of vehicles (see Table 14). Each vehicle had a field of view that corresponded to 5x5 cells around each vehicle.

Table 14. Fire-fighting vehicle distribution.

Vehicle	Participant X	Participant Y	Participant Z
Fire engines	5	2	2
Water truck		3	
Fuel truck			3

Procedure

When the participants arrived, in groups of three, they were given a cup of coffee and information about the experiment. Then there was a training session, three 5 minute training sittings, that increased in complexity and difficulty with each sitting, forcing the participants to train the skills that were needed for fighting the fire. After the training session the teams were briefed on the scenario. Then they answered the questionnaire PreGQ. The scenario was started and every 7 minutes the participants answered the DGQ. The session finished after the fire was put out or after 20 minutes of playing. The orders of the 4 sessions were balanced over all teams. The next session continued immediately. After all 4 sessions were finished the participants immediately answered the PostGQ. After that they were thanked and given a cinema ticket as compensation for their time. The experiment took approximately 120 minutes.

Scoring

Aggregating individual level data to achieve a team level measure comes with some methodological problems. There are also some practical concerns related to this, how should individual level data be combined. For the first two studies (study 1 and study 2) an analytical scoring computation derived from how dyadic comparisons within a team can be used as basis for an aggregated measure of overlap and as a measure for calibration was used (cf. Andersson et al, 2002a). Another approach is to use Kendall's measure of concordance (Kendall & Babington Smith, 1939). In order to compare more than two team members' lists of priorities, Spearman's rank order correlation could not be used. Another advantage with Kendall's W , apart from enabling m raters to rate N items, is that output is a number between 0 and 1, where 0 is no concordance at all and 1 equals a perfect fit. A correlation would produce a number between -1 and 1, something that is much harder to understand. A 1 is of course, a perfect correlation, 0 is no correlation, while -1 is a perfect negative correlation. When this is about, for example, height and weight, the relation is easy to understand. A positive correlation would be that as weight increases so does height. A negative correlation would say that as one factor increases, the other decreases – the heavier, the shorter. No correlation would say that height does not have anything to do with weight. Applying this to rank lists is a little bit trickier. The higher the “correlation” value the more similar in order the rank items. The closer to 0, the lower the similarity. However, a negative “correlation” of -1 would be that the lists are contrariwise each other. If the lists have a natural order, like size, amount or some other quantitative aspect, this could be logical. But when the content is qualitative, a -1 correlation will be confusing.

Simulation score was calculated by giving points to all cells that had been affected by fire within the scenario. A house gave 5 points, a slowly burning tree gave 2 points, a fast burning tree gave 3 points, a cell without objects gave 1 point, and a burnt-out cell gave 2 points. The sum of these points gave the simulation score for the scenario. The higher score a team achieved the worse its team performance.

Results

Firstly, the dependent measures are presented in Table 15 providing the descriptive analysis. Occasion 1, 2, and 3 are the occasions that the DGQ was responded to after 7, 14, and 20 minutes, respectively.

Table 15. Descriptive analysis of dependent variables from the C3Fire1 study (M and SD within parenthesis).

Dependent variable	Occasion	Condition	
		Full view	Limited view
W for 4-item list	1	0.57 (0.33)	0.61 (0.28)
	2	0.55 (0.26)	0.62 (0.30)
	3	0.57 (0.27)	0.62 (0.29)
W for 5-item list	1	0.56 (0.18)	0.54 (0.30)
	2	0.45 (0.25)	0.53 (0.26)
	3	0.51 (0.21)	0.51 (0.25)
W for 6-item list	1	0.61 (0.27)	0.53 (0.29)
	2	0.51 (0.26)	0.61 (0.25)
	3	0.58 (0.24)	0.54 (0.29)
Shared situation awareness	1	5.31 (1.22)	4.69 (1.05)
	2	5.50 (1.12)	4.96 (0.81)
	3	4.89 (1.05)	4.43 (1.10)
Team performance	1	5.43 (1.13)	4.85 (1.11)
	2	5.47 (1.19)	5.13 (0.87)
	3	4.86 (1.11)	4.42 (1.02)
Simulation score ¹¹		1696.6 (831.5)	1906.4 (714.1)

The next step was to correlate all dependent variables with each other. This was done using the occasion 3 data since these supposedly reflected simulation score, which was the total score for the whole scenario. The correlation matrix can be seen Table 16.

¹¹ Simulation score is presented as the total score after the full 20 minutes scenario was finished.

Table 16. Correlation matrix for dependent measures.

	W4	W5	W6	Shared situation awareness	Team performance	Simulation score
W for 4-item list	1					
W for 5-item list	.390**	1				
W for 6-item list	.524**	.483**	1			
Shared situation awareness	0.219	0.212	0.146	1		
Team performance	0.25	0.168	0.118	.905**	1	
Simulation score	0.033	0.019	0.107	-.549**	-.555**	1

** . Correlation is significant at the 0.01 level (2-tailed).

The shared understanding measures all correlate significantly with each other. However, they do not correlate with simulation score, team performance, or shared situation awareness. On the other hand, simulation score, team performance and shared situation awareness all correlate significantly with each other. That simulation score is negatively correlated to team performance and shared situation awareness is because a lower score is a better simulation performance.

The third step in the analysis was to compare the different dependent measures over the conditions using an ANOVA for repeated measure (see Table 17).

Table 17. Repeated measures ANOVA, main effects and interaction effects.

Dependent measure	Occasion	Condition	Interaction
W for 4-item list	ns.	ns.	ns.
W for 5-item list	ns.	ns.	ns.
W for 6-item list	ns.	ns.	$F_{(2,44)}=4.12, p=.023$
Shared situation awareness	$F_{(2,46)}=11.02, p<.001$	$F_{(1,23)}=7.45, p=.012$	ns.
Team performance	$F_{(2,46)}=15.59, p<.001$	$F_{(1,23)}=4.79, p=.039$	ns.

The interaction effect between occasion and condition for shared understanding on 6-item lists (seen in Figure 10) shows that when one of the conditions goes one way the other goes in the opposite direction. The main effect of occasion for shared situation awareness is seen in differences between occasion 1 and 3 (mean difference 0.34, $p=.012$ where 1 is larger), and between occasion 2 and 3 (mean difference 0.57, $p<.001$ where 2 is larger). The main effect of condition for shared situation awareness is seen in the difference between full view ($M = 5.23$, $SD = 0.22$) and limited view ($M = 4.70$, $SD = 0.17$). The main effect of occasion for team performance is seen in differences between occasion 1 and 3 (mean difference 0.50, $p=.002$), and between occasion 2 and 3 (mean difference 0.66, $p<.001$). The main effect of condition for team performance is seen in the difference between full view ($M = 5.26$, $SD = 0.22$) and limited view ($M = 4.80$, $SD = 0.18$).

Comparing the simulation score using a repeated measures ANOVA gave that $F_{(1,23)}=3.24$, $p=.085$.

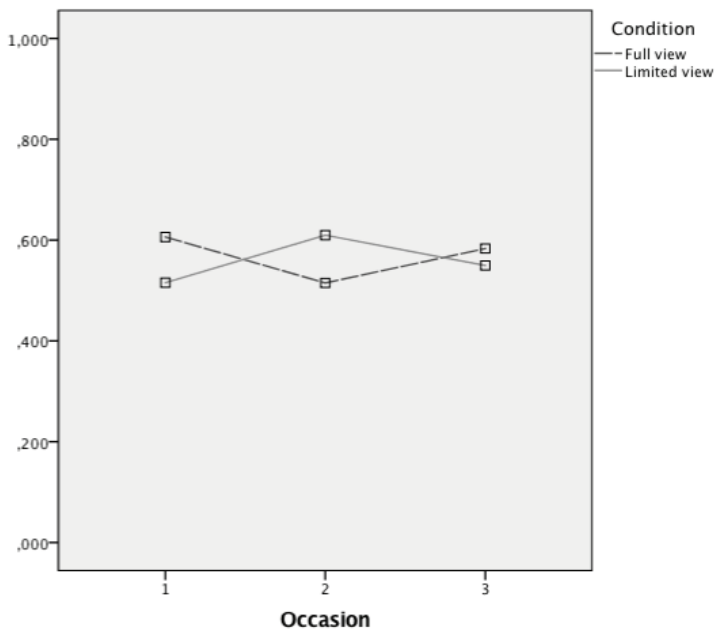


Figure 10. Plot for shared understanding using \mathcal{W} based on 6-item lists for both conditions.

Discussion

Considering team performance and shared situation awareness, the conditions can be seen as verified, i.e., in the full view condition the teams perceived that they performed better, and that they had a better shared situation awareness. The simulation score did not match these findings

as the result was not significant, even though the result pointed in the same direction. In addition, team performance, shared situation awareness, and simulation score all correlated significantly.

The shared understanding measure using pre-defined items was not able to distinguish between different conditions with different amounts of available information. Shared understanding, measured this way, did not explain performance, either as simulation score, or as perceived team performance. Shared understanding did not correlate with shared situation awareness either. All versions of the shared understanding measure (using 4, 5, or 6-item lists) correlated significantly. The number of items in the lists did not have any effect on the results, however there was a significant interaction effect for the 6-item list between condition (full view vs limited view) and occasion (7, 14, and 20 minutes into the scenario). This interaction effect is considered less important as the other shared understanding measures do not reveal the same pattern.

The C3Fire microworld had proven to provide an experimental platform that allowed for dynamic change and complexity while preserving experimental control and offering performance measures that reflected the actual behaviour in the simulation.

The conclusion here is that the instrument in this version did not work. On the other hand, some new lessons were learned:

- Generating items that are general enough to use at different times within a developing scenario might not be feasible. The results show that general pre-defined items did not capture the shared understanding within the team, a shared understanding that could be related to performance.
- There was no evident difference between the lists with different numbers of items. However, it is important to maintain a certain amount of variation to allow for statistical variance. Therefore, too few items should be avoided since rank ordering such a list might become trivial. This would provide an outcome that is always ordered in the same way.
- Preparing items prior to utilising them in a certain scenario takes a lot of time and effort, including the use of subject matter experts who understand the domain and task well. This is a cost for any assessment project or research project that needs to be considered when designing a study.

General items might not be such a good idea, even though a goal of an instrument development project should be to come up with an instrument that is useful in more than one situation. Specific items that reflect a certain situation might very well reflect the shared understanding

within a team, a shared understanding that also has an impact on performance. A problem then is to allow for lists with specific items that can be used in unspecified situations.

By letting the participating team members rank items that were important for the team to perform well, the shared understanding within the team would be possible to compute by comparing the team members' individual lists of items. These comparisons were computed using Kendall's W (Kendall & Babington Smith, 1939). In the present study teams with three members acted in an artificial environment, namely the C3Fire microworld. Lists with predetermined items were presented to the team members who were asked to rank order the items. Even though there were some shortcomings with the method, the approach had several advantages: it was easy for the participants to understand and it was easy to achieve a value that was straightforward to interpret. It was expected that a measure assessing shared understanding in this context would relate to the situation awareness concept or at least be able to say something about performance. Therefore one considered drawback was that the measure did not correlate to performance or situation awareness. Another drawback was that it took a lot of effort from subject matter experts to generate the list. One concern was the same as had been raised after the first study, that the participants were students who did not really understand the task, the skills needed to perform, or each other with respect to skills and the task. Another worry was that students who participated in an experiment might consider it as too much of a gaming situation, and consequently would not approach the task seriously – rather than consider options in order to carry out the best possible activities from the team's point of view they would try to outsmart the game and test plans without taking into consideration consequences for the team in a realistic manner.

Summary

Considering the shortcomings of earlier studies on how to capture the shared awareness within the team, the research was brought back to the laboratory with student participants. Instead of letting the participants subjectively explicate what they thought of each other's possibility to contribute to team performance, the team member's priorities were assessed. In the present study, C3Fire1, the priorities were chosen by a subject matter expert. The fire-fighting teams performed better in environments where they had better sensor range coverage compared to low sensor coverage. The results from the ranking of important factors did not reveal any significant results. Again, it was suspected that the student participants were not involved enough with the task. Would professionals performing in their own domain do better? In addition there were some experiences worth mentioning:

- The participants articulated that it was easy to rank order the factors.
- The analysis was quick to carry out using Kendall's W (Kendall & Babington Smith, 1939).
- Producing the factors took a lot of time for the subject matter expert. The factors needed to be relevant for the team and fit in with how the scenario progressed.

Acknowledgement and publication

The experimental design was planned together with Jan Andersson. Analyses were conducted together with Jens Alfredson. Scenarios were constructed and items were developed with the help of Rego Granlund. The data for this study were collected in 2006 with the help of Therese Allvar.

Material from this study has been published as follows:

Berggren, P., Alfredson, J., Andersson, J., & Granlund, R. (2008). Assessing shared situational awareness in dynamic situations. In *NATO RTO HFM-142 Symposium* (pp. 1–7). Copenhagen, DK: NATO RTO.

Berggren, P., Alfredson, J., Andersson, J., & Modéer, B. (2007). Comparing measurements of shared situational awareness in a microworld. In *Proceedings of IFAC-HMS*. Seoul, Korea.

Chapter 6. Study 4: The Tank Crew Training Facility – Shared priorities

The previous study used pre-defined rank ordered lists, which had not been successful. The use of the C3Fire microworld was a step forward, as experimental control was increased. The pre-defined lists had been time-consuming and demanded a deep understanding of the scenario and situation. Even so, the lists had not been able to reflect the actions that the teams needed to take in order to reach their common and shared goals. How could specific items be generated, without losing the possibility of using the lists in unspecified situations? Considering the outcomes of the last study it was realised that it had been advantageous to use the lists that could be rank ordered by the team members, an act that was quickly executed. The disadvantage was that the lists had taken subject matter expertise and time to generate. If the team members themselves could generate the items that the lists consisted of, several benefits could be expected; the need for subject matter experts could be eliminated, the preparation time could be shortened, and most importantly, the lists could reflect the current situation in a more precise fashion. That is, the lists would reflect the responding team members' views both in terms of items on the lists and in terms of the order of the items. Instead of forcing a worldview (the subject matter expert's view) on the participants, the team's view on the situation would be provided, a view that reflected the current situation as a consequence of the evolving situation which the team was trying to control. Such lists could also be used without specifying the situation.

The next study was carried out at the Tank Crew Training Facility (BTA – Besättningsträning) at the Swedish Armed Forces in Skövde.

Purpose

The purpose of this study was to test a method for assessing a team's shared understanding. One goal with such a method would be that it could be used by commanders, in the field, to assess their own teams. Therefore, the method could not take too much time to prepare before deployment, shouldn't require a trained research administrator to use, and should provide results that are useful for a commander as well as for research purposes.

Background

In the command and control domain the operators often have to interact with different types of technologies to coordinate with other actors on different hierarchical levels and from different organisational units. Therefore there is a real-world problem in using mixed types of technology and methods for command and control. It is of interest to assess how these mixed conditions affect the C2-teams. Today military commanders, especially on lower tactical levels, assume that the use of different C2-systems for interaction with different units influences performance negatively. Tactical command and control is a complex, time-critical, stressful, dynamic task with high risk/high stakes with changing goals (cf. naturalistic decision making; Klein et al., 1993). Using different C2-systems while performing in these environments might very well slow mission progress down; changing between different systems and cognitive mind-sets takes time. Not just due to transforming information between different media, but also since mental transformations and input of information into systems take time.

As pointed out by Salas et al. (2008), synthetic task environments do provide a context for studying team cognition that is complex and dynamic. Since the processes within the teams are complex, as are the systems they interact with, and too the dynamic environments they work in - using synthetic task environments provides a somewhat controlled setting for team research where tasks can be presented the same way over and over again.

Method

Participants

Nine officers (ranging from lieutenant to major) from the Swedish Armed Forces participated. They were all trained tank commanders. The participants were split up into three teams, each team controlling a tank. Each team consisted of a commander, a driver, and a shooter. The three tank commanders were also the command team of the tank company, one as a company commander, the other two as platoon leaders.

The two platoon leaders each commanded 3 tanks within the platoon (controlled from the scenario control room). The company commander had two tasks to carry out. The first task was to command the tank and the second was to lead the company.

Design

A within-group design was used with two conditions: condition A) same command method – all teams used the same C2-system and radio, versus condition B) mixed command methods – one team used radio, paper, map, and pen while the other teams used the C2-system and radio. Condition A was considered to be easier than condition B. The dynamic updating of the situation picture required less from the commander in condition A, compared to condition B where the commander had to use the audio information and update the paper map by hands (while still continuing with the normal tasks). Since this required more mental effort, this was considered more difficult.

Assessments were made three times for each condition, i.e., after each phase. The dependent measures were the shared priorities measure of shared understanding, subjectively assessed shared understanding, and team performance.

Material

A questionnaire with questions concerning the perceived shared situation awareness and team performance, using a 7-graded rating scale (1 meant little or poor and 7 meant much or well). A questionnaire was used to capture the shared understanding within the team using a shared priorities measure. One of the commanders was asked to generate and rank order five factors that were important for the company in that moment, to be successful in fulfilling the company's task. After he had finished, the ordered list was collected. The factors that the commander had generated were the basis for a new form in which the order of the factors was scrambled. The new form was then distributed to the remaining two commanders who were asked to rank the factors in order of importance for the company to be successful in fulfilling the company's task at that moment. In total, the administration of the rank order instrument took approximately 5 minutes on every occasion it was used.

Apparatus/Platform

The experiment was performed at the Tank Crew Training Facility at the Swedish Armed Forces in Skövde. The BTA is a simulator platform with three battle tank 122 cabins with full crew training possibility and an instructors' control room from where the scenario and computer generated forces are controlled. The tank cabins are the same size as the Swedish Tank 122

(modified Leopard 2). All three crews can train together in the same virtual environment. In each tank the positions of the commander, the pilot, the shooter, and the loader can be trained.

Scenario

The scenario was developed in collaboration with officers who trained personnel at the command training facility. The participants were engaged in a multinational peace enforcing mission and were collaborating and coordinating their actions to hinder enemy activity. It consisted of three phases; a movement phase, an observation phase, and an attack phase. The phases within the two conditions were equivalent.

Procedure

After initial information about the purpose of the experiment, the company commander gave a briefing to the platoon leaders about the plan and how he intended to solve the task. The platoon leaders then briefed their platoons. After that a training session was carried out with all participants seated at their positions in the cabins. This included a training session using the shared priorities measure. After that the first scenario was initiated (same command method). After each phase the scenario was halted and the questionnaires were answered. Then the scenario continued. After all three phases were finished and the questionnaires were answered a break was made for lunch. After lunch the scenario for the second condition started (mixed command method). The routine was the same as for the first condition. After the second condition was finished, including questionnaires, a debriefing with the instructors took place. In total the scenario took approximately 6 hours for both conditions, including questionnaire response and a lunch break.

Scoring

The team's correspondence score of the ranked factors was calculated using Kendall's measure of concordance (Kendall & Babington Smith, 1939). Shared situation awareness and Team performance was an aggregated measure of the team members' ratings, respectively.

Results

The dependent measures over the three events are presented in Table 18. Shared situation awareness and Team performance are based on the mean of the commanders' assessments of the measures respectively (*SD* in parenthesis).

Table 18. Descriptive statistics for the dependent measures.

Phase	Condition	Shared understanding (W)	Shared situation awareness	Team performance
1. Movement	Same method	0.56	3.33 (1.53)	4.00 (1.00)
	Mixed method	0.27	4.67 (0.58)	5.00 (1.00)
2. Observation	Same method	0.78	5.00 (1.73)	5.00 (1.00)
	Mixed method	0.38	5.33 (0.58)	4.33 (1.53)
3. Attack	Same method	0.62	5.00 (0)	4.67 (0.58)
	Mixed method	0.35	4.33 (0.58)	5.00 (0)

Shared understanding

The shared priorities measure could not be analysed using repeated measures analysis due to a small N. However, a few observations about the results can be pointed out. There were no roof or ceiling effects (a concern could otherwise have been that the order of factors would have been too obvious for trained professionals). Also, there seemed to be a difference (not statistically tested) between the two conditions where the agreement within the team seemed lower for the mixed command method condition than for the same command method condition. This could be expected, since it took more time and effort to coordinate actions within the team with a compromised capacity and different media for the common operational picture.

Shared situation awareness

Using the individual participants' ratings of the shared situation awareness for the whole team (the company) as dependent variable a two-way repeated measures ANOVA was performed (phase X condition). There were no main effects ($F_{\text{Condition}}(1,2)=1.00$, $p=0.423$, ns., $F_{\text{Phase}}(2,4)=2.31$, $p=0.215$, ns.), nor any interaction effect ($F(2,4)=1.50$, $p=0.327$, ns.).

Team performance

Using the individual participants' ratings of team performance for the whole team (the company) as dependent variable a two-way repeated measures ANOVA was performed (phase X

condition). There were no main effects ($F_{\text{Condition}}(1,2)=0.57, p=0.529, \text{ns.}, F_{\text{Phase}}(2,4)=0.25, p=0.790, \text{ns.}$), nor any interaction effect ($F(2,4)=4.75, p=0.088, \text{ns.}$).

Feedback on the instrument

The participants considered the instrument easy to respond to (even though it was mentioned that coming up with items could be challenging) and that it was easy to rank order the items.

Discussion

The purpose with this study was to test a method for assessing a team's shared understanding. From this point on the instrument is called shared priorities as it focuses on factors that would direct the team towards their shared goals. It turned out to be easy to use, require little preparation time, and was easily understood by the participants. However, statistical analyses were problematic as the number of participants was small. When examining the descriptive statistics in Table 18 it seemed promising.

These findings are supported with the experiences that came up during the debriefing. That is, the shared understanding was lower in the mixed methods condition than in the same methods condition. This indicated that the team (the company) had a lower degree of agreement on which actions (items) to prioritise. This could lead to degraded coordination, and in turn to worse performance. Even though the participants considered that they had acted better (a more correct behaviour) during the mixed method, according to the discussions during the debriefing, they all agreed that the time it had taken was substantially longer. Taking performance into account, that was really bad, as a tank battalion's way of handling situations is to keep the element of surprise. This is lost when within company coordination takes too much time,

This is a small step towards the developing and validating methods for shared cognition, as discussed by Salas et al. (2008). This kind of measure can be very useful to the people working in real world C2 situations; they can use the results to better understand how teams differ in shared understanding, and how different technologies and systems affect a team's shared understanding. They can also value how training and exercises affect teams.

Conclusion

Assessing shared understanding using the shared priorities instrument approach used this way was easy to deploy and easy to understand. A preliminary interpretation of the results from shared understanding seems promising, as it indicates that the instrument might be able to distinguish between the different conditions. The next step would then be to test the instrument with a bigger population.

In this study the participants themselves were requested to generate items that were important for the team to perform well, reflecting the current situation. As this was done at a pause in the simulation, the items reflected the current situation and what had led up to it. Also, the perception of what was important for the team was based on the current need and status of the involved units. Even though the number of team members was small, the results were promising (even if not too much emphasis should be put on the statistical tests). There were several positive sides of the instrument (simple to administer, easy to understand, and easy to score) and the participating team members understood the shared priorities scores.

Summary

Study four presented further evidence of the benefits of the ranking approach. The big development step was that the priority lists were generated by the participants. This captured how the team goals were perceived by the participants in the present situation. By letting the participants generate the items that were to be prioritised several issues were solved:

- There was no need to spend valuable subject matter expert time on producing the items.
- The items were per definition from the team's point of view.
- The items were related to the scenario.

Even though the number of participants was low, the instrument seemed promising. Participants said that it was easy to generate items for the lists and that it was easy to rank other members' items. Administering the instrument was easy, quick, and only bothered the participants minimally. The name *shared priorities* for the instrument started to stick.

Acknowledgement and publication

The data for this study were collected in 2008 with the help of Ulf Hörberg, Jonathan Svensson, Sandra Jonsson, & Fredrik Höglund.

Material from this study has been published as follows:

Berggren, P., & Johansson, B. J. E. (2010). Developing an instrument for measuring shared understanding. In *Proceedings of ISCRAM 2010* (pp. 1–8). Seattle, WA.

Berggren, P., Svensson, J., & Hörberg, U. (2008). *Mätning av gemensam lägesbild vid ledning på stridsteknisk och taktisk nivå. Studie genomförd på TCCS* (FOI-R--2647--SE). Linköping: FOI.

Berggren, P., Svensson, J., Hörberg, U., Jonsson, S., & Höglund, F. (2009). Shared priorities as a measure of shared understanding. *Proceedings of the Europe Chapter Human Factors and Ergonomics Society (HFES) Conference 2009*. Linköping.

Chapter 7. Study 5: C3Fire 2 – Validating the shared priorities measure

The previous study, carried out at the Tank Crew Training Facility, had indicated that using rank ordered lists with items that were generated by the participants was a promising step forward. In that study, the conditions that had been manipulated were same vs mixed methods for command. That was, the different crews used the same command and control systems and methods for tactical command and control versus one of the crews within the company which had a failing C2-system and had to use a paper map instead. The tank company commander had been interested in if it was possible to achieve results that confirmed or disproved that different methods of command would have an effect on the company's performance.

Since the number of participants at the Tank Crew Training Facility study had been small, due to accessibility of the facility and the number of available tank officers, a study was needed to confirm that the results were not due to chance. The present study was therefore designed to test the shared priorities measure in a controlled fashion, with a sufficient number of participants and using established measures of related concepts, such as mutual awareness and shared situation awareness. To do this, the laboratory was the appropriate environment to return to. Here experimental control could be sustained and a large number of participants could be found. The C3Fire microworld (Granlund, 2003) had already been used and students were a cost-effective and suitable step for testing the version of shared priorities that was developed in the previous study.

Distributed Assessment of Team Mutual Awareness (DATMA)

DATMA is a method for measuring the mutual awareness in teams (MacMillan et al., 2005). They define mutual awareness as “the extent to which team members are informed of other team members’ behaviours” (ibid., p. 51-1). DATMA utilises three different aspects of “mutual awareness” as measurements: *Taskwork Awareness*, *Workload Awareness*, and *Teamwork Awareness*. Different questionnaires are used for each of these facets. The DATMA method was chosen because it is partly based on well-validated measures (NASA TLX; Hart, & Staveland, 1988), which, together with the CARS questionnaire (McGuinness & Foy, 2000), will provide a rich representation of the studied system. DATMA is administered post-trial and therefore less intrusive than on-line or real-time questionnaires or measurement. It does not require any training on the part of the subjects, and the use of subject matter experts is limited to creating the first questionnaire. The Taskwork Awareness questionnaire focuses on the task that the team is handling. The Workload Awareness questionnaire is aimed at assessing the workload level within the team. The Teamwork Assessment questionnaire is directed at how each individual experienced the teamwork.

The crew awareness rating scale (CARS)

Salmon et al. (2009) reviewed several methods for assessing situation awareness. From their description, the method that seemed best suited for this study was the Crew Awareness Rating Scale (CARS) developed by McGuinness and Foy (2000). CARS was originally developed for generic use (McGuinness, 1999), and was easy to adapt to the C3Fire environment. Salmon et al. (2009) lists the strength of the method as; “1) Developed for use in infantry environments 2) Less intrusive than on-line techniques 3) Quick, easy to use requiring little training” (Salmon et al. 2009, p. 53). Being a post-trial questionnaire, CARS is less intrusive than freeze-probes and similar techniques, such as SAGAT (Endsley, 2000). CARS does not require subject matter experts (SMEs) for analytical purposes as other methods do, such as SABARS (Matthews & Beal, 2002) or SASHA (Jeanott et al., 2003). CARS is based on Endsley’s three-step model (Salmon et al. 2009). It uses 8 questions to assess individual SA. For each of these the participant is asked to rate themselves on a scale ranging from 1 to 4 where 1 is the best case scenario and 4 the worst case (McGuinness & Ebbage 2002). The four-point scale for the “content” ratings therefore ranges from a 1 which is a definitive and certain “Yes, I have a good SA” to 4 which is a definite and certain “No, I do not have good SA” (ibid.).

Using these kinds of questionnaires (such as CARS and DATMA) involves a number of recognised problems (Stanton et al. 2005). The collected data are based on the subjective

opinions of the participants, where each member is likely to have a different frame of reference than the others. Participants are also more likely to remember significant events and overlook events considered less important. In addition, perceived workload is likely to correlate with performance (ibid.). The advantages are that the methods are easy to employ, requiring minimal training for the participants. They are easy to score and inexpensive.

Purpose

The present study intended to validate the shared priorities measure in a more controlled experiment where the number of participants was large enough to draw valid conclusions. Also, the shared priorities measure needed to gain support in relation to other established measures of mutual awareness and shared understanding.

The overall research question was *Could the results from the BTA study be repeated?* which was split into:

- A) Does the shared priorities measure correlate to performance?
- B) Does the shared priorities measure correlate with other established measures?
 - a. Mutual awareness (DATMA)
 - b. Shared situation awareness (CARS)?
- C) Can the shared priorities measure distinguish between conditions with different levels of difficulty?

Method

Delimitation

Not all of the available questionnaire questions have been used in the analysis for this chapter (these results can be found in Berggren et al., 2011 and Prytz et al., 2010). Instead, the focus here is on the results that are central to the purpose of the thesis and to the research questions in this chapter.

Participants

Thirty-six students from Linköping University were recruited to participate in the experiment, with an average age of 23.6 years (SD=2.1), where 23 were males and 13 were females. The participants were randomly assigned into two-person teams. Average time the team members had known each other was 16.1 months (SD=14.6). All participants were Swedish citizens with native-level proficiency in Swedish. Eight participants had previous experience of C3Fire, whereas the remaining 28 participants had never used C3Fire before. Self-rated previous

computer experience was $M=4.28$ ($SD=0.61$), and self-rated previous computer gaming experience was $M=3.42$ ($SD=1.22$). Both computer and gaming experience was rated on a 5-graded scale ranging from no experience (1) to extensive experience (5). In addition, three operators were recruited to run the C3Fire scenario, two regular and one back-up.

Design

A within-group design with three conditions balanced over runs was used. The conditions were:

- i) Both commanders use the C3Fire graphical user interface (GUI) to manage units and communicate with each other and to the operators who controlled the units. This condition was called GUI/GUI.
- ii) One commander uses the C3Fire GUI while the other commander uses a regular paper map and an electronic messaging system to communicate with his/her operator and the other commander. This condition was called GUI/Map.
- iii) Both commanders use paper maps and the messaging system for communication. This condition was called Map/Map.

For each condition two participants were tasked with stopping a fire in C3Fire. The participants assumed the roles of commanders, one responsible for the fire-fighting vehicles (Commander 1) and one responsible for the water and fuel trucks (Commander 2). Each commander communicated his/her orders to a trained operator via electronic messages and the operator carried out the orders (i.e. moved fire, water and fuel vehicles) in the microworld (see Figure 11). Both participants and operators were located separately to hinder verbal communication.

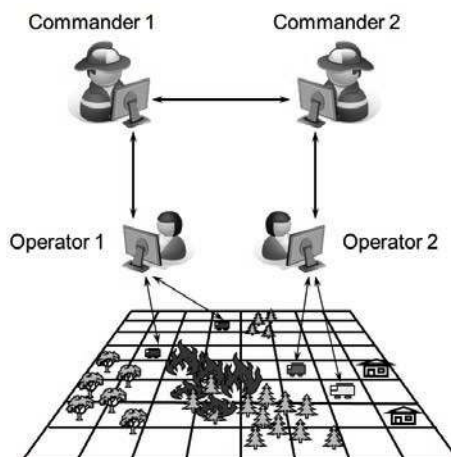


Figure 11. Experimental setup with participants (commanders) and operators.

The role of the operators in the experiment was to control the units in the microworld and, in the GUI/Map and Map/Map conditions, relay information back to the participant(s) using a map. The rationale for using operators in this experiment was that when the commanders were restricted to a paper map they would not have any way of interacting with the microworld. In order to keep the different conditions comparable operators were used in all conditions. A communication protocol was used for communication between operator and commander to ensure that the communication was consistent over trials and groups, as well as controlling what the operators could and could not do.

Dependent Measures

The dependent measures were: Simulation Score, perceived Team Performance, Shared Priorities, CARS, and DATMA (Task Awareness, Team Workload Awareness, and Teamwork Assessment).

Team performance

Simulation Score was captured from the C3Fire system recordings. In all conditions, the goal was to extinguish the fires in the microworld. More specifically, the goal was to prevent the fire from spreading and in particular spreading to buildings. Each square in the grid with a fire (burned out or active) was worth one point at the end of the scenario. If the square contained a building (a house or pump station) it was worth an additional two points. The maximum score was calculated for each scenario. The simulation performance score was then computed by using the formula $\text{simulation performance} = 1 - (\text{achieved score} / \text{maximum score})$. This gave that 0 was the worst performance, indicating that all cells that could burn out did so and 1 was the optimal performance, indicating that all cells had been rescued (which is impossible to achieve). This way of presenting the simulation performance score was considered easier to understand (to compare with study 3).

Perceived team performance was calculated using the team performance question from the DATMA Team Workload Awareness questionnaire. This question was answered on a scale, ranging from 0 to 20, which was multiplied by 5 to achieve a value between 0 and 100. The mean of the team members' responses was used as the dependent measure perceived team performance.

Material

All material was presented in Swedish. The materials used were paper maps (exact copies of the maps used in C3Fire), transparent plastic sheets and coloured markers for annotation. Scrap paper and pens were also distributed to the participants. All questionnaires but the shared

priorities instrument were printed on paper. A printout of the communication protocol was also given to each participant. Before each scenario the participants were briefed about the scenario and the conditions (for instance if their teammate used paper map or GUI).

Shared Priorities

To collect the shared priorities measure both team members were requested to individually generate five items that were important for the team to perform in the current situation. They were then asked to rank those items in order of importance. This was the target list. The experiment leader then scrambled the order of the items in the target list and the new list was redistributed to the other team member who was asked to rank the items in order of importance. The shared priorities score was calculated using Kendall's W (Kendall & Babington Smith, 1939) on each of these lists. The shared priorities score for the team was the average of the individual W values.

CARS

The CARS questionnaire was slightly modified. The questionnaire was translated into Swedish and the four specific answer alternatives (from “Yes, definitely” to “No, definitely not”) were changed into a more general four point Likert-scale with “Best possible case” at one end (1) and “Worst possible case” (4) at the other. This was done to space the alternatives in a more equidistant manner, similar to the other Likert-scales used in other questionnaires.

CARS was scored by aggregating the individual team members' judgments. That is, the average of the eight CARS questions was computed for all team members.

DATMA

The DATMA method was adapted slightly to better fit this study. The team workload questionnaire was not modified. The teamwork awareness questionnaire was also left unchanged. The taskwork awareness questionnaire was modified from asking about a salient event to, instead, mark the activities of the units in the current situation. Nine categories were developed by a subject matter expert (for more information, see Prytz et al., 2010). These were used by the participants to classify what action each unit was performing at that specific time in the scenario. Further, the participants were asked to indicate, on a paper map, the location of each vehicle as best they could remember as well as mark down each square that was burning or had been burning during the scenario up to this point.

Each DATMA subscale was scored individually. DATMA team workload awareness used the questions focused on the team level. The responses to these questions were done on a scale

(ranging from 0 to 20) which was multiplied by 5 to achieve a value between 0 and 100 on each workload subscale. The mean for all team members on the team workload awareness questionnaire (mental demand, time pressure, effort, and frustration) was used as a measure of team workload. Team performance was not included, as it was used as a measure for perceived team performance.

DATMA task awareness used two aspects; task score and position score. The maximum points a respondent could achieve for task score was 8 points, and for position score 24. To get a score that was comparable to the other task awareness aspect, task score was divided by 8 (max points) and position score was divided by 24 (max points). The mean for the task awareness aspects scores was then computed over all team members to achieve the DATMA task awareness score for the team.

DATMA teamwork assessment was computed using the mean for all team members on the included teamwork assessment questionnaire.

Apparatus/Platform

Four computers were used to run the simulation. One was a server computer that ran the C3Fire server. All computers used Windows XP Professional as operating system.

Scenario

Each scenario was functionally equivalent to allow comparisons to be made. This was achieved by rotating and mirroring the original scenario (the terrain, location of vehicles and fires) while using the same map. The map used consisted of 40x40 cells.

The participants had access to two chat/messenger windows, one for communication with the other participant and one for communicating with the operator (receiving information and giving orders).

In condition 1 (GUI/GUI) both commanders had access to the graphical user interface (GUI) of C3Fire. That meant that they could see the position of each vehicle and also get a moment-to-moment update on what the vehicles could “see”, such as fire. Each vehicle had a 3x3 line of sight while the rest of the map was covered in a “fog of war” where only the static map, i.e. terrain and structures, could be seen.

In condition 2 (GUI/Map), Commander 1 had access to the C3Fire GUI while Commander 2 had a regular paper map and a system for electronic messaging on the computer. The paper map was the same as in C3Fire using the same grid. Commander 2 was given markers in different

colours (one green, one red, one blue, and one black) that he/she could use together with plastic overhead sheets in order to annotate on top of the map. The operator provided feedback to Commander 2 regarding what the vehicles could see. Commander 1, using the C3Fire GUI, could see his/her own vehicles and their line of sight, but not the vehicles that were controlled by the other commander (unless they were within view of his/her units). In addition, Commander 2 had access to a screen providing information about fuel and water levels and what activities the units were carrying out.

In condition 3 (Map/Map) both commanders used a paper map with the same coordinate system as in the C3Fire GUI map. In this condition both commanders had the same equipment as commander 2 in condition 2 (see above).

Procedure

Initially, the participants were given a presentation about the purpose of the experiment, instructions and a demonstration of C3Fire and how to communicate with their operators. This included an explanation of performance points and the shared priorities method which was demonstrated using examples. Each participant received one paper with specific instructions for how to communicate with the operators.

The two participants were then seated in separate rooms with one computer each. The two operators were placed in a third room with one computer each. Each operator communicated with one commander. Prior to a scenario round being started (also before the training scenario) each participant was given written information about the location (coordinates) and fuel- and water status of each unit he/she controlled. This document also informed them whether they and/or their teammate were using paper map(s) or GUI, how many points each terrain type was worth and also the approximate location of the first fire.

The first training session was run using a unique map. For the training session, both commanders used the C3Fire GUI. The training session lasted 2 x 10 minutes. During this first training scenario the commanders could also see the fire at all times. This time the participants could not see the fire at all times but had to scout for it using their units.

After training, the first of the three scenarios was run. Each scenario took 15 minutes to complete. Near the end of the scenario (after 14 minutes of play) it was “frozen” and the first DATMA questionnaire was distributed. The monitors were turned off and the paper maps covered when this questionnaire was filled out. Once completed, the shared priorities data were collected. After the two questionnaires were finished the scenario was resumed for another

minute. After the 15 minutes the remaining questionnaires were distributed. The procedure was the same for all of the three scenarios.

Results

The research questions were:

- A) Does the shared priorities measure correlate to performance?
- B) Does the shared priorities measure correlate with other established measures?
 - a. Mutual awareness (DATMA)
 - b. Shared situation awareness (CARS)?
- C) Can the shared priorities measure distinguish between conditions with different levels of difficulty?

To respond to research question A and B, the correlations were computed and are presented below (see Table 19). This is based on all conditions (N=54).

Table 19. Correlation matrix for dependent measures.

	Sim. Score	Team Perf.	DATMA Team Workload	DATMA Task Awareness	DATMA Teamwork Assessment	CARS
Sim. Score	1					
Team Perf.	.691**	1				
DATMA Team Workload	-.511**	-.547**	1			
DATMA Task Awareness	0.129	0.072	-0.127	1		
DATMA Teamwork Assessment	.541**	.607**	-.547**	0.007	1	
CARS	-.675**	-.687**	.636**	0.04	-.661**	1
Shared Priorities	0.008	-0.011	0.141	0.206	0.124	-0.066

** Correlation is significant at the 0.01 level (2-tailed).

Notes: Simulations score (Sim. Score), perceived team performance (Team Perf.), DATMA team workload awareness (DATMA TWL), DATMA task awareness (DATMA Task Awareness), DATMA teamwork assessment (DATMA Teamwork Assessment), crew awareness rating scale (CARS), and shared understanding (Shared Priorities).

Shared priorities did not correlate with performance either as simulation score, or as perceived team performance. Shared priorities does not correlate with other established measures of shared awareness such as CARS and DATMA.

For research question C a repeated measures ANOVA was performed for each of the dependent measures. A descriptive analysis is presented in Table 20 and the repeated ANOVAs are presented in Table 21 below.

Table 20. Descriptive analysis of the dependent measures. Means and standard deviations, SD within parenthesis).

Measure	GUI/GUI	GUI/Map	Map/Map
Sim. Score	0.55 (0.14)	0.37 (0.10)	0.20 (0.09)
Team Perf.	70.69 (11.56)	54.03 (17.34)	41.11 (11.51)
DATMA TWL	55.625 (15.72)	69.06 (9.01)	73.40 (10.35)
DATMA Task			
Awareness	0.30 (0.10)	0.27 (0.10)	0.24 (0.12)
DATMA Teamwork			
Assessment	4.99 (0.68)	3.99 (0.66)	3.78 (0.76)
CARS	1.88 (0.34)	2.44 (0.35)	2.82 (0.30)
Shared Priorities	0.68 (0.18)	0.65 (0.21)	0.63 (0.15)

Table 21. Summary of repeated ANOVA results and the pairwise comparisons between the different conditions..

	ANOVA	GUI/GUI vs GUI/Map*	GUI/GUI vs Map/Map*	GUI/Map vs Map/Map*
Sim. Score**	$F_{(1.52, 25.84)} = 55.90,$ $p < .001$	0.18 (0.04), $p < .001$	0.35 (0.04), $p < .001$	0.18 (0.02), $p < .001$
Team Perf.	$F_{(2,34)} = 23.77,$ $p < .001$	16.67 (4.43), $p = .005$	29.58 (3.86), $p < .001$	12.92 (4.59), $p = .036$
DATMA TWL	$F_{(2,34)} = 14.24,$ $p < .001$	-13.44 (3.34), $p = .003$	17.78 (4.10), $p = .001$	-4.34 (2.88), $p = .450$ (ns)
DATMA Task Awareness	$F_{(2,34)} = 1.29,$ $p = .288$ (ns)			
DATMA Teamwork Assessment**	$F_{(1.48, 25.07)} = 19.11,$ $p < .001$	1.07 (0.18), $p < .001$	1.22 (0.18), $p < .001$	0.21 (0.27), $p = 1.0$
CARS	$F_{(2,34)} = 43.43,$ $p < .001$	-0.55 (0.11), $p < .001$	-0.94 (0.10), $p < .001$	-0.39 (0.10), $p = .005$
Shared Priorities	$F_{(2,34)} = 0.43,$ $p = .652$ (ns)			

* Mean difference and significance level is presented. ** Adjusted using Greenhouse-Geisser because of assumed sphericity.

From Table 21 it can be seen that there are some differences among the dependent measures. When both commanders have the GUI they perform better than when one commander uses the GUI and the other uses the paper map, which is better than when both commanders use paper maps. This pattern is true for both simulation score and perceived team performance, as well as for DATMA Teamwork Assessment. In addition CARS shows the same pattern, although the scale is inverted, that is, a higher number indicates a worse crew awareness. DATMA team workload awareness indicates the same pattern, even though there is no significant difference between the GUI/Map condition and the Map/Map condition (a higher value on DATMA TWL

indicates a higher workload). DATMA task awareness and the shared priorities measure did not differentiate between the conditions. In conclusion, the different measurements show that there are differences between conditions.

Discussion

In short, to respond to the research questions: the shared priorities measure does not correlate to performance. Nor does it correlate to other measures of shared understanding (DATMA and CARS). Shared priorities cannot distinguish between different conditions with different levels of difficulty.

The study showed a difference between the different conditions regarding team performance, shared situation awareness, and to some extent mutual awareness.

There are differences between different command conditions. This is a replication of the findings from the previous study (described in study 4). In the present study the team performance measures (simulation score and perceived team performance) both showed that different command methods (GUI/GUI, GUI/Map, and Map/Map) gave different levels of performance. This was similar to the previous result where the time it took for the mechanised battalion to act was delayed considerably when the commanders had to use different command methods.

One reflection was stated that could explain this result – that the teams were not trained for the task, nor used to working together. Another concern was the size of the teams, dyads might not demand as much as teams with more members in terms of coordination and collaboration.

Summary

Since the number of participants in the previous study (study 4) was low, a study to validate the findings regarding the shared priorities instrument was carried out. The purpose of the study was also to validate the shared priorities instrument against other well-documented measures of mutual awareness (DATMA; MacMillan et al., 2005) and shared situation awareness (CARS; McGuinness & Foy, 2000; McGuinness & Ebbage, 2002). The shared priorities instrument consisted of items generated by the team members. The experimental setup included three conditions where the participating teams either had access to all available information through the command and control system, where the command and control system was severely degraded for one team member, or where the command and control system was severely degraded for all team members. The experiment was carried out in the C3Fire microworld with 36 students who formed two-person teams. Results for team performance, mutual awareness, and shared SA (CARS) showed a difference between the experimental conditions. The results did not support

that the shared priorities instrument provided any useful information. That the participants had not been trained sufficiently was one plausible explanation for the results from the instrument, as the experiments with the earlier instrument had previously indicated that it provided interesting outcomes.

A reflection on the results was that since the participants differed from the tank commanders in terms of understanding of the task, training on the task, group maturity, and risk of losing face in the light of the other team members (who were both friends and colleagues) might have affected the outcome.

Acknowledgement and publication

The experimental design was planned together with Björn JE Johansson and Erik Prytz. The data for this study were collected in 2010 by Erik Prytz who implemented the scenarios into C3Fire.

Material from this study has been published as follows:

Berggren, P., Prytz, E. G., Johansson, B. J. E., & Nählinder, S. (2011). The relationship between Workload, Teamwork, Situation Awareness, and Performance in Teams: A microworld study. In *Proceedings of the Human Factors and Ergonomics Society 55th Annual Meeting* (pp. 851–855). Las Vegas, Nevada.

Prytz, E. G., Berggren, P., & Johansson, B. J. E. (2010). *Performance and shared understanding in mixed C2-systems* (FOI-R--3155--SE). Linköping: FOI.

Chapter 8. Study 6: C3Fire 3 – Trained teams

From earlier studies we have learned that: A) questions that are too complicated cannot be answered (cf. study 1 and 2), B) Rank ordered list with pre-defined items were not successful (cf. study 3), C) However, open-ended items with professional teams seemed useful (cf. study 4), whereas teams with students who were not trained could not distinguish between the different conditions using the shared priorities measure (cf. study 5).

The rationale for study 6 was to examine if trained teams would be able to use the shared priorities measure. Hence, a team experiment was designed where participants a priori were trained as teams to take on the C3fire microworld task.

Purpose

The purpose of this study was to test if trained teams had a higher degree of agreement than non-trained teams. This would be measured using the shared priorities measure.

The research questions where:

- A) Is it possible to distinguish between trained and non-trained teams using the shared priorities measure?
- B) Does the shared priorities measure correlate with performance?
- C) Does the shared priorities measure correlate with other established measures of mutual awareness and shared situation awareness?

This study used the same instruments that had been used for the previous study (cf. study 5), DATMA and CARS.

Method

Six teams were trained for 10 sessions in C3Fire prior to the experiment. The members of the trained teams were kept constant, that is, the teams were composed of the same members over all training sessions and for the experiment. The training study is described in Baroutsi et al. (2013).

Delimitation

All of the available questionnaire questions have not been used in the analysis for this chapter. Focus has been on results that are central to the purpose of the thesis and in this chapter. Also, the mid-session assessments have been excluded because there were no simulation performance scores to match with the shared priorities measurements.

Participants

Twelve teams with three members in each team participated. Six teams were trained prior to the experiment (cf. Baroutsi et al., 2013) and six teams were the control group (non-trained teams). There were 28 men and 8 women participating. The mean age of the participants was 23.9 (SD= 3.6) years. There was no significant difference between trained and non-trained teams regarding age or gaming experience.

Design

The experiment used a split plot design, with two factors: sensor range (limited view vs. full map view) x team type (trained vs. non-trained). This type of design is also called mixed within-subjects factorial design (Keppel, 1991). All analysis is performed on team level data.

Team type refers to if the team were trained or non-trained prior to the experiment. The trained team had received extensive training. The team members had been trained as a team for 10 sessions in the C3Fire microworld in scenarios resembling the experiment scenario (each training session took approximately 60 minutes including simulation, questionnaires, and discussion). Focus had been on collaboration, communication, roles, and strategies. The non-trained teams (control group) went through an introduction where they learned the functionality of the game. The training for the control group consisted of two separate rounds in C3Fire, during which they were encouraged to ask questions to better understand the environment and their task. For the first training round the full map view was used with a 40x40 cells map and the scenario was 11 minutes long. All participants had the same types of units to control: one water truck, one fuel truck, and three fire engines. This configuration was used to allow the participants to try all units. During the second training round the limited map view was used with a 60x60 cells map. The

configuration of roles was the same as during the actual experiment. This round took 7 minutes, resulting in a total of 18 minutes of practice for the non-trained teams.

Sensor range is the amount of information available to the participants during the simulation run. Two different configurations were used for this study; full view and limited view. For the full view condition all information was available on the interactive map, including locations of the other team members' units and the spread of the fire. In the limited view condition neither the locations of the other team members' units nor the spread of the fire was visible unless they were in an adjoining cell to the participant's own units (3x3 cells vision). Sensor range was a within-groups variable that was balanced over the experiment to control for any effect of order.

Dependent measures

Several dependent measures were used to assess the teams. These included Simulation score, perceived team performance, Shared priorities, Content Analysis, CARS, DATMA Team workload, DATMA Task awareness, and DATMA Teamwork assessment.

Roles

The organisation in the experiment was a flat organisation, that is, all team members could communicate with all other team members. The organisation in the experiment consisted of three roles: Fire Chief, Water Chief, and Fuel Chief. The participants choose which role to act during the experiment, a role that was then held throughout the experiment.

There were three kinds of units involved in the experiment: fire engines, water trucks and fuel trucks (see Table 22). The Fire Chief controlled six fire engines, where two were faster than the other vehicles but had smaller water tanks. The faster vehicles were intended for reconnaissance, for example, to locate new fires in the limited view condition, or for fighting smaller fires. The Water Chief controlled two fire engines and three water trucks. The Fuel Chief controlled two fire engines and three fuel trucks. The configuration forced the teams to coordinate actions.

Table 22. Role configuration and assigned vehicles.

Role	Fire engines	Fast fire engines	Water trucks	Fuel trucks
Fire Chief	4	2	-	-
Water Chief	2	-	3	-
Fuel Chief	2	-	-	3

Material

All questionnaires (see Baroutsi et al., 2013) were available in both Swedish and English. All questionnaires were printed on paper.

Performance measures

Simulation score was gathered from the C3Fire system recordings. It was calculated from the amount of cells that had been burnt down or put out. The scoring procedure is described in Baroutsi et al. (2013). The simulation score ranged from 0 (worst performance) to 1 (optimal performance).

Perceived team performance was extracted from the DATMA team workload questionnaire. It is the mean of all team members' assessment of the team's performance. The response is given on a 20-graded scale.

Shared Priorities

The shared priorities instrument was used in the same way as in study 4 and 5. All three members were asked to generate five items each and rank order these items in order of importance for the team to reach the team's shared goal. Each list was then rank ordered by the other two team members (hence three ranked ordered lists). An example is given in Table 23. Shared priorities were collected using an electronic document that automatically collected the lists, scrambled those and re-distributed them to the intended team members. This was a solution that enabled the experiment leader to quickly administer the shared priorities instrument.

Table 23. Example from experiment used for Shared Priorities and Content Analysis measures. This includes the list of items that member A generated and the rank order of those items done by member B and member C.

A's list of items	A	B	C
Save the village in the very top left.	1	2	5
Save the village in the middle left.	2	1	4
Get better communication with my teammates	3	5	3
Reorganise the units to the new areas of interest	4	4	2
Prevent any more houses/schools from catching fire	5	3	1

The shared priorities score was calculated using Kendall's W (Kendall & Babington Smith, 1939) on each of these lists. The example in Table 23 gives that $W = 0.11$. The team shared priorities score was the average of the three sets.

Content Analysis

In addition to how the previous versions of the shared priorities instrument were used, the content of the items that were generated were analysed. The aim was to see if trained teams were focusing on other aspects than non-trained teams. Content Analysis is a measure based on the content of the ranked items. That is, each statement was evaluated against defined criteria. An item had to reflect a specific future action that could be evaluated (for example, “extinguish second fire from east side”, while “put out fire” was not accepted) and have a team focus (for example, “use fast moving vehicles to scout third fire”, while “I need water” was not accepted). All team members' ranked items were evaluated, i.e., for every member there were five goals for each occasion. If an item fulfilled the criteria the team received one point. Maximum for each team and occasion was a content analysis score of 15.

Questionnaires

A background questionnaire was distributed before the experiment. A modified version of CARS was used (same as in study 5; cf. Prytz et al., 2010). The three DATMA questionnaires included Teamwork Assessment (same as in study 5; cf. *ibid.*), Team Workload Awareness (same as in study 5; cf. *ibid.*), and Task Awareness. The version of the Task Awareness questionnaire used in study 5 (cf. *ibid.*) was adjusted. The questionnaire was changed so that the participants did not only have to specify what activities his/her own units were performing, but also what activities they thought that the other members' units were performing. In study 5 (cf. *ibid.*) nine categories were used, these were reduced to seven (see Table 24). Categories “Moving to fight fire” and “Moving with other purposes than to fight fire” were altered into “Moving” and categories “Inactive due to lack of fuel and/or water” and “Inactive due to any other reason” were altered into “Inactive” because it is difficult to assess another person's intentions.

Table 24. List of activities the participants chose from for the Task Awareness assignment.

Number	Task
1	Fighting fire
2	Moving
3	Inactive
4	Refilling the unit's own fuel or water
5	Refilling another unit's fuel or water
6	Don't know
7	Other - please specify

The scoring process for the DATMA questionnaires was the same as for study 5 with exception for the task awareness questionnaire where the participant received 1 point for each type of unit assigned with the correct activity. This gave a maximum score of 16 points for each participant. The average for the three team members was calculated and used as the team's score.

Apparatus

Four computers were used to run the C3Fire simulation, including a service computer that was used to control and run C3Fire. All four computers were equipped with Windows XP Professional. C3Fire version 3.2.7 was used in this experiment. Each role was assigned to a computer. The participants were all separated by dividers to make sure that they could not see each other's screens.

Scenario

Both sensor range conditions used the same map configuration and scenario script. However, the map was rotated and mirrored to avoid familiarity. The map was 60x60 cells, whereas the interface only allows for participant to see a 40x40 cells view at a time. Over the course of 25 minutes 5 fires started within the mission area.

Procedure

The participants arrived in teams of three and were informed about the study and the purpose. They were informed about the experiment and given a written description of C3Fire and its current configuration and information. Before the experiment started they answered a background questionnaire. This was followed by a briefing on how to respond to the shared priorities instrument. The control group (the non-trained teams) went through an introduction where they learned the functionality of the game. This resulted in a total of 18 minutes hands-on practice for the non-trained teams.

The participants decided on which role to operate. All teams played two rounds in C3Fire; each round took 25 minutes. Fifteen minutes into the rounds a pause was implemented and the shared priorities instrument was administered. The round then continued. After the round was finished the participants responded to the shared priorities instrument, CARS, and DATMA. In this study only the second administration of the shared priorities instrument is used. A questionnaire about training and motivation was also used but is outside of the scope of this thesis. Map type was balanced over runs, i.e. the first team started with the full map view and then played the limited map view. The second team started with the limited map view and then played the full map view, and so on.

Analysis

Structural equation modelling (SEM) allows for analysis and modelling of complex relationships between a number of interdependent variables. LISREL (Jöreskog & Sörbom, 1993) is a statistical program for SEM-analysis. In a LISREL model, the linear structural relationship and the factor structure are combined into one comprehensive model. The relations between the latent constructs compose the structural equation model; the relationships between the latent constructs and their observable indicators or outcomes compose the factor model. All parts of the comprehensive model may be represented in a path diagram and all factor loadings and structural relationships appear as coefficients of the path. LISREL gives a series of Goodness of Fit measures of the whole model.

Results

The following research questions were addressed:

- A) Is it possible to distinguish between trained and non-trained teams using the shared priorities measure?
- B) Does the shared priorities measure correlate with performance?

- C) Does the shared priorities measure correlate with other established measures of mutual awareness and shared situation awareness?

Is it possible to distinguish between trained and non-trained teams using the shared priorities measure?

All dependent measures were analysed using a 2 (team type) x 2 (sensor range) ANOVA (seen in Table 25). For several of the dependent measures there was a main effect of team type (simulation performance, shared priorities, and DATMA Task Awareness), for workload $p=0.073$. Trained teams performed better than non-trained teams on simulation performance. Trained teams also rated themselves higher on CARS, experienced lower workload on the DATMA Team workload instrument, and scored higher on DATMA Teamwork assessment than non-trained teams

There was a significant main effect of sensor range for all dependent measures except shared priorities. The teams performed better in the full view condition. Also, the teams rated themselves higher on CARS, experienced lower workload on the DATMA Team workload instrument, and scored higher on both DATMA Task awareness and Teamwork assessment in the full view condition. For the shared priorities measure there was no main effect of sensor range.

There were several significant interaction effects present. The interaction effect for perceived team performance is presented in Figure 12, the interaction effect for CARS is presented in Figure 13, and the interaction effect for DATMA Teamwork assessment is presented in Figure 14. The pattern is that trained teams differ more between limited and full view.

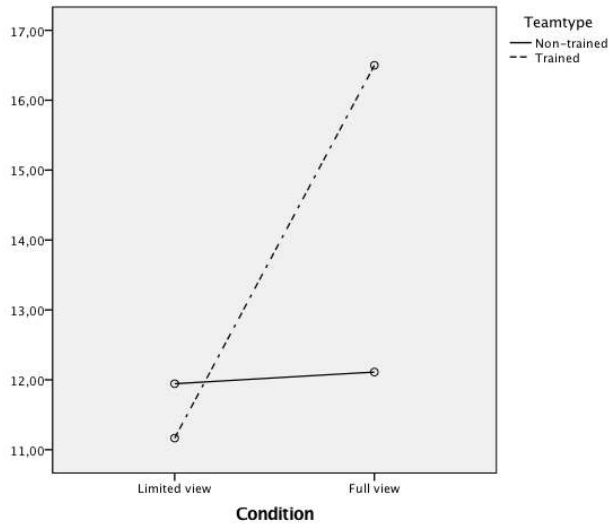


Figure 12. Interaction effects for perceived team performance.

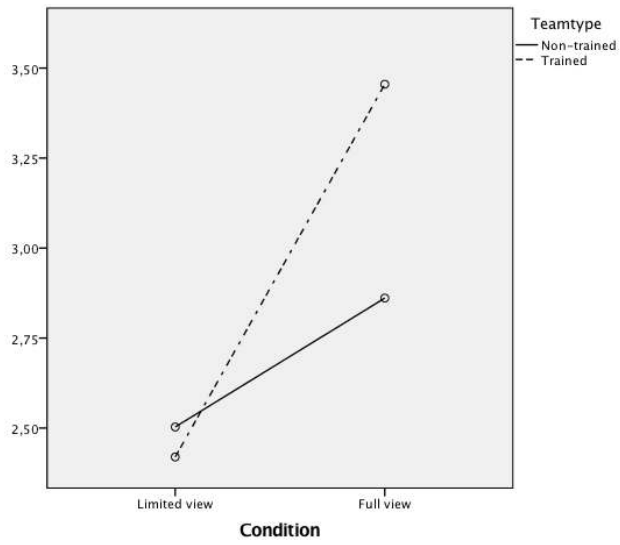


Figure 13. Interaction effects for CARS.

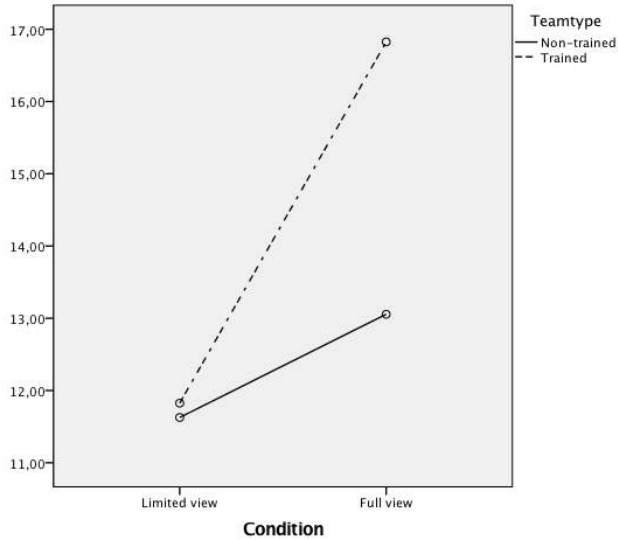


Figure 14. Interaction effects for DATMA Teamwork assessment.

Does the shared priorities measure correlate with performance?

All dependent measures were correlated. The correlation matrix can be seen in Table 26. The performance measures all correlated positively. The shared priorities measure correlated positively with simulation score. Content analysis correlated positively with both performance measures. Both shared priorities and content analysis correlate negatively with workload. Shared priorities and content analysis are strongly correlated.

Does the shared priorities measure correlate with other established measures of mutual awareness and shared situation awareness?

From Table 26 it can be seen that Shared priorities correlates negatively with the DATMA team workload measure. Content analysis, on the other hand, correlates with CARS, DATMA task awareness, DATMA teamwork assessment, and negatively with the DATMA team workload measure. Shared priorities does not correlate with CARS, nor with task awareness or teamwork assessment. CARS correlates significantly with simulation performance, task awareness, and teamwork assessment. There is also a significant negative correlation with workload.

Table 25. Summary of ANOVA results. All measures are analysed using a split plot ANOVA with sensor range as the repeated measure and team type as the independent measure.

Dependent measure	Main effect of		Trained teams,		Non-trained teams, M (SD)		Main effect of sensor range		Full view,		Limited view,		Interaction effect	
	team type	M (SD)	M (SD)	M (SD)	teams, M (SD)	sensor range	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)
Simulation performance	F (1, 10) = 15.38, p = 0.003	0.65 (0.16)	0.34 (0.21)	F (1, 10) = 29.63, p < 0.001	0.60 (0.25)	0.39 (0.19)	F (1, 10) = 1.72, p = 0.219 (n.s.)							
Perceived team performance	F (1, 10) = 3.37, p=0.096 (n.s.)	13.83 (3.49)	12.03 (2.01)	F (1, 10) = 12.98, p = 0.005	14.31 (2.58)	11.56 (2.69)	F (1, 10) = 11.45, p = 0.007							
DATMA Team workload	F (1, 10) = 4.02, p = 0.073	10.80 (2.73)	12.96 (2.03)	F (1, 10) = 13.81, p = 0.04	10.78 (3.04)	12.98 (1.48)	F (1, 10) = 3.57, p = 0.088 (n.s.)							
DATMA Task awareness	F (1, 10) = 5.28, p = 0.044	9.44 (3.35)	5.90 (2.54)	F (1, 10) = 10.65, p = 0.009	8.68 (3.64)	6.67 (3.01)	F (1, 10) = 2.14, p = 0.174 (n.s.)							
DATMA Teamwork assessment	F (1, 10) = 2.01, p = 0.178 (n.s.)	14.33 (3.09)	12.34 (3.18)	F (1, 10) = 27.82, p < 0.001	14.94 (2.68)	11.73 (2.99)	F (1, 10) = 8.59, p = 0.015							
CARS	F (1, 10) = 1.26, p = 0.288 (n.s.)	2.94 (0.64)	2.68 (0.52)	F (1, 10) = 38.20, p < 0.001	3.16 (0.48)	2.46 (0.47)	F (1, 10) = 9.02, p = 0.013							
Shared priorities	F (1, 10) = 8.44, p = 0.016	0.58 (0.20)	0.40 (0.14)	F (1, 10) = 2.93, p = 0.118 (n.s.)	0.55 (0.17)	0.43 (0.20)	F (1, 10) = 0.71, p = 0.419 (n.s.)							
Content analysis	F (1, 10) = 10.27, p = 0.009	12.17 (3.01)	6.58 (3.70)	F (1, 10) = 4.08, p = 0.071 (n.s.)	10.25 (4.49)	8.50 (4.23)	F (1, 10) = 1.12, p = 0.315 (n.s.)							

Table 26. Correlation matrix of all dependent measures. ** Correlation is significant at the 0.01 level (2-tailed). * Correlation is significant at the 0.05 level (2-tailed).

	Simulation performance	Perceived team performance	Shared priorities	Content Analysis	CARS	DATMA Team workload	DATMA awareness
Perceived team performance	0.601**						
Shared priorities	0.496*	0.336					
Content Analysis	0.637**	0.514*	0.658**				
CARS	0.536**	0.693**	0.19	0.429*			
DATMA Team workload	-0.611**	-0.490*	-0.456*	-0.490*	-0.644**		
DATMA Task awareness	0.618**	0.326	0.335	0.525**	0.457*	-0.729**	
DATMA Teamwork assessment	0.661**	0.800**	0.199	0.518**	0.908**	-0.577**	0.373

Modelling the data

Structural equation modelling is a way of reducing the complexity in the data (Svensson & Nählinder, 2014), while making sure that validity is not compromised. For example, if observations would randomly vary a significant model would not form.

By means of factor analysis and structural equation modelling (LISREL) (Jöreskog & Sörbom, 1993) a causal re-cursive model was created out of the data. Using performance variables, shared priorities, content analysis, and DATMA measures of workload a model was developed. Three latent variables were found: Shared understanding, General performance, and Team workload.

Shared understanding relates negatively to Team workload ($-.62, p < .05$) and positively to General performance ($.50, p < .05$). In turn, Team workload relates negatively to General performance ($-.48, p < .05$). The model can be seen in Figure 15 below.

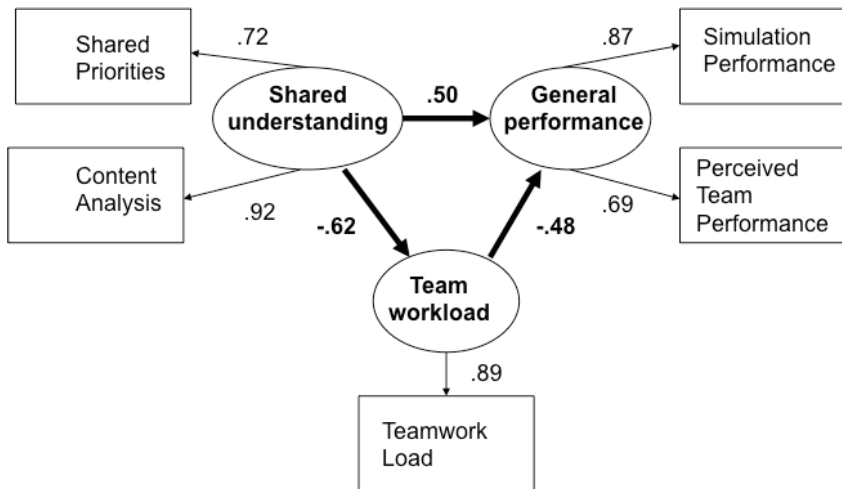


Figure 15. Model of causal relations between Shared understanding, Team workload, and General performance. Ovals represent latent variables or factors, and squares manifest variables. Thick arrows represent effects and thin arrows factor loadings. All effects and factor loadings are significant ($p < .05$).

That is, if the team members rank order items similarly and have a global view on the items this affects general performance positively and also reduces team workload. A reduced team workload also increases general performance.

The Weighted Least Square Chi-square equals 1.76 and has a p-value of 0.62. The Root Mean Square Error of Approximation (RMSEA) equals 0.00. The Adjusted Goodness of Fit Index (AGFI) was 0.94, and Comparative Fit Index (CFI) was 1.00, which means that 100 percent of the information in the input correlation matrix is explained by the model. Accordingly, the

adjustment of the model is almost perfect and the correlations between the 5 markers could fully be explained by the three different factors and their relations.

Discussion

First, a response to the research questions:

i) *Is it possible to distinguish between trained and non-trained teams using the shared priorities measure?*

Yes, the shared priorities measure and also the content analysis can separate trained teams from non-trained teams.

ii) *Does the shared priorities measure correlate with performance?*

Yes it does. The higher degree of agreement on the shared priorities measure in a team the better that team performs. The same is true for the content analysis, although even more so since the correlations are stronger.

iii) *Does the shared priorities measure correlate with other established measures of mutual awareness and shared situation awareness?*

The shared priorities measure correlates negatively with the DATMA team workload measure, that is, the higher the team workload the lower the degree of agreement on the shared priorities measure. This result might seem obvious, that the higher the stress is on the team the harder a time the team has to maintain a shared understanding.

The relation between team workload and general performance found in the model replicates patterns from earlier findings and thus provides credibility to the model (cf. Nählinder et al., 2004; Svensson & Wilson, 2002). These models show how workload correlates negatively with performance, sometimes using a situation awareness factor as mediator. Even though these models are based on the individual operator's assessments the argument has been put forward earlier (cf. Berggren et al., 2014b). In addition, the shared understanding factor in the present model adds value. It can be interpreted that a team which maintains a shared understanding will be able to coordinate actions more effectively, and thus reduce workload. Directly, this shared understanding affects performance, and indirectly by reducing workload, accordingly reducing the negative effect of workload on performance. The total effect (direct and indirect via workload) of shared understanding on performance is .80 ($p < .05$), i.e. shared understanding explains more than 60 percent of the variance of general performance.

From the model proposed in this paper, it is worth noting that in addition to subjective ratings, the simulation performance score was included. This measure is based on information in the C3Fire log files, hence an objective measure of performance. A model based on both subjective and objective measures reflects both the inner world of team members and performance reflected in events in the environment outside of the team. In statistical terms, general performance represents the shared variance of the two different measures. Although such luxuries aren't always possible when conducting measures in the field, where objective performance indicators are scarce, it does strengthen the validity of the model, given that the microworld reflects aspects of real-world tasks.

Since the model is significant, it indicates that the measures are reliable and valid. If the measures and responses fluctuated randomly between participants and collections the model would not appear. This suggests that the reliability is satisfactory (the composite or construct reliability of shared understanding is .81, and the corresponding reliability measure for general performance is .76). That the model reflects earlier findings (cf. models on the relation between workload and performance) indicates validity, as the model corresponds to an outside world represented by other data collections and other participants.

The scenario setup was successful according to the analyses of performance. This was confirmed from simulation performance, a measure that is directly linked to outcome from the simulation, and also from the subjective performance assessments done by experimenter and participants. It was possible to distinguish between an easier scenario (full view) and a more difficult one (limited view). Also, the trained teams performed better, which can be seen as a validation of the pre-experiment training of the teams. This was seen in both simulation performance and in the performance assessed by the experimenter. It is not seen in the individually assessed performance of the team, something that might be attributed the fact that individuals not trained to assess themselves are not very good at doing self-assessments (Podsakoff et al., 2003), especially with vague criteria (i.e., evaluate your team's performance on a 20-graded scale from 'very low' to 'very high'). When it comes to assessing the team in relation to sensor range, the individuals have a natural point of comparison – did they do better or worse than they did in the prior scenario round. Since team type did not change, and the teams could not see how other teams had behaved or performed, there was no natural point of comparison.

The shared priorities instrument focuses on the agreement on objectives/goals rather than perceptions of the current situation. Implicitly, this relates to level three SA, which concerns "projection of future states", but it goes beyond this as the items describe what is to be achieved

rather than what is going to happen. Intent is thus an important aspect of the shared priorities measure compared to the definition of SA as defined by Endsley (1995a). Our view is that “team awareness” emerges from the interactions of the team (cf. Cooke, et al., 2013b) in relation to reality, and also as a consequence of team (group) formation (Tuckman, 1965). From this point of view, there is no “shared situational awareness” in the sense that all team members have the same understanding of the current situation. Rather, each team member has his/her keyhole view of events, which is based on the individual’s competence, organisation of work and technological constraints (Hollnagel & Woods, 2005). From that point of view, the idea of “shared situational awareness” becomes less relevant as it actually would be contra-productive to strive towards a completely shared view of ongoing events. Instead, a well-functioning team will act as a joint system that serves the purpose of achieving a common goal, something that does not demand identical situation awareness on the individual level.

Summary

This study aimed at reiterating the previous study, with a few alterations. These were: larger teams, trained participants, and focus on the shared priorities instrument. Again both CARS and DATMA were used. The study was conducted in the C3Fire microworld environment with student participants who formed three-person teams. Six teams were trained on ten occasions, without changing the members within the teams. During the experiment the trained teams were compared to non-trained teams. In this controlled experiment the trained teams performed better and had a higher level of shared understanding than the non-trained teams. The shared priorities measure correlated to performance (measured as simulation performance – calculated from the simulation logs). The shared priorities measure did not correlate with other, well-established measures.

The model that emerged indicates that the measures are both valid and reliable, although further validation studies are required.

In a controlled study, the findings suggest that the shared priorities instrument provides valid outcomes that can, to an extent, explain performance. The experiment indicates that the instrument can be used to make a distinction between different teams regarding level of experience.

Acknowledgement and publication

The experimental design was planned together with Björn JE Johansson. The data for this study were collected in 2013 with the help of Nicoletta Baroutsi who implemented the scenarios into C3Fire. Erland Svensson helped in the statistical modelling.

Material from this study has been published as follows:

- Baroutsi, N. (2014). *Communication in Joint Activity. Investigating Teams' Communication Pattern in a Dynamic Decision Making Environment* (LIU-IDA/KOGVET-A--14/012--SE). Linköping: Linköping University.
- Berggren, P., Baroutsi, N., Johansson, B., Turcotte, I., & Tremblay, S. (2014). Assessing team focused behaviors in emergency response teams using the shared priorities measure. In *Proceedings of the 11th ISCRAM*. The Pennsylvania State University, University Park, PA.
- Berggren, P., Johansson, B. J. E., Baroutsi, N., & Dahlbäck, N. (2014a). The shared priorities measure as a way of assessing team strategic awareness – a bridge between self-assessment and the deep blue sea of field recordings. In *Proceedings of the ECCE 32nd annual conference*. Vienna, Austria.
- Berggren, P., Johansson, B., Svensson, E., Baroutsi, N., & Dahlbäck, N. (2014b). Statistical modelling of team training in a microworld study. In *Proceedings of the HFES 58th annual meeting*. Chicago, Illinois.
- Berggren, P., Johansson, B. J. E., & Baroutsi, N. (2016). Assessing the quality of Shared Priorities in teams using content analysis in a microworld experiment. *Theoretical Issues in Ergonomics Science*.

Chapter 9. Discussion

In this chapter the results are discussed. Firstly, the findings from each study are summarised. Secondly, the research questions are answered. Thirdly, how the findings can be understood in the light of the theoretical background (cf. team cognition, shared mental models, and strategic consensus) is presented. Next, methodological considerations are discussed. Lastly, future research is suggested.

Findings

Study 1 – Operation Flashpoint

In the first study, overlap and calibration were tested with three-person teams consisting of students in a first-person shooter game. A difference was found when comparing friends with non-friends regarding team performance, that is, friends outperformed non-friends. Since the teams did not differ in any other respect, the friends were assumed to have a more developed language about how to talk to each other and knowledge of each other than non-friends, who had not met before. However, this friendship effect was not seen using the overlap and calibration instrument. The respondents considered the instrument difficult to understand and respond to. One assumption was that the results regarding the concepts overlap and calibration reflected that the teams were naïve participants, i.e., this was not what they were trained for and they were not working with colleagues.

Analysing study 1 by considering the ambitions that were formulated for the purpose of the instrument (usable, understandable, objective, flexible, and self-explanatory) it could be seen that this version of the instrument was **usable**, i.e., it was easy to operate and to employ. It was not **understandable**, this audience found the instrument difficult to understand and respond to. The instrument was not **objective** as it relied on the participants' perceptions of each other (cf. "How will member A perform?"). Whether the instrument was valid and reliable is difficult to say since

there was no difference between different conditions (high vs low) for either overlap or calibration. The instrument was **flexible**; the questions about performance and behaviour could be used in different situations or domains, although data collections in field settings might prove difficult, as the participants needed some time to comprehend the questions. This version of the instrument could not be seen as **self-explanatory**. The questions were difficult to respond to and the outcome of the instrument was calculated so that the initial meaning was aggregated into a value which was difficult to relate to.

Study 2 – the Swedish Air Force Combat Simulation Centre

The second study used the same instrument as in the first study, to assess overlap and calibration in teams (fourships) of fighter pilots at the Swedish Air Force Combat Simulation Centre. The result of the second study was similarly unsatisfactory as the instrument did not reveal anything about sharedness within teams, even though the participants were highly trained professionals, working with colleagues that were well-known to them. The questions in the instrument might be too vague to capture central aspects that actually related to performance.

Analysing study 2 by considering the stated ambitions it could be seen that this version of the instrument was **usable**, i.e., it was easy to operate and to employ. It was not **understandable**, as in the prior study this audience found the instrument difficult to understand and respond to, even though these were trained fighter pilots who were used to collaborating and who knew each other well. The instrument was not **objective** as it relied on the participants' perceptions of each other (cf. "How will member A perform?"). Again it was not possible to see if the instrument was valid and reliable since there was no difference between different conditions (high vs low) for either overlap or calibration. That the instrument was **flexible** was shown using the instrument in a new setting. Again, this version of the instrument could not be seen as **self-explanatory**.

Study 3 – C3Fire1

For study 3 a new approach was tested. Explicitly, the approach was to let the participants rank order items that were pre-defined by a subject matter expert. The experiment was carried out using the microworld C3Fire, in which the participating teams were to collaborate and coordinate their actions to stop forest fires. It was not possible to differentiate between conditions using this version of the rank order instrument. The conditions were i) that the participants could see all information in the operational picture (full view), and ii) each participant could only see what was happening around the vehicles s/he controlled (limited view). The latter condition forced the team members to communicate to a larger extent in order to be able to coordinate actions and control the fires. A disadvantage of this version of the shared understanding instrument was that

it took considerable time to prepare the instrument to match the scenario and task. Nonetheless, it was perceived as quite easy to respond to.

Analysing study 3 by considering the ambitions it could be seen that this version of the instrument was somewhat **usable**, i.e., it was easy to operate and to employ after the lists had been created. However, it took quite some time to generate the items for the lists as these were to reflect the situation (which was dependent on the participants' actions) as it looked when the simulation was paused. It was **understandable**, this audience had no problems in understanding and responding to the instrument. The instrument was **objective** as focus is on how important an item is compared to other items for the team to reach its goals. As these ranked lists were compared the outcome was not the perception of an individual that was reflected, it was the shared understanding of what was important for the team to succeed. It was not possible to see if the instrument was valid and reliable since the results from the instrument could not discriminate between the different conditions (full vs limited view). This version of the instrument was not **flexible** as it needed to be adjusted to the certain situation that a team was acting in. This version of the instrument required a subject matter expert to generate the items for the lists and could not be seen as **self-explanatory**.

Study 4 – the Tank Crew Training Facility

The fourth data collection was at the Swedish Armed Forces Tank Crew Training Facility. Three 3-person teams of tank officers participated in a mechanised battle scenario that included two conditions. The first condition, characterised as easy, was that all tank commanders could use the battle management system (electronic maps, communication systems, and blue-force tracking) in full. The second condition, characterised as difficult, was that the battle management system for one of the three tank commanders was down. This meant that the commander could only communicate with other commanders over radio and plan actions using a paper map. This time the rank order instrument had been changed to let the participants themselves generate the items that were to be rank ordered. This saved time and money (the effort from the subject matter expert), and also saved time in the respect that the experimenter did not have to spend time understanding the domain. The new instrument, since then called shared priorities, seemed to be able to distinguish between two conditions. However, the number of participating teams was only 3, hence the experimental power was very low and the results needed to be interpreted with caution.

Analysing study 4 by considering the stated ambitions it could be seen that this version of the instrument was **usable**, i.e., it was easy to operate and to employ. It was **understandable**, this

audience had no problems in understanding and responding to the instrument. Like in the previous version, this version of the instrument was **objective** as focus was on how important an item was compared to other items for the team to reach its goals. It was not possible to conclude if the instrument was valid and reliable since the number of participating teams was very low. This version of the instrument was **flexible** as it did not need to be adjusted to the certain situation that a team was acting in. The instrument could be seen as **self-explanatory** as there was no need for any experts in order to use the instrument.

Study 5 – C3Fire2

This study was planned to validate the findings from the Tank Crew Training Facility in a controlled laboratory experiment. The shared priorities instrument was used together with CARS and DATMA to validate against the outcome variable (simulation performance). CARS and DATMA relate to team performance (cf. MacMillan et al., 2005; McGuinness & Foy, 2000). Earlier studies have shown that a relation exists between performance and the CARS and DATMA measures respectively (Salmon et al., 2006; Entin & Entin, 2000). The experiment was run with 2-person teams of students in the C3Fire microworld. No significant effects were found for the shared priorities instrument. Shared priorities could not distinguish between different conditions, and the outcome of the instrument did not correlate with CARS or DATMA. A reflection was stated that could explain this result – the teams were not trained for the task, nor were they used to working together. Another concern was the size of the teams, dyads might not demand as much as teams with more members in terms of coordination and collaboration.

Analysing study 5 by considering the ambitions it could be seen that this version of the instrument was **usable**, i.e., it was easy to operate and to employ. It was **understandable**, this audience had no problems in understanding and responding to the instrument. Like in the previous version, this version of the instrument was **objective** as focus was on how important an item was compared to other items for the team to reach its goals. The results from the instrument could not distinguish between the different conditions. This version of the instrument was **flexible** as it did not need to be adjusted to the certain situation that a team is acting in. This instrument, just like in study 4, could be seen as **self-explanatory**.

Study 6 – C3Fire3

The sixth study was aimed at, in part, repeating the fifth study. Alterations to the setup were that the teams consisted of more members (3 persons instead of 2 persons) and that 50% of the teams were extensively trained. The aim was to investigate if the shared priorities instrument could be used to detect differences between conditions and also to validate it against CARS and

DATMA. The experiment platform used was the C3Fire microworld using a mixed design with four conditions; sensor range (full sensor range vs limited sensor range) and team type (trained vs non-trained teams). The sensor range conditions could be considered more and less difficult to handle. This meant that the participants had full access to information presented in the operational picture vs where the participants could only see around the vehicles that they controlled themselves. This time it was possible to distinguish between the trained teams and the non-trained teams using the shared priorities instrument outcome. Trained teams had a higher score than non-trained teams. The shared Priorities instrument outcome correlated with performance. CARS and aspects of DATMA could distinguish between the two sensor range conditions but did not correlate with the shared priorities instrument. Hence, the shared priorities instrument explains different aspects of performance than CARS and DATMA. The structural model, computed from the data in study 6, reveals two things: (A) how the concept shared understanding relates to team workload and to general performance, and (B) that the instrument is valid and reliable, providing credibility to the instrument.

Analysing study 6 by considering the ambitions it could be seen that this version of the instrument was **usable**, i.e., it was easy to operate and to employ. It was **understandable**, this audience had no problems in understanding and responding to the instrument. Like in the previous version, this version of the instrument was **objective** as focus is on how important an item is compared to other items for the team to reach its goals. The results from the instrument could distinguish between the different team types (trained vs non-trained), while not being able to distinguish between the view conditions (full vs limited). The statistical model (see Figure 15) showed that the instrument was valid and reliable. This version of the instrument was **flexible** as it does not need to be adjusted to the certain situation that a team was acting in. The instrument could be seen as **self-explanatory**.

The research questions

This section provides answers to and discuss the research questions.

1. How can shared understanding be measured without the disadvantages of existing methods?

The downsides are that existing measures take time, need to be prepared by subject matter experts, that the analysis is complicated, and that the instruments/techniques are difficult to employ in real world situations, hence providing techniques mainly suitable for laboratory settings.

Introducing the concept of having team members generate individual items to be ranked by other team members resolves several issues. Fewer persons (and less working hours) are needed to prepare material suited for a particular situation as the need for subject matter experts is eliminated. This also reduces the cost associated with preparation. The analysis, using Kendall's W , is quite straightforward and provides easily interpretable results. There is no need for the administrators of the data collection to interpret the list of items created by the participants and the results; these are best understood by the participating team members who have both the domain knowledge and understanding of the developing situation. The shared priorities instrument is an instrument that should be possible to use in both laboratory experiments and quasi-experimental field settings. This is due to the fact that the technique is not limited to situations that are fully controlled and where every change is completely understood by the data collection administrator(s). As the change and development are captured using the team members' perspective, there is no need for the administrators of the data collection to understand this completely. In addition, the instrument does not require extensive training to use nor to understand the outcome from the assessments.

2. How can shared understanding be assessed without the bias of self-ratings and/or experts/observers?

The second research question is answered by looking at all the studies that were carried out. During the studies, the need for subject matter experts to be involved in the use of the instrument has been reduced. Hence, using the shared priorities removes bias from subject matter experts.

Using self-ratings to capture shared understanding, for example when letting respondents answer questions such as "To what extent does your team have a shared understanding of phenomenon X?", only reflects the individual's perceptions of the shared understanding. If self-rating measures actually measure shared understanding, or if it is a (possibly distorted) personal view, is difficult to disentangle. Even when the team members' perceptions are combined to achieve a team level measure it is still just an aggregated perception and still not clear if it reflects the actual shared understanding or not. In the case of the shared priorities instrument, a measure of the shared strategic understanding within the team is achieved (regarding the created items) by calculating Kendall's W .

3. Can team performance be better understood by the outcomes of an instrument that measures shared understanding?

A basic assumption of this work aligns with the words written by Mohammed and Dumville (2001:89) “The general thesis of the shared mental model literature is that team effectiveness will improve if team members have an adequate shared understanding of the task, team, equipment, and situation”. The strength of many measures and techniques used to measure shared/team mental models is that they correlate with team performance. Study six provides results where outcomes from the shared priorities instrument correlate with performance. Here, higher scores on the shared priorities instrument correlate with better team performance. This connection between shared understanding, measured with the shared priorities instrument, and performance is important as it demonstrates that shared understanding and performance are empirically related to each other.

General discussion

Revisiting the ambitions with the instrument development (usable, understandable, objective, flexible, and self-explanatory) over the different instrument iterations, a summary is presented below (see Table 27).

Table 27. Ambitions of the instrument related to the different studies.

	Study 1	Study 2	Study 3	Study 4	Study 5	Study 6
Usable	Yes	Yes	Yes	Yes	Yes	Yes
Understandable	No	No	Yes	Yes	Yes	Yes
Objective	No	No	Yes*	Yes*	Yes*	Yes
Flexible	Yes	Yes	No	Yes	Yes	Yes
Self-explanatory	No	No	No	Yes	Yes	Yes
Correlation with performance	No	No	No	-	No	Yes
Correlation with other measures	No	No	No	-	No	Yes

* The results from the instrument could not distinguish between different conditions.

This shows that all of the ambitions have been reached over the instrument design iterations for the subsequent studies.

Another perspective on the results is to ask what the results would have looked like if the final version of the shared priorities instrument could have been used in study 1, 2, and 3. Study 1, where teams of friends battled teams of non-friends (all students) in a computer game, probably wouldn't have shown any significant results regarding the shared strategic understanding. This conclusion is based on the findings in study 5 where non-trained teams' results on the shared priorities instrument did not reveal any significant effects. To have co-trained teams seem to be a pre-requisite for achieving non-random scores when using the shared priorities instrument. Study 2, where teams of fighter pilots combated other teams of fighter pilots, could have shown a difference in performance between teams who scored higher on the shared priorities instrument compared to teams who scored lower. This due to the findings in study 6 where shared priorities correlated to simulation performance. Study 3, where 3-person teams stopped wildfires in the C3Fire microworld, would probably have the same issues as study 5, that the teams were not co-trained sufficiently to show any significant effects of the shared priorities instrument.

Examining what types of teams have participated another pattern is visible, indicating that one cell of a four square grid is not covered (seen in Table 28).

Table 28. Distribution of studies regarding if the teams consist of trained participants or not, and if the members know each other or not.

	Professional	Non-professional
Co-trained	2, 4	6
Not (co-)trained		1, 3, 5, 6

This cell concerns professionals (persons exercising their professional role, who have domain knowledge and clearly defined role) who are not trained together, hence they do not know each other. This would for example be crisis response ad-hoc teams where the members of the team are not familiar with each other while they have clearly defined roles and responsibilities.

Theoretical implications

To let team members rank order items is a technique to capture sharedness on a team level of analysis. The measure is not assessing the actual knowledge in the team. This is a measure of team cognition that is not based on the team's behaviour (cf. communication), instead the measure is based on the similarity of how the team members perceive the strategic goals and how

to reach these. The individual member's perception of the strategic goals probably controls how each individual team member acts and tries to coordinate actions with other team members, while the similarity or deviation in the whole team's shared understanding affects team performance. This approach is closely related to the interactive team cognition perspective described in the theoretical background above (Cooke et al., 2013b; Wildman et al., 2014). The interactive team cognition perspective states that team interaction *is* team cognition, that team cognition is directly observable and directly linked to team effectiveness. Interactive team cognition is also constructed dynamically, in real time, through team members' interaction (Gorman & Cooke, 2011). The other view of team cognition, presented in the theoretical background is shared cognition (*ibid.*), where team cognition is considered in the form of shared knowledge representations. Here, a shared mental model is about how similar the team members' models of the team as a system are. The knowledge underlying these structures can be argued to change at a slower pace than a situation is evolving. For example, what each team member knows about another team member's set of skills and knowledge (for example, a person's ability to solve maths problems) probably does not change from minute to minute. On the other hand, a complex situation unfolds dynamically, changing from minute to minute (for example, the development of a wildfire). The shared priorities instrument considers the strategic goals and how these are prioritised. This can be seen as a bridge between the shared cognition perspective and the interactive team cognition perspective (see Figure 16). The circles (lower left side) represent the individual's knowledge structures. These knowledge structures are the input that a team member uses as prior understanding when interpreting the situation and the common goals of the team and how to prioritise these. In addition these shared priorities may influence an individual's knowledge structures (hence the double directed left arrow). The figures (lower right side) represent the team members' interactions within the team. These interactions affect how goals are viewed, the other team members' needs, and how actions a coordinate actions within the team. The individual team member's view on priorities of goals is also used in the interaction with other team members (the double directed right arrow). On one hand the shared priorities method has similarities with assessing shared mental models (in terms of similarity between knowledge structures), on the other hand the shared priorities method captures to what extent the understanding of the strategic goals are shared on a team level – something that is only present on a team level and emerging from the interactions within the team. Complementary to this, the awareness among team members raised by the articulated shared strategic understanding might affect the knowledge structures on whom to collaborate with and which actions to

coordinate. Hence, the outcome of the shared priorities instrument might be responsive to changes in the situation and within the team.

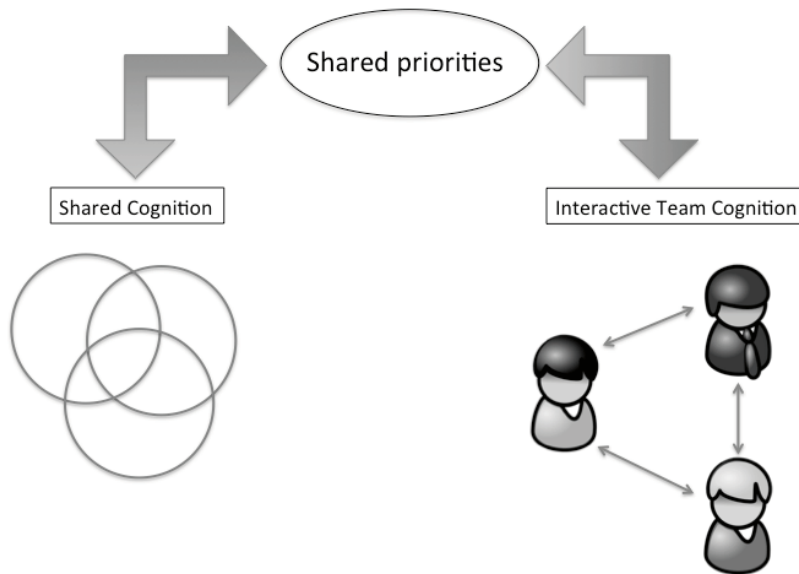


Figure 16. Using shared priorities as a bridge between shared cognition and interactive team cognition.

A potential bridging effect of the shared priorities method is also that the shared understanding of the team goals will help coordinate actions within the team to move closer towards the team goals. This shared understanding is affected by the actions carried out by the team members, and as such changes the order of which goal is most important. When the team has achieved one goal, another goal becomes more important as the situation develops. Hence, the shared priorities instrument is sensitive to this kind of more rapid change, compared to phenomena that change slowly (cf. knowledge structures regarding individual knowledge as measured in terms of shared mental models).

The shared priorities instrument is focusing on factors in the situation that the team members can have an opinion about, whereas both the CARS and DATMA methods are directed at the subjective experiences that the individual team members have. As was seen in study 6, the shared priorities instrument captures other aspects than CARS and DATMA teamwork assessment. What these aspects are is a question that is not covered in this work. CARS is based on Endsley's (1995a) model of situation awareness, a model that is focusing on the present situation. Deriving from the situation that fighter pilots are facing, situation awareness is focusing on here and now, and the projected future states in the fighter pilot domain are a couple of minutes away ("Will I have time to avoid that missile"?, "If I fire this missile at my opponent will that force him/her

into a less advantageous situation”?). The shared priorities instrument is focusing further away, on the shared goals of the team. These goals are less sensitive to what is happening in the present situation. In the same study (study 6) it was shown that the more mature teams (the trained teams) were better at formulating the priorities that concerned team level aspects (cf. content analysis) and also had a better shared understanding than the non-trained teams.

Shared strategic understanding

The choice to introduce the concept *strategic* with shared understanding was a consequence of the results that indicated that the shared priorities instrument was measuring something else than what situation awareness measures or what shared mental models did. Measures capturing situation awareness is focused on assessing the current situation to predict the development. Since the shared priorities instrument is addressing team goals further ahead in time, these can be considered as strategic for the team. Here strategic is used in the sense long-term, on a time scale relative the phenomenon which the team is attempting to handle. In comparison to shared mental models, the shared priorities instrument captures aspects of team cognition that is sensitive to dynamic change, whereas the shared mental models mainly deal with knowledge structures that are changing at a slower pace.

The definition of shared understanding suggested by Smart et al. (2009) is used as a starting point for a definition of shared strategic understanding. Hence, shared *strategic* understanding is *the ability of multiple agents to exploit common bodies of causal knowledge over time for the purpose of accomplishing common (or shared) goals.*

This definition indicates that members (agents) who have a shared understanding will be able to coordinate their behaviours to reach the shared goals, acknowledging that shared goals might change over time. That is, to maintain understanding and direction towards the team's shared goals. By adapting the definition of Smart et al. (ibid), both the interactive, co-construction of understanding of what team members see as important enablers of success and the dynamic reconstruction of the same understanding are recognized.

Ecological validity

A consideration has been how to take the findings from lab experiments and make use of them in applied settings. How can findings from studies with teams consisting of students be generalised to be used by teams with professionals who have been trained for years, working in specific naturalistic environments? Mathieu et al. (2000) say that it is the underlying construct relationships that corresponds to other situations, and also:

In effect, we argue that while the use of low-fidelity simulations limits the point-to-point correspondence that is critical for direct generalisability to real-world settings, they do provide for a controlled examination of the critical factors influencing team processes and performance in dynamic environments. In this sense we are interested in studying the construct interrelationships and their applicability to teams in other situations. (p.280)

In the pursuit to develop the theories on team cognition it is essential to have techniques and instruments that, in addition to being valid and reliable, are usable in the field. The area of team cognition is dependent on data from the applied settings, as many areas in this field are rooted in applied problems, and the complexity in these settings is difficult to truthfully represent in the laboratory. Hence, data often need to be collected in the wild, where the phenomena are found and where the teams' behaviours are real, to really understand what is happening and why. The shared priorities instrument seems to have the characteristics that allow for being used in both laboratory- and applied settings, which need to be tested in future studies. In relation to the participants in a field study, other attractive features are that the participants supply their own content using their own terminology, content that is appropriate for the situation, instead of the researcher imposing a predetermined perspective.

Instrument rationale

Traditionally, instruments and measures in the field of Human Factors/team cognition are based on a theoretical model (cf. Endsley, 1995a; Lewis, 2003). In the case of shared priorities the instrument development has been driven by an applied need, where empirical results drive progress. That is, the requirements and aim of the instrument were defined, after which the instrument was iteratively developed to reach these requirements and goals.

The thesis argues that the shared priorities instrument is domain, hierarchical level, and theoretically, independent. That is, the instrument seems to be independent of which domain it is used in or on what hierarchical organisational level a team is located. Also, the instrument does not bring a lot of theoretical assumptions along; it can be used in harmony with most team/group theories. Therefore, as a user of the instrument you do not have to accept (concealed) assumptions that follow with the use of the instrument. For example, the theory of situation awareness (Endsley, 1995a) considers the human operator as an information-processing unit. The concept is not directly measurable in the environment as it is something that occurs inside the head of the operator. As such, the concept can only be measured indirectly. A similar claim can be made for mental workload, a concept that is indirectly measured in various ways (cf. Young et al., 2015). The shared priorities instrument is based on the assumption that teams that

are trained together prioritise team goals in a similar way, but it does not make claims about the internal cognitive processes of the team members. Other assumptions are that it is possible to formulate priorities in the domain in focus, that the respondents can verbalise priorities, and that they are willing to share them.

The main idea of this thesis was to develop an assessment method for assessing shared understanding. Why measure? Methods and techniques to measure and assess concepts are central in order to test models and hypotheses, and in turn advance theories within a field. Empirical testing of a hypothesis without proper assessment techniques is a waste of time and money. If instruments and assessment approaches are not valid and reliable, the conclusions drawn from studies using such instruments are not valid. Sound theory development depends on instruments that are valid and reliable.

Criticism

The teams participating in the studies presented in this thesis consist of relatively few members. Real-world teams range from small teams, like the ones in the studies, to relatively large teams consisting of many members, for example sports teams, gaming teams in on-line multiplayer games, large project teams, etc. Many of those kinds of teams can also be viewed as teams of teams, which is a different research area (cf. Davison et al., 2012; DeChurch & Zaccaro, 2010; Marks et al., 2005). Since the teams in the studies presented in this thesis consist of teams with between 2 and 4 members this limits the generalisability to small teams. However, the instrument will probably be less useful to teams with more than 8-10 members, as these teams might have problems in acquiring a shared strategic understanding. Factors that affect a teams shared strategic understanding can be team size, task types that the team is involved in, and ways of interaction within the team. In larger teams the ability to maintain a shared strategic understanding is harder to uphold as the number of possible interactions among team members is multiplied with an increased number of team members. The direct communication in smaller teams provides an environment where individual members have the opportunity to become acquainted with each other. For example, when comparing face-to-face teams with distributed teams, face-to-face teams were found to have better internal leadership and coordination than virtual teams (Burke & Chidambaram, 1994; Eveland & Bikon, 1988). This can be because a distributed team's ability to handle non-verbal communication cues is severely limited and this can increase misunderstanding among members (Warkentin et al., 1997).

Team research requires a large number of individuals to reach a sufficient number of teams for statistical analysis. This adds an extra dimension of complexity to team research, especially when

there is a longitudinal aspect to a data collection. It takes quite an effort to motivate participants to come back on several occasions (as in the trained condition in study 6), where all team members have to show up or the data collection opportunity must be postponed until all members are willing and able to be present. Some of the studies have a small number of participating teams, which is reflected in the conclusions for these studies.

Future work

There are several possible paths to continue from this point:

- To conduct an experiment with ad-hoc teams using the shared priorities instrument (see Table 28). This constellation of teams has not yet been evaluated using the shared priorities instrument. While naïve trained (trained as teams) and non-trained teams, and (even though with a small sample size) professional teams of colleagues have been assessed, ad-hoc professional teams have not been tested. Ad-hoc teams are teams where the team members are not acquainted with each other, while they are there to conduct the jobs they are trained for. Hence, they do not have the shared mental models of their ad-hoc colleagues' skills, knowledge, and ability to work as a team. A hypothesis here would be that their shared strategic understanding would be lower compared to teams where the members are colleagues who are co-trained to do their jobs.
- To assess validity it would be interesting to use the instrument to test its sensitivity to some intervention (that is, test before and after some changes in the situation/context that are expected to alter the shared strategic understanding).
- To further develop the instrument use would be to dig deeper into the content of the items. Apart from the apparent inexperienced/experienced team it would be interesting to explore if the content of the items changes over time as a reaction to changed team goals. A longitudinal study where it would be possible to study this would also shed light on interesting aspects of shared strategic understanding development.

Chapter 10. Conclusions

A central idea throughout this work has been that teams who have a higher degree of shared understanding perform better than teams with a lower degree of shared understanding, that is, the degree of shared understanding affects the level of performance. Earlier techniques that assess shared mental models in teams have been problematic to use, difficult to understand, and demanding to analyse. The purpose of this thesis was to develop an instrument that would be *usable, understandable, objective, flexible, and self-explanatory*. To meet this purpose three research questions were raised and which have been answered in this thesis. These were:

1. How can shared understanding be measured without the disadvantages of existing methods?

By letting the team members generate items and rank order these in order of importance several issues are resolved. An item is a factor that is considered by the team member to be important for the team in order to reach the team's shared goals. Fewer people (and less working hours) are needed to prepare material suited for a particular situation and analyse outcomes compared to using traditional team cognition methods for assessing a team's shared understanding (for example, shared mental models and communication analysis). In turn this reduces the costs associated with preparation. The analysis, using Kendall's W , is quite straightforward and provides easily interpreted results. There is no need for the administrators of the data collection to interpret the lists and the results. The shared priorities instrument is not tied to the experimental situation, it is an instrument that can be used in both laboratory experiments and applied settings. Thus, the technique is not limited to situations that are fully controlled and where every change is completely understood by the data collection administrator(s). As the change and development is captured using the team members' perspective, there is no need for the administrators of the data collection to understand this completely. In addition, the

instrument does not require extensive training to use nor to understand the outcome from the assessments.

2. How can shared understanding be assessed without the bias of self-ratings and/or experts/observers?

Over the studies the need for subject matter experts to be involved in the use of the instrument has been reduced. Hence, using the shared priorities instrument removes bias from subject matter experts, and also the bias of self-ratings. Comparison of the rank ordered lists provides a picture of the shared understanding (regarding the items in question).

3. Can team performance be better understood by the outcomes of an instrument that measures shared understanding?

The sixth study provides results where outcomes from the shared priorities instrument correlate with performance. The model from study 6 shows how the concept shared understanding relates to both team workload and general performance. As the shared priorities instrument results do not correlate with CARS nor with DATMA Teamwork Assessment, while all three measure correlate with performance, it is clear that the shared priorities measure captures different aspects of team cognition. This can be considered as a link between shared cognition and interactive team cognition as the shared priorities instrument, in some respect, assesses knowledge about the situation and desired future states. However, the view on these future states are highly affected by what is happening, what other team members are planning to do, and how they manage to communicate this. In addition the prioritisation is an interactive process, as the items of one team members is ranked by other team members.

In comparison to other measures of team cognition, the shared priorities measure requires little preparation, demands no expertise in the domain in which the participants partake on part of the researcher, and is fairly straight-forward to analyse. That the structural model from study 6 is significant shows that the instrument is valid and reliable.

The presented six studies involved teams ranging from two to four members. Regarding teams from the professional domains that participated, this was what was feasible for the data collection opportunities. However, the teams that are the focus of this research concern teams of small size, that is 2 to 10 members. In larger teams there is a possibility that team members will have problems with having insightful opinions about other team members.

For this thesis shared understanding has been explored in several different domains and with many different types of participants. Students (study 1, 3, 5, and 6), fighter pilots (study 2), and tank commanders (study 4) have all contributed to the development of the shared priorities instrument. Domains that have been included are dismounted soldiers (study 1), forest fire-fighting (study 3, 5, and 6), air-to-air combat (study 2), and tank combat (study 4). This has been on the short time scale (cf. military tactical level; study 1 and 2) and mid time scale levels (cf. military operational level; study 3, 4, 5, and 6) of command. This diversity allows for some generalisation, namely that the problem of maintaining a shared strategic understanding within a team seems quite general. Also, most domains, if not all, lack an instrument for assessing the shared understanding in a team that is easy to use, easy for the participants to respond to, does not consume too much time from other activities, and which correlates to team performance.

Contribution

An instrument that manages to measure shared understanding in teams, a sharedness that is dynamically changing due to circumstances in the task or situation, would provide answers to research questions that depend on data that reflect team dynamics with a temporal resolution that is difficult to achieve with techniques requiring a deeper understanding of the situation, task and domain. The outcome of such an instrument should thus, at least qualitatively, reflect the level of goals present in the team, team formation, and explain outcome in terms of performance.

The shared priorities instrument is not context dependent, that is, the instrument does not need to be prepared for the context in which it will be used. There is no reason to believe that the instrument is limited for use on a certain hierarchical organisational level. It can be used with operative teams as well as tactical teams. The instrument can probably be used with top-level teams, although this has not yet been tested. The use of the instrument is quick and straightforward, consequently it can be used repeatedly to collect data that reflect changes over time. All these assertions point to the instrument being a useful addition to the measurement techniques that are proposed and used by the team cognition community.

This research is rooted in the area of human factors with a clear focus on applied benefit. Consequently the contribution of this thesis is mainly of an applied character, in that it provides an instrument that assesses shared strategic understanding in teams.

Instructions and recommendations on how to use the shared priorities instrument and also some lessons learnt from handling the instrument are found in appendix 1.

Concluding remarks

This thesis has focused on the development of an instrument for assessing shared understanding. The work has spanned several studies where each study is based on the previous findings in order to develop and refine the instrument. The development of the instrument has been achieved with the help of participants from different backgrounds, using teams of different sizes, and iterating versions of the shared priorities instrument in different domains.

Much effort has been put into developing an instrument that is usable, understandable, objective, flexible, and self-explanatory so that the developed instrument can be employed in applied settings such as military command and control or emergency crisis management.

This work has been discussed in the light of the team cognition literature. The team cognition literature mainly uses two perspectives (shared cognition and interactive team cognition) to understand team cognition. This shared priorities instrument is suggested as a way of bridging these two perspectives as the instrument can be used to capture shared strategic understanding, a phenomenon that is changing dynamically as the team members interact, without paying attention to how team members interact.

Chapter 11. References

- Alberts, D. S., & Hayes, R. E. (2003). *Power to the edge - Command and control in the information age*. Washington DC: CCRP Publication Series.
- Alberts, D. S., Garstka, J. J., & Stein, F. P. (1999). *Network centric warfare. Developing and leveraging information superiority* (2nd ed.). Washington DC: CCRP Publication Series.
- Alberts, D. S., Huber, R. K., & Moffat, J. (2010). *NATO NEC C2 Maturity Model*. Washington DC: CCRP Publication Series.
- Allberg, H., Berggren, P., Johansson, B. J. E., Oskarsson, P.-A., Svenmarck, P., Wikberg, P., Gustavsson, P. & Thyni, D. (2014). *Brigade Staff Exercise 2013V: Assessment of Brigade Staff Efficiency, C2-Agility, Dynamic Measurement and C2 Systems*. (FOI-RH-1472-SE). Linköping: FOI.
- Andersson, J., & Berggren, P. (2003, 8th-11th September). Team Performance: The relation between shared mental model measures and team performance. *Paper presented at the Human Factors of Decision Making in Complex Systems*, Hilton Dunblane Hydro Hotel.
- Andersson, J., Berggren, P., Castor, M., Magnusson, S., & Svensson, E. (2002a). *Development of an instrument for measuring team performance potential* (FOI-R-0429-SE). Linköping: FOI. (In Swedish).
- Andersson, J., Berggren, P., Grönkvist, M., Magnusson, S., & Svensson, E. (2002b). Oxygen saturation and cognitive performance. *Psychopharmacology*, (162), 119–128.
- Andersson, J., & Rönnerberg, J. (1997). Cued Memory Collaboration: Effects of Friendship and Type of Retrieval Cue. *European Journal of Cognitive Psychology*, 9(3), 273–287.
- Andriole, S. J., & Halpin, S. M. (1986). Information Technology for Command and Control. *IEEE Transactions on Systems, Man, and Cybernetics*, 16(6), 762–765.
- Aronson, E., Wilson, T. D., & Akert, R. M. (2013). *Social psychology* (8th ed.). Upper Saddle River, N.J.: Pearson Education.
- Artman, H. (1999). Situation Awareness and Co-operation within and between Hierarchical Units in Dynamic Decision Making. *Ergonomics*, 42(11), 1404–1417.

- Artman, H. (2000). Team situation assessment and information distribution. *Ergonomics*, 43(8), 1111–28.
- Artman, H., & Granlund, R. (1999). Team Situation Awareness using Graphical or Textual Databases in Dynamic Decision Making. In *Proceedings of the 9th European Conference on Cognitive Ergonomics* (pp. 13–26). University of Limerick, Ireland.
- Artman, H., & Wærn, Y. (1998). Creation and Loss of Cognitive Empathy at an Emergency Control Centre. In Y. Wærn (Ed.), *Co-operative process management - Cognition and information technology* (pp. 69–76). London: Taylor & Francis.
- Aubé, C., Rousseau, V., & Tremblay, S. (2014). Perceived shared understanding in teams: The motivational effect of being “on the same page.” *British Journal of Psychology*, 1–19.
- Avnet, M. S., & Weigel, A. L. (2012). The Structural Approach to Shared Knowledge: An Application to Engineering Design Teams. *Human Factors*, 55(3), 581–594.
- Baber, C., Baker, K., Berggren, P., Bruseberg, A., Burov, A., Chatelier, P., ... Vandebroek, H. (2010). *Human Systems Integration for Network Centric Warfare* (RTO-HFM-RTG-155). Paris: NATO RTO.
- Banko, K., Berggren, P., Eikelboom, A. R., Estabrooke, I. V., Howard, T., Lau, R., Marklund, J., Marklund, J., van Vliet, T. & Williams, A. (2013). *Processes for Assessing Outcomes of Multi-national Missions* (STO TR-HFM-185). Paris: NATO STO.
- Banks, A. P., & Millward-Purvis, L. J. (2009). Distributed mental models: Mental models in distributed cognitive systems. *The Journal of Mind and Behavior*, 30(4), 249–266.
- Baroutsi, N. (2014). Communication in Joint Activity. Investigating Teams’ Communication Pattern in a Dynamic Decision Making Environment (LIU-IDA/KOGVET-A--14/012--SE). Linköping University, Linköping.
- Baroutsi, N., Berggren, P., Nählinder, S., & Johansson, B. J. E. (2013). *Training teams to collaborate as cohesive units* (FOI-R--3830--SE). Linköping: FOI.
- Bates, K. E., Bird, G. L., Shea, J. A., Apkon, M., Shaddy, R. E., & Metlay, J. P. (2014). A tool to measure shared clinical understanding following handoffs to help evaluate handoff quality. *J Hosp Med*, 9(3), 142–147.
- Benson, J. (2000). *Working More Creatively with Groups*. London: Routledge.
- Berggren, A. W. (2004a). Humanperspektivet i Försvarsmaktens flexibla insatsförsvär. Med särskilt fokus på interoperabilitet här och nu. Stockholm: FHS.
- Berggren, A. W. (2004b). Humanperspektivet i Försvarsmaktens flexibla insatsförsvär. Med särskilt fokus på interoperabilitet här och nu. Stockholm: FHS.
- Berggren, A. W. (2005). Människan i NBF. Med särskilt fokus på internationella insatser. Stockholm: FHS.
- Berggren, A. W. (2006). Människan i NBF. Med ökad internationell förmåga i fokus. Stockholm: FHS.

- Berggren, P., Alfredson, J., Andersson, J., & Granlund, R. (2008). Assessing shared situational awareness in dynamic situations. *Paper presented at the NATO RTO HFM-142 Symposium.*
- Berggren, P., Alfredson, J., Andersson, J., & Modeer, B. (2007). Comparing measurements of shared situational awareness in a microworld with a developmental environment. *Paper presented at the IFAC-HMS 2007, Seoul, Korea.*
- Berggren, P., & Andersson, J. (2003). Team Performance in a Simulated Low-cost Military Environment. Paper presented at the Human Factors of Decision Making in Complex Systems, Hilton Dunblane Hydro Hotel.
- Berggren, P., Baroutsi, N., Johansson, B., Turcotte, I., & Tremblay, S. (2014). Assessing team focused behaviors in emergency response teams using the shared priorities measure. In *Proceedings of the 11th ISCRAM.* The Pennsylvania State University, University Park, PA.
- Berggren, P., Hammervik, M., Klum, P., Eriksson, L., Castor, M., Lif, P., Andersson, H., Tydén, L. (2008). Användartest LKS nr 6 och 7 (FOI-R--2583--SE). Linköping: FOI.
- Berggren, P., & Johansson, B. J. E. (2010). Developing an instrument for measuring shared understanding. In *Proceedings of ISCRAM 2010* (pp. 1–8). Seattle, WA.
- Berggren, P., Johansson, B. J. E., & Baroutsi, N. (2016). Assessing the quality of Shared Priorities in teams using content analysis in a microworld experiment. *Theoretical Issues in Ergonomics Science.* Online.
- Berggren, P., Johansson, B. J. E., Baroutsi, N., & Dahlbäck, N. (2014a). The shared priorities measure as a way of assessing team strategic awareness – a bridge between self-assessment and the deep blue sea of field recordings. In *Proceedings of the ECCE 32nd annual conference.* Vienna, Austria.
- Berggren, P., Johansson, B., Svensson, E., Baroutsi, N., & Dahlbäck, N. (2014b). Statistical modelling of team training in a microworld study. In *Proceedings of the HFES 58th annual meeting.* Chicago, Illinois.
- Berggren, P., Nählinder, S., Alfredson, J., & Svensson, E. (2009). The quasi-dynamic approach to measuring complex systems. *Proceedings of the Europe Chapter Human Factors and Ergonomics Society (HFES) Conference 2009.* Linköping.
- Berggren, P., Nählinder, S., & Svensson, E. (2014). *Assessing Command and Control Effectiveness – Dealing with a changing world.* Farnham: Ashgate.
- Berggren, P., Prytz, E., Johansson, B., & Nählinder, S. (2011). The relationship between Workload, Teamwork, Situation Awareness, and Performance in Teams: A microworld study. In *Proceedings of the Human Factors and Ergonomics Society 55th Annual Meeting* (pp. 851–855). Las Vegas, Nevada.
- Berggren, P., Svensson, J., & Hörberg, U. (2008). *Mätning av gemensam lägesbild vid ledning på stridsteknisk och taktisk nivå. Studie genomförd på TCCS* (FOI-R--2647--SE). Linköping: FOI.
- Berggren, P., Svensson, J., Hörberg, U., Jonsson, S., & Höglund, F. (2009). *Shared priorities as a measure of shared understanding.* Paper presented at the Human Factors and Ergonomics Society Europe Chapter Annual Meeting.

- Biemann, T., Ellwart, T., & Rack, O. (2014). Quantifying similarity of team mental models: An introduction of the rrg index. *Group Processes & Intergroup Relations*, 17(1), 125–140.
- Bittner, E. A. C., & Leimeister, J. M. (2014). Creating Shared Understanding in Heterogeneous Work Groups: Why It Matters and How to Achieve It. *Journal of Management Information Systems*, 31(1), 111–144.
- Bohemia Interactive Studio. (2001). Operation Flashpoint. Codemasters.
- Borgatti, S. P., Everett, M. G., & Freeman, L. C. (1992). UCINET IV network analysis software [Computer software]. Columbia, SC: Analytic Technologies.
- Bowman, C., & Ambrosini, V. (1997). Perceptions of strategic priorities, consensus and firm performance. *Journal of Management Studies*, 34(2), 241–258.
- Brannick, M. T., & Prince, C. (1997). An overview of team performance measurement. In M. T. Brannick, E. Salas, & C. Prince (Eds.), *Team Performance Assessment and Measurement: Theory, Methods, and Applications* (pp. 3–16). Mahwah, NJ: Lawrence Erlbaum Associates.
- Brehmer, B. (1992). Dynamic decision making: Human control of complex systems. *Acta Psychologica*, 81, 211–241.
- Brehmer, B., & Dörner, D. (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, 171–184.
- Brehmer, B., & Svenmarck, P. (1995). Distributed decision making in dynamic environments: Time scales and architectures of decision making. In J.-P. Caverni, M. Bar-Hillel, F. H. Barron, & H. Jungermann (Eds.), *Contributions to decision making – I* (pp. 155–174). Amsterdam, The Netherlands: Elsevier Science.
- Burke, K., & Chidambaram, L. (1994). Development in Electronically-Supported Groups: A Preliminary Longitudinal Study of Distributed and Face-to-Face Meetings. In *Proceedings of the Twenty-Seventh Annual Hawaii International Conference on System Science* (pp. 104–113).
- Cacciabue, P. C., & Hollnagel, E. (1995). Simulation of Cognition: Applications. In J.-M. Hoc, P. C. Cacciabue, & E. Hollnagel (Eds.), *Expertise and Technology: Cognition and Human-Computer Cooperation* (pp. 55–73). New Jersey: Lawrence Erlbaum Associates.
- Cannon-Bowers, J. A., & Salas, E. (2001). Reflections on shared cognition. *Journal of Organizational Behavior*, 22(2), 195–202.
- Cannon-Bowers, J. A., Salas, E., & Converse, S. A. (1993). Shared mental models in expert team decision making. In N. J. Castellan (Ed.), *Individual and group decision making* (pp. 221–246). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Castor, M. (2009). *The use of structural equation modeling to describe the effect of operator functional state on air-to-air engagement outcomes*. Linköping University: Linköping University Electronic Press.
- Castor, M., Hanson, E., Svensson, E., Nählinder, S., Blaye, P. Le, MacLeod, I., Wright, N., Alfredson, J., Ågren, L., Berggren, P., Juppet, V., Hilburn, B., & Ohlsson, K. (2003). *GARTEUR Handbook of Mental Workload Measurement* (GARTEUR FM AG13 FINAL REPORT - GARTEUR TP 14). GARTEUR.

- Castor, M., Lindoff, J., & Hammervik, M. (2008). Användartest LKS demonstrator version 3. Linköping: FOI.
- Clark, H. (1992). *Arenas of language*. Chicago: The University of Chicago Press.
- Clark, H. (1996). *Using language*. Cambridge: University Press.
- Cook, T. D., & Campbell, D. T. (1979). *Quasi-experimentation: Design and analysis issues for field settings*. Boston: Houghton Mifflin Company.
- Cooke, N. J., Champion, M., Rajivan, P., & Jariwala, S. (2013a). Cyber situation awareness and teamwork. *EAI Endorsed Transactions on Security and Safety*, 13(1-6), 1–6.
- Cooke, N. J., & Gorman, J. C. (2013). Microworld experimentation with teams. In J. D. Lee & A. Kirlik (Eds.), *The Oxford handbook of cognitive engineering* (pp. 327–335). New York, NY: Oxford University Press.
- Cooke, N. J., Gorman, J. C., Duran, J. L., & Taylor, A. R. (2007a). Team cognition in experienced command-and-control teams. *Journal of Experimental Psychology: Applied*, 13(3), 146–57.
- Cooke, N. J., Gorman, J. C., & Kiekel, P. A. (2008). Communication as team-level cognitive processing. In M. P. Letsky, N. W. Warner, S. M. Fiore, & C. A. P. Smith (Eds.), *Macro-cognition in teams. Theories and methodologies* (pp. 51–64). Aldershot: Ashgate.
- Cooke, N. J., Gorman, J. C., Myers, C. W., & Duran, J. L. (2013b). Interactive team cognition. *Cognitive Science*, 37(2), 255–85.
- Cooke, N. J., Gorman, J. C., & Rowe, L. J. (2009). An ecological perspective on team cognition. In *Team effectiveness in complex organizations: Cross-disciplinary perspectives and approaches* (pp. 157–182). New York: Routledge/Taylor & Francis Group.
- Cooke, N. J., Gorman, J. C., & Winner, J. L. (2007b). Team cognition. In F. T. Durso, R. S. Nickerson, S. T. Dumais, S. Lewandowsky, & T. J. Perfect (Eds.), *Handbook of applied cognition* (2nd ed., pp. 239–268). Chichester: John Wiley & Sons, Ltd.
- Cooke, N. J., Kiekel, P. A., Bell, B., & Salas, E. (2002). Addressing Limitations of the Measurement of Team Cognition. *Proceedings of the HFES 46th Annual Meeting*, 403–407.
- Cooke, N. J., Kiekel, P. A., Salas, E., Stout, R., Bowers, C., & Cannon-Bowers, J. A. (2003). Measuring team knowledge: A window to the cognitive underpinnings of team performance. *Group Dynamics: Theory, Research, and Practice*, 7(3), 179–199.
- Cooke, N. J., Salas, E., Cannon-Bowers, J. A., & Stout, R. J. (2000). Measuring Team Knowledge. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 42(1), 151–173.
- Cooke, N. J., Salas, E., Kiekel, P. A., & Bell, B. (2004). Advances in measuring team cognition. In E. Salas & S. M. Fiore (Eds.), *Team cognition: Understanding the factors that drive process and performance*. Washington, DC: American Psychological Association.
- Corvera Charaf, M., Rosenkranz, C., & Holten, R. (2013). The emergence of shared understanding: applying functional pragmatics to study the requirements development process. *Information Systems Journal*, 23, 115–135.
- Danielsson, E., & Larsson, G. (2014). From common operating picture to situational awareness. *International Journal of Emergency Management*, 10(1), 28–47.

- Davison, R. B., Hollenbeck, J. R., Barnes, C. M., Slesman, D. J., & Ilgen, D. R. (2012). Coordinated action in multiteam systems. *The Journal of Applied Psychology, 97*(4), 808–824.
- DeChurch, L. A., & Mesmer-Magnus, J. R. (2010a). Measuring shared team mental models: A meta-analysis. *Group Dynamics: Theory, Research, and Practice, 14*(1), 1–14.
- DeChurch, L. A., & Mesmer-Magnus, J. R. (2010b). The cognitive underpinnings of effective teamwork: a meta-analysis. *The Journal of Applied Psychology, 95*(1), 32–53.
- DeChurch, L. A., & Zaccaro, S. J. (2010). Perspectives: Teams Won't Solve This Problem. *Human Factors, 52*(2), 329–334.
- Dekker, S. W. A., Jacobsson, M., Ohlsson, K., Roessing, J., Zon, R., Leggatt, A., Neary, C., Andersson, J., Berggren, P., Eisele, P., Svensson, E., Andersen, H., Andersen, H., & Eden, B. (2001). *State of the art* (VINTHEC-WP1-TR 01). VINTHEC.
- Donelson, R. F. (2006). *Group Dynamics* (4th ed.). Belmont, CA: Thomson Wadsworth Publishing.
- Duff, S. N., Giudice, K. Del, Johnston, M., Flint, J., & Kudrick, B. (2014). A systems approach to diagnosing and measuring teamwork in complex sociotechnical organizations. In *Proceedings of the Human Factors and Ergonomics Society 58th Annual Meeting* (pp. 573–577). Chicago, IL.
- Dörner, D., Kreuzig, H., Reiter, F., & Stäudel, T. (1983). *Lobhausen. Vom Umgang mit Unbestimmtheit und Komplexität*. Bern: Huber.
- Dörner, D., Stäudel, T., & Strohschneider, S. (1986). *Moro. Programmdokumentation* (Memorandum XX, Psychologie II). Bamberg: University of Bamberg.
- Elliott, T., Welsh, M., & Nettelbeck, T. (2007). Investigating naturalistic decision making in a simulated microworld: What questions should we ask? *Behavior Research Methods, 39*(4), 901–910.
- Emery, F. E., & Trist, E. L. (1960). Socio-technical systems. In C. W. Churchman & M. Verhulst (Eds.), *Management Science Models and Techniques* (Vol. 2, pp. 83–97). Oxford: Pergamon.
- Endsley, M. R. (1995a). Toward a Theory of Situation Awareness in Dynamic Systems. *Human Factors, 37*(1), 32–64.
- Endsley, M. R. (1995b). Measurement of Situation Awareness in Dynamic Systems. *Human Factors, 37*(1), 65–84.
- Endsley, M. R. (2000). Theoretical underpinnings of situational awareness: A critical review. In D. Garland & M. Endsley R (Eds.), *Experimental Analysis and Measurement of Situational Awareness*. Daytona Beach: Embry-Riddle Aeronautical University Press.
- Endsley, M. R., & Garland, D. J. (2000). *Situation awareness analysis and measurement*. Mahwah, NJ.: Lawrence Erlbaum Associates.
- Endsley, M. R., & Jones, W. M. (1997). *Situation Awareness Information Dominance & Information Warfare* (AL/CF-TR-1997-0156).
- Endsley, M. R., & Jones, W. M. (2001). A model of inter- and intrateam situation awareness: Implications for design, training, and measurement. In M. McNeese, E. Salas, & M. Endsley (Eds.), *New trends in cooperative activities: Understanding system dynamics in complex environments*. Santa Monica: Human Factors and Ergonomics Society.

- Entin, E. B., & Entin, E. E. (2000). Assessing team situation awareness in simulated military missions. In *Proceedings of the International Ergonomics Association and 44th Annual Meeting of the Human Factors and Ergonomics Society* (pp. 73–76). San Diego, CA.
- Essens, P., Vogelaar, A., Mylle, J., Baranski, J., Goodwin, G. F., Van Buskirk, W., ... Hof, T. (2010). *CTEF 2.0—Assessment and Improvement of Command Team Effectiveness: Verification of Model and Instrument* (RTO TR-HFM-127). Paris: NATO RTO.
- Essens, P., Vogelaar, A., Mylle, J., Blendell, C., Paris, C. R., Halpin, S. M., & Baranski, J. (2005). *Military Command Team Effectiveness: Model and instrument for assessment and improvement* (NATO RTO TR-HFM-087). Paris: NATO RTO.
- Eveland, J. D., & Bikson, T. K. (1988). Work group structures and computer support: a field experiment. *ACM Transactions on Information Systems*, 6(4), 354–379.
- Field, A., & Hole, G. J. (2003). *How to design and report experimnts*. London: SAGE Publications Ltd.
- Fiske, S., & Taylor, S. E. (2013). *Social Cognition: From Brains to Culture* (2nd ed.). London: SAGE Publications.
- Fitts, P. M. (1947). *Psychological research on equipment design* (AD651795). Washington DC: Aero-Medical Laboratory.
- Fogel, A., Berggren, P., & Jander, H. (2004). BA rescue team performance - Exploring team situation awareness, mental models, and team processes in breathing apparatus rescues (FOI-R--1508--SE). Linköping: FOI.
- Fracker, M. L. J. (1991). *Measures of situation awareness: review and future directions* (AL-TR-1991-0128). Dayton, OH: Human Engineering Division.
- Fransen, J., Weinberger, A., & Kirschner, P. A. (2013). Team Effectiveness and Team Development in CSCL. *Educational Psychologist*, 48(1), 9–24.
- Fullagar, C. J., & Egleston, D. O. (2008). Norming and Performing□ : Using Microworlds to Understand the Relationship Between Team Cohesiveness and Performance. *Journal of Applied Social Psychology*, 38(10), 2574–2593.
- Gentner, D., & Stevens, A. L. (1983). *Mental models*. London: Lawrence Earlbaum Associates.
- Gonzalez, C., Vanyukov, P., & Martin, M. K. (2005). The use of microworlds to study dynamic decision making. *Computers in Human Behavior*, 21(2), 273–286.
- Gonzalez-Benito, J., Aguinis, H., Boyd, B. K., & Suarez-Gonzalez, I. (2010). Coming to Consensus on Strategic Consensus: A Mediated Moderation Model of Consensus and Performance. *Journal of Management*, 38(6), 1685–1714.
- Gorman, J. C., & Cooke, N. J. (2011). Changes in team cognition after a retention interval: the benefits of mixing it up. *Journal of Experimental Psychology. Applied*, 17(4), 303–19.
- Gorman, J. C., Cooke, N. J., & Winner, J. L. (2006). Measuring team situation awareness in decentralized command and control environments. *Ergonomics*, 49(12-13), 1312–25.
- Granlund, R. (2002). *Monitoring distributed teamwork training*. Linköping University.

- Granlund, R. (2003). Monitoring experiences from command and control research with the C3Fire microworld. *Cognition, Technology & Work*, 5(3), 183-190.
- Granlund, R., Johansson, B. J. E., & Persson, M. (2001). C3Fire a Micro-world for Collaboration Training and Investigations in the ROLF environment. In *Proceedings to the 42nd Conference on Simulation and Modeling, Simulation in Theory and Practice, SIMS 2001*. Porsgrunn, NO.
- Granåsen, M., & Andersson, D. (2015). Measuring team effectiveness in cyber-defense exercises: a cross-disciplinary case study. *Cognition, Technology & Work, Online*.
- Gustavsson, B., & Bäckman, C. (2006). *Vilka samband finns mellan teamkaraktäristika och teamprestation?* (ISSN 1401-565X ILM serie T:34). Stockholm: Försvarshögskolan.
- Hackman, J., R. (1987). The design of work teams. In J. W. Lorsch (Ed.), *Handbook of Organizational Behavior* (pp. 315–341). Englewood Cliffs, NJ: Prentice Hall.
- Hambrick, D. C. (1981). Strategic Awareness within Top Management Teams. *Strategic Management Journal*, 2(3), 263–279.
- Hammervik, M., Lindoff, J., & Castor, M. (2006). Användartest LKS demonstrator version 1 (FOI-R--2034--SE). Linköping: FOI.
- Hammervik, M., Lindoff, J., Castor, M., Berggren, P., & Kylesten, B. (2007). LKS demonstrator som plattform för att mäta effekter av ledningskrigföring på ledningsförmåga - Användartest LKS demonstrator version 2 (FOI-R--2254--SE). Linköping: FOI.
- Hart, S. G., & Staveland, L. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In P. A. Hancock & N. Meshkati (Eds.), *Human Mental Workload*. Amsterdam: North Holland Press.
- Hauland, G. (2008). Measuring Individual and Team Situation Awareness During Planning Tasks in Training of En Route Air Traffic Control. *The International Journal of Aviation Psychology*, 18(3), 290–304.
- Heath, C., & Luff, P. (1992). Collaboration and Control: Crisis Management and Multimedia Technology in London Underground Line Control Rooms. *Journal of Computer Supported Cooperative Work*, 1(1), 24–48.
- Hof, T., de Koning, L., & Essens, P. (2010). Measuring effectiveness of teams and multi-team systems in operations. In *Proceedings of the 15th ICCRTS*.
- Hollnagel, E. (1998). *Cognitive reliability and error analysis CREAM*. Oxford: Elsevier Science.
- Hollnagel, E., & Woods, D. D. (1983). Cognitive systems engineering: New wine in new bottles. *International Journal of Man-Machine Studies*, 18(18), 583–600.
- Hollnagel, E., & Woods, D. D. (2005). *Joint cognitive systems: Foundations of cognitive systems engineering*. Boca Raton, FL: CRC Press.
- Howie, D. E., & Vicente, K. J. (1998). Measures of operator performance in complex, dynamic microworlds: advancing the state of the art. *Ergonomics*, 41(4), 485–500.
- Hubert, L., & Schultz, J. (1976). Quadratic assignment as a general data analysis strategy. *British Journal of Mathematical and Statistical Psychology*, 29, 190–241.

- Hutchins, E. (1995). *Cognition in the wild*. Cambridge, MA: MIT Press.
- Hutchins, E., & Klausen, T. (1996). Distributed Cognition in an Airline Cockpit. In Y. Engeström & D. Middleton (Eds.), *Cognition and Communication at Work* (pp. 15–34). New York: Cambridge University Press.
- Jajodia, S., Liu, P., Swarup, V., & Wang, C. (2010). *Cyber situational awareness: advances in information security*. Berlin: Springer.
- Jeannot, E., Kelly, C., & Thompson, D. (2003). *The Development of Situation Awareness Measures in ATM Systems* (HRS/HSP-005-REP- 01). Brussels, Belgium: EUROCONTROL.
- Jensen, E. (2009). Sensemaking in military planning: a methodological study of command teams. *Cogn Tech Work*, 11, 103–118.
- Jentsch, C., & Beimbom, D. (2014). Shared understanding among business and IT – a literature review and research agenda. In *Twenty Second European Conference on Information Systems* (pp. 1–15). Tel Aviv.
- Jobidon, M.-E., Breton, R., Rousseau, R., & Tremblay, S. (2006). Team response to workload transition: The role of team structure. In *Proceedings of the Human Factors and Ergonomics Society 50th Annual Meeting—2006* (pp. 1769–1773).
- Johansson, B. J. E., Persson, M., Granlund, R., & Mattsson, P. (2003). C3Fire in command and control research. *Cognition, Technology & Work*, 5(3), 191–196.
- Johnson, T. E., Lee, Y., Lee, M., O'Connor, D. L., Khalil, M. K., & Huang, X. (2007). Measuring Sharedness of Team-Related Knowledge: Design and Validation of a Shared Mental Model Instrument. *Human Resource Development International*, 10(4), 437–454.
- Jones, P. E., & Roelofsma, P. H. (2000). The potential for social contextual and group biases in team decision-making: biases, conditions and psychological mechanisms. *Ergonomics*, 43(8), 1129–52.
- Jöreskog, K., & Sörbom, D. (1993). *LISREL: Structural Equation Modeling with the SIMPLIS Command Language*. Hillsdale: Lawrence Erlbaum Associates, Inc.
- Kellermanns, F. W. (2005). The Lack of Consensus About Strategic Consensus: Advancing Theory and Research. *Journal of Management*, 31(5), 719–737.
- Kellermanns, F. W., Walter, J., Floyd, S. W., Lechner, C., & Shaw, J. C. (2011). To agree or not to agree? A meta-analytical review of strategic consensus and organizational performance. *Journal of Business Research*, 64(2), 126–133.
- Kendall, M. G., & Babington Smith, B. (1939). The problem of m rankings. *Annals of Mathematical Statistics*, 10(3), 275–287.
- Keppel, G. (1991). *Design and analysis. A researcher's handbook* (Third ed.). Upper Saddle River, NJ: Prentice Hall, Inc.
- Kim, B., Bush, L. A. M., & Shah, J. (2013). Quantitative estimation of the strength of agreements in goal-oriented meetings. In *2013 IEEE International Multi-Disciplinary Conference on Cognitive*

- Methods in Situation Awareness and Decision Support (CogSIMA)* (pp. 38–44). San Diego, CA: IEEE.
- Klein, G. A. (1999). *Sources of power: How people make decisions*. Cambridge, MA: MIT Press.
- Klein, G. A. (2009). *Streetlights and shadows. Searching for the keys to adaptive decision making*. Cambridge, MA: MIT Press.
- Klein, G. A., Orasanu, J. M., Calderwood, R., & Zsombok, C. E. (1993). *Decision making in action: Models and methods*. Norwood, NJ: Ablex Publishing Corporation.
- Klein, G. A., Ross, K. G., Moon, B. M., Klein, D. E., Hoffman, R. R., & Hollnagel, E. (2003). Macrocognition. *IEEE Intelligent Systems*, 18(3), 81–85.
- Klein, G. A., & Wright, C. (2016). Macrocognition: From Theory to Toolbox. *Frontiers in Psychology*, 7(54), 1–5.
- Kleinsmann, M., & Valkenburg, R. (2008). Barriers and enablers for creating shared understanding in co-design projects. *Design Studies*, 29(4), 369–386.
- Klimoski, R., & Mohammed, S. (1994). Team mental model: Construct or metaphor? *Journal of Management*, 20(2), 403–437.
- Kwon, W., Clarke, I., & Wodak, R. (2014). Micro-Level Discursive Strategies for Constructing Shared Views around Strategic Issues in Team Meetings. *Journal of Management Studies*, 51(2), 265–290.
- Kylesten, B., & Nählinder, S. (2011). The effect of decision-making training: results from a command-and-control training facility. *Cognition, Technology & Work*, 13, 93–101.
- Lagerlöf, J., & Pallin, K. (1999). Doctrine and command in the Swedish Armed Forces. In *Proc. of the CCRTS Command and Control Research and Technology Symposium*. Newport, RI.
- Lane, R., McMaster, H., Adnum, J., & Cavanagh, M. (2014). Quality reflective practice in teacher education: a journey towards shared understanding. *Reflective Practice*, 15(4), 481–494.
- Langan-Fox, J., Code, S., & Langfield-Smith, K. (2000). Team mental models: Techniques, methods, and analytic approaches. *Human Factors*, 42(2), 242–271.
- Langan-Fox, J., Wirth, A., Code, S., Langfield-Smith, K., & Wirth, A. (2001). Analyzing shared and team mental models. *International Journal of Industrial Ergonomics*, (28), 99–112.
- Lewis, K. (2003). Measuring transactive memory systems in the field: Scale development and validation. *Journal of Applied Psychology*, 88(4), 587–604.
- Lindgren, I., Hirsch, R., & Berggren, P. (2007). It takes three points to define a common ground: Breathing apparatus firefighters' communication during rescue operations. *Journal of Pragmatics*, (39), 1482–1502.
- Lindgren, I., & Smith, K. (2006). Using Microworlds to Understand Cultural Influences on Distributed Collaborative Decision Making in C2 Settings. *Paper presented at the 11th International Command and Control Research and Technology Symposium, Cambridge, UK*.
- Løvborg, L., & Brehmer, B. (1991). *NEWFIRE - A flexible system for running simulated fire-fighting experiments* (Riso-M-2953). Roskilde: Risø National Laboratory.

- MacDermott, J. (2009). *Humanitära organisationers villkor för civil-militär samverkan* (FOI-R--2808--SE). Stockholm: FOI.
- MacMillan, J., Paley, M. J., Entin, E. B., & Entin, E. E. (2005). Questionnaires for Distributed Assessment of Team Mutual Awareness. In N. Stanton, K. Hedge, E. Brookhuis, E. Salas, & H. Hendricks (Eds.), *Handbook of Human Factors Methods* (pp. 51-1-9). London: Taylor and Francis.
- Magnusson, S., & Berggren, P. (2002). Dynamic Assessment of Pilot Mental Status. In *Proceedings of the HFES 46th Annual Meeting* (pp. 1997-2001). Baltimore, MD.
- Marks, M. A., DeChurch, L. A., Mathieu, J. E., Panzer, F. J., & Alonso, A. (2005). Teamwork in Multiteam Systems. *The Journal of Applied Psychology, 90*(5), 964-71.
- Mathieu, J. E., Heffner, T. S., Goodwin, G. F., Salas, E., & Cannon-Bowers, J. A. (2000). The influence of shared mental models on team process and performance. *Journal of Applied Psychology, 85*(2), 273-283.
- Mathieu, J. E., Marks, M. A., & Zaccaro, S. J. (2001). Multiteam Systems. In N. Anderson, D. S. Ones, H. K. Sinangil, & C. Viswesvaran (Eds.), *Handbook of Industrial, Work and Organizational Psychology Volume 2: Organizational Psychology* (pp. 289-313). SAGE.
- Mathieu, J. E., Maynard, M. T., Rapp, T., & Gilson, L. (2008). Team Effectiveness 1997-2007: A Review of Recent Advancements and a Glimpse Into the Future. *Journal of Management, 34*(3), 410-476.
- Matteson, M. (2015). Capturing shared mental models: An approach for bona fide groups. *Journal of Librarianship and Information Science, 47*(1), 56-70.
- Matthews, M. D., & Beal, S. A. (2002). *Assessing Situation Awareness in Field Training Exercises* (Research Report 1795). West Point, NY: ARI
- Maynard, M. T., & Gilson, L. L. (2014). The Role of Shared Mental Model Development in Understanding Virtual Team Effectiveness. *Group & Organization Management, 39*(1), 3-32.
- McGuinness, B. (1999). Situational awareness and the crew awareness rating scale (CARS). *1999 Avionics Conference*. Heathrow, London.
- McGuinness, B., & Ebbage, L. (2002). Assessing Human Factors in Command and Control: Workload and Situational Awareness Metrics. In *2002 Command and Control Research and Technology Symposium*. Monterey, CA.
- McGuinness, B., & Foy, L. (2000). Subjective measure of SA: the Crew Awareness Rating Scale (CARS). In *The Human Performance, Situational Awareness an Automation Conference*.
- McKendrick, R., Shaw, T., de Visser, E., Saqer, H., Kidwell, B., & Parasuraman, R. (2014). Team Performance in Networked Supervisory Control of Unmanned Air Vehicles: Effects of Automation, Working Memory, and Communication Content. *Human Factors, 56*(3), 463-475.
- McNeese, M., Bains, P., Brewer, I., Brown, C. E., Connors, E. S., Jefferson, T., ... Terrell, I. S. (2005). The NeoCITIES Simulation: Understanding the design and methodology used in a team emergency management simulation. In *Proceedings of the Human Factors and Ergonomics Society 49th Annual Meeting* (pp. 591-594). Santa Monica, CA.

- Meister, D. (1985). *Behavioral analysis and measurement methods*. New York: John Wiley & Sons.
- Mohammed, S., & Dumville, B. C. (2001). Team mental models in a team knowledge framework: expanding theory and measurement across disciplinary boundaries. *Journal of Organizational Behavior*, 22(2), 89–106.
- Mohammed, S., Ferzandi, L., & Hamilton, K. (2010). Metaphor No More: A 15-Year Review of the Team Mental Model Construct. *Journal of Management*, 36(4), 876–910.
- Mohammed, S., Klimoski, R., & Rentsch, J. R. (2000). The measurement of Team Mental Models: We have no shared schema. *Organizational Research Methods*, 3(2), 123–165.
- Nonose, K., Kanno, T., & Furuta, K. (2010). An evaluation method of team situation awareness based on mutual belief. *Cognition, Technology & Work*, 12(1), 31–40.
- Nählinder, S., Berggren, P., & Persson, B. (2005). Increasing training efficiency using embedded pedagogical tools in a combat flight simulator. In *Proceedings of the HFES 49th Annual meeting*. Orlando, FL.
- Nählinder, S., Berggren, P., & Svensson, E. (2004). Reoccurring LISREL patterns describing mental workload, situation awareness and performance. In *Proceedings of the HFES 48th Annual meeting*. New Orleans.
- Omodei, M. M., Wearing, A. J., McLennan, J., Elliot, G. C., & Clancy, J. M. (2004). “More is better?”: Problems of Self-Regulation in Naturalistic Decision Making Settings. In B. Brehmer, H. Montgomery, & R. Lipshitz (Eds.), *How Professionals make decisions*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.
- Onag, Z., & Tepeci, M. (2014). Team Effectiveness in Sport Teams: The Effects of Team Cohesion, Intra Team Communication and Team Norms on Team Member Satisfaction and Intent to Remain. *Procedia - Social and Behavioral Sciences*, 150, 420–428.
- Oppl, S., & Stary, C. (2013). Facilitating shared understanding of work situations using a tangible tabletop interface. *Behaviour & Information Technology*, 33(6), 619–635.
- Orasanu, J. M., & Connolly, T. (1993). The reinvention of decision making. In G. Klein, J. Orasanu, R. Calderwood, & C. Zsombok E. (Eds.), *Decision making in action: Models and methods* (pp. 3–20). Norwood, NJ: Ablex.
- Paris, C., Banko, K., Berggren, P., Burov, O., Clark, E., Davis, K., Galster, S., Halpin, S., Hartt, D., Kermarrec, Y., Lussier, J., Quiram, T., Schaab, B., Ward, J. & Wikberg, P. (2014) Measuring and Analyzing Command and Control Performance Effectiveness. (Final report. NATO STO-TR-HFM-156). Paris: NATO.
- Patterson, E. S., & Miller, J. E. (2010). *Macro-cognition metrics and scenarios. Design and evaluation for real-world teams*. Farnham: Ashgate.
- Podsakoff, P. M., MacKenzie, S. B., Lee, J.-Y., & Podsakoff, N. P. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. *The Journal of Applied Psychology*, 88(5), 879–903.
- Prytz, E., Berggren, P., & Johansson, B. (2010). *Performance and shared understanding in mixed C2-systems* (FOI-R--3155--SE). Linköping.

- Rencrantz, C., Lindoff, J., & Andersson, J. (2005). Är det viktigt att förstå varandra för att prestera bra? (FOI-R-- 1885 --SE). Stockholm: FOI.
- Rencrantz, C., Lindoff, J., Andersson, J., & Svensson, E. (2007). Utvärdering incidentberedskap Utvärdering av av incidentberedskap Luft i StriC Luft i StriC Utvärdering av incidentberedskap Luft i StriC (FOI-R--2261--SE). Linköping: FOI.
- Rousseau, D. M. (1985). Issues of level in organizational research: Multi-level and cross-level perspectives. In L. L. Cummings & B. M. Staw (Eds.), *Research in Organizational Behavior* (Vol. 7, pp. 1–37). Greenwich: JAI Press Inc.
- Saetrevik, B., & Eid, J. (2013). The “Similarity Index” as an Indicator of Shared Mental Models and Situation Awareness in Field Studies. *Journal of Cognitive Engineering and Decision Making*, 8(2), 119–136.
- Salas, E., & Cannon-Bowers, J. A. (2001). Special issue preface. *Journal of Organizational Behavior*, 22, 87–88.
- Salas, E., Cooke, N. J., & Rosen, M. A. (2008). On Teams, Teamwork, and Team Performance: Discoveries and Developments. *Human Factors*, 50(3), 540–547.
- Salas, E., Dickinson, T. L., Converse, S. A., & Tannenbaum, S. I. (1992). Toward an understanding of team performance and training. In *Teams: Their training and performance* (pp. 3–29). Norwood, NJ: Ablex.
- Salas, E., & Fiore, S. M. (2004). *Team cognition: understanding the factors that drive process and performance*. Washington, DC: American Psychological Association.
- Salas, E., Fiore, S., Warner, N., & Letsky, M. (2010). Emerging multi-disciplinary theoretical perspectives in team cognition: an overview. *Theoretical Issues in Ergonomics Science*, 11(4), 245–249.
- Salas, E., Prince, C., Baker, D. P., & Shresta, L. B. (1995). Situation Awareness in Team Performance: Implications for Measurement and Training. *Human Factors*, 37(1), 123–136.
- Salmon, P. M., Stanton, N. A., Walker, G. H., Baber, C., Jenkins, D. P., McMaster, R., & Young, M. S. (2008). What really is going on? Review of situation awareness models for individuals and teams. *Theoretical Issues in Ergonomics Science*, 9(4), 297–323.
- Salmon, P. M., Stanton, N. A., Walker, G. H., & Green, D. (2006). Situation awareness measurement: A review of applicability for C4i environments. *Applied Ergonomics*, 37(2), 225–238.
- Salmon, P. M., Stanton, N. A., Walker, G. H., & Jenkins, D. P. (2009). *Distributed Situation Awareness. Theory, measurement and application to teamwork*. Farnham: Ashgate.
- Saner, L. D., Bolstad, C. A., Gonzalez, C., & Cuevas, H. M. (2009). Measuring and Predicting Shared Situation Awareness in Teams. *Journal of Cognitive Engineering and Decision Making*, 3(3), 280–308.
- Sapateiro, C., Ferreira, A., & Antunes, P. (2010). Using Microworlds to Study Teamwork at the Cognitive Level. In *SBSFC: 8º Simpósio Brasileiro de Sistemas Colaborativos*.
- Schober, M. F., & Spiro, N. (2014). Jazz improvisers’ shared understanding: a case study. *Frontiers in Psychology*, 5(Article 808), 1–21.

- Slater, P. E. (1958). Contrasting Correlates of Group Size. *Sociometry*, 21(2), 129–139.
- Smart, P. R., Mott, D., Sycara, K., Braines, D., Strub, M., & Shadbolt, N. R. (2009). Shared Understanding within Military Coalitions: A Definition and Review of Research Challenges. In *Knowledge Systems for Coalition Operations*. Southampton, UK.
- Smith-Jentsch, K. A., Mathieu, J. E., & Kraiger, K. (2005). Investigating linear and interactive effects of shared mental models on safety and efficiency in a field setting. *The Journal of Applied Psychology*, 90(3), 523–35.
- Stahl, G. (2010). Perspectives: Group Cognition Factors in Sociotechnical Systems. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 52(2), 340–343.
- Stanton, N. A., Baber, C., & Harris, D. (2008). *Modelling command and control. Event analysis of systematic teamwork*. Aldershot: Ashgate.
- Stanton, N. A., Salmon, P. M., Walker, G. H., Baber, C., & Jenkins, D. P. (2005). *Human factors methods. A practical guide for engineering and design*. Aldershot: Ashgate.
- Steen, M., Buijs, J., & Williams, D. (2014). the Role of Scenarios and Demonstrators in Promoting Shared Understanding in Innovation Projects. *International Journal of Innovation and Technology Management*, 11(1), 1–21.
- Strohschneider, S., & Gerdes, J. (2004). MS ANTWERPEN: Emergency Management Training for Low-Risk Environments. *Simulation & Gaming*, 35(3), 394–413.
- Šumanski, M. M., Kolenc, I., & Markic, M. (2007). Teamwork and defining group structures. *Team Performance Management: An International Journal*, 13(3/4), 102–116.
- Svenmarck, P. (1998). *Local Co-ordination in Dynamic Environments: Theories and Co-ordination Support*. Unpublished Liu-Tek-Lic-1998:52, Linköping University; Linköping.
- Svenmarck, P., & Brehmer, B. (1991). D3Fire - An experimental paradigm for the study of distributed decision making. In B. Brehmer (Ed.), *Proceedings of the 3rd MOHLAWC Workshop on Distributed Decision Making* (pp. 47–77). Belgirate, Italy: Risö National Laboratory.
- Svenson, P., Forsgren, R., Kylesten, B., Berggren, P., Fah, W. R., Chon, M. S., & Hann, J. K. Y. (2010). Swedish-Singapore studies of Bayesian modelling techniques for tactical intelligence analysis. *13th International Conference on Information Fusion*. Edinburgh.
- Svensson, E., & Nählinder, S. (2014). Dynamic Measures of Effectiveness in Command and Control. In P. Berggren, S. Nählinder, & E. Svensson (Eds.), *Assessing Command and Control Effectiveness – Dealing with a changing world* (pp. 49–70). Farnham: Ashgate.
- Svensson, E., Rencrantz, C., Lindoff, J., Berggren, P., & Norlander, A. (2006). Dynamic measures for performance assessment in complex environments. In *Proceedings of the HFES 50th Annual meeting*. San Francisco, CA.
- Svensson, E., & Wilson, G. F. (2002). Psychological and Psychophysiological Models of Pilot Performance for Systems Development and Mission Evaluation. *The International Journal of Aviation Psychology*, 12(1), 95–110.
- Swezey, R. W., & Salas, E. (1992). *Teams: Their training and performance*. Norwood, NJ: Ablex.

- Teo, T. (2013). An initial development and validation of a Digital Natives Assessment Scale (DNAS). *Computers & Education*, 67, 51–57.
- Thunholm, P., Berggren, P., & Wikberg, P. (2014). Evaluating the Effectiveness of an Armoured Brigade Staff. In P. Berggren, N. Staffan, & E. Svensson (Eds.), *Assessing Command and Control Effectiveness – Dealing with a changing world*. Farnham: Ashgate.
- Thunholm, P., Tan, D., Wikberg, P., Tey, F., Gullberg, U., Larsson, S.-Å., ... Thorstensson, M. (2011). C2 Agility: Evaluating a New Brigade Standard Operating Procedure for Information Management and Coordination. In *16th ICCRTS - Collective C2 in Multinational Civil-Military Operations*. Québec, Canada.
- Trnka, J. (2009). *Exploring Tactical Command and Control: A Role-Playing Simulation Approach*. Linköping University: Linköping University Electronic Press.
- Tuckman, B. W. (1965). Developmental Sequence in Small Groups. *Psychological Bulletin*, 63(6), 384–99.
- Uhr, C. (2014). Att åstadkomma inriktning och samordning - 7 analyser utifrån hantering av skogsbranden i Västmanland 2014. Lund: Lunds Universitet.
- Van den Bossche, P., Gijssels, W. H., Segers, M., & Kirschner, P. A. (2006). Social and Cognitive Factors Driving Teamwork in Collaborative Learning Environments. Team Learning Beliefs and Behaviors. *Small Group Research*, 37(5), 490–521.
- van der Haar, S., Segers, M., Jehn, K. A., & Van den Bossche, P. (2015). Investigating the Relation Between Team Learning and the Team Situation Model. *Small Group Research*, 46(1), 50–82.
- Warkentin, M. E., Sayeed, L., & Hightower, R. (1997). Virtual Teams versus Face-to-Face Teams: An Exploratory Study of a Web-based Conference System. *Decision Sciences*, 28(4), 975–996.
- Wegner, D. M. (1987). Transactive memory: A contemporary analysis of the group mind. In B. Mullen & G. R. Goethals (Eds.), *Theories of group behaviour* (pp. 185–208). York: Springer-Verlag.
- Wheelan, S. A. (2009). Group Size, Group Development, and Group Productivity. *Small Group Research*, 40(2), 247–262.
- Wijngaards, N., Kempen, M., Smit, A., & Nieuwenhuis, K. (2006). Towards Sustained Team Effectiveness. In O. B. et Al (Ed.), *Lecture Notes in Computer Science, Coordination, Organizations, Institutions, and Norms in Multi-Agent Systems* (pp. 35–47). Berlin: Springer-Verlag.
- Wildman, J. L., Salas, E., & Scott, C. P. R. (2014). Measuring Cognition in Teams A Cross-Domain Review. *Human Factors*, 56(5), 911–941.
- Winner, J. L., Freeman, J. T., Cooke, N. J., & Goodwin, G. F. (2007). A metric for the shared interpretation of commander's intent. In *Proceedings of the Human Factors and Ergonomics Society 51st annual meeting* (pp. 122–126).
- Wolbers, J., & Boersma, K. (2013). The Common Operational Picture as Collective Sensemaking. *Journal of Contingencies and Crisis Management*, 21(4), 186–199.

- Woods, D. D., & Hollnagel, E. (2006). *Joint cognitive systems. Patterns in cognitive systems engineering*. London: CRC Press Inc.
- Yin, R. K. (1994). *Case study research: Design and methods* (2nd ed.). Thousand Oaks, CA: Sage Publications.
- Young, M. S., Brookhuis, K. A., Wickens, C. D., & Hancock, P. A. (2015). State of science: mental workload in ergonomics. *Ergonomics*, 58(1), 1–17.
- Zajac, S., Gregory, M. E., Bedwell, W. L., Kramer, W. S., & Salas, E. (2014). The cognitive underpinnings of adaptive team performance in ill-defined task situations. *Organizational Psychology Review*, 4(1), 49–73.
- Zon, R., Neary, C., Taylor, M., Berggren, P., Jacobsson, M., Andersen, H., Gopher, D., & Ober, J. (2004). *Full-scale simulator experiment - Final report* (VINTHEC II-WP8-TR 01). VINTHEC.
- Åsberg, R., Hummerdal, D., & Dekker, S. W. A. (2011). There are no qualitative methods – nor quantitative for that matter: the misleading rhetoric of the qualitative–quantitative argument. *Theoretical Issues in Ergonomics Science*, 12(5), 408–415.

Chapter 12. Appendix 1

This appendix provides an instruction on how to use the instrument and also some lessons learnt from using it.

Instructions on using the shared priorities instrument

There are 7 steps to go through.

1. Choose a team member to generate the primary list. This could be the team leader or one of the members who is important in coordinating the team. Let the team member write down five items that are important in order for *the team* to reach the team's goal/objective. It is important to stress that the focus is on the team level.
 - o Alternatively, let all team members individually write down five items that are important for the team to reach the team's goal/objective. Having all team members' views might sometimes be desirable. For the researcher this might be the most informative alternative. However, one list means less work and less time spent responding for the team members.
2. Ask the person who generated the list to rank order these items from 1 to 5, where 1 is the most important, and 5 is the least important of these five items.
3. Collect the generated list.
4. Copy the items from the list and write these down on a new paper in a scrambled order. Duplicate the list so that each team member (except for the person who generated the list in the first place) can have a copy of the new list.
5. Distribute the new list to all team members (except for the person who generated the list in the first place) and ask them to individually rank order the items from 1 to 5, where 1 is the most important, and 5 is the least important of these five items.

6. Collect the rank ordered list. The team can now continue with their business.
7. Compute Kendall's W for the lists that all team members have rank ordered.

The lessons learnt from using the instrument are summarised below:

- Do not use pre-defined lists, i.e., lists that have been prepared in advance. This will limit the perspective and views that could be captured when the team members are making the lists up with factors that they perceive as important in the situation. Also, do not use standard operating procedures as factors, the order in which to handle these are well known in trained teams. This will lead to no variance in the order of the lists. Nor will lists with SOP factors reflect the present situation that the team is dealing with.
- It is important to stress that the focus is on the team level, on *shared* understanding. This will lead to the team coordinating actions to reach the desired goal.
- The responding team members should be instructed that the items should not be about common knowledge. It is about factors that are important for the team to reach its goal.
- Consider who should generate the primary list. Is it the team leader? Is someone else better suited for this task? Is there an informal leader? Should all team members be equally suitable for the task?
- Remember that the situation needs to be meaningful to the team, otherwise it will be very difficult to come up with important factors for the primary list.

Dissertations

Linköping Studies in Science and Technology

Linköping Studies in Arts and Science

Linköping Studies in Statistics

Linköpings Studies in Information Science

Linköping Studies in Science and Technology

- No 14 **Anders Haraldsson:** A Program Manipulation System Based on Partial Evaluation, 1977, ISBN 91-7372-144-1.
- No 17 **Bengt Magnhagen:** Probability Based Verification of Time Margins in Digital Designs, 1977, ISBN 91-7372-157-3.
- No 18 **Mats Cedwall:** Semantisk analys av process-beskrivningar i naturligt språk, 1977, ISBN 91-7372-168-9.
- No 22 **Jaak Urmi:** A Machine Independent LISP Compiler and its Implications for Ideal Hardware, 1978, ISBN 91-7372-188-3.
- No 33 **Tore Risch:** Compilation of Multiple File Queries in a Meta-Database System 1978, ISBN 91-7372-232-4.
- No 51 **Erland Jungert:** Synthesizing Database Structures from a User Oriented Data Model, 1980, ISBN 91-7372-387-8.
- No 54 **Sture Hägglund:** Contributions to the Development of Methods and Tools for Interactive Design of Applications Software, 1980, ISBN 91-7372-404-1.
- No 55 **Pär Emanuelson:** Performance Enhancement in a Well-Structured Pattern Matcher through Partial Evaluation, 1980, ISBN 91-7372-403-3.
- No 58 **Bengt Johnsson, Bertil Andersson:** The Human-Computer Interface in Commercial Systems, 1981, ISBN 91-7372-414-9.
- No 69 **H. Jan Komorowski:** A Specification of an Abstract Prolog Machine and its Application to Partial Evaluation, 1981, ISBN 91-7372-479-3.
- No 71 **René Reboh:** Knowledge Engineering Techniques and Tools for Expert Systems, 1981, ISBN 91-7372-489-0.
- No 77 **Östen Oskarsson:** Mechanisms of Modifiability in large Software Systems, 1982, ISBN 91-7372-527-7.
- No 94 **Hans Lunell:** Code Generator Writing Systems, 1983, ISBN 91-7372-652-4.
- No 97 **Andrzej Lingas:** Advances in Minimum Weight Triangulation, 1983, ISBN 91-7372-660-5.
- No 109 **Peter Fritzson:** Towards a Distributed Programming Environment based on Incremental Compilation, 1984, ISBN 91-7372-801-2.
- No 111 **Erik Tengvald:** The Design of Expert Planning Systems. An Experimental Operations Planning System for Turning, 1984, ISBN 91-7372-805-5.
- No 155 **Christos Levcopoulos:** Heuristics for Minimum Decompositions of Polygons, 1987, ISBN 91-7870-133-3.
- No 165 **James W. Goodwin:** A Theory and System for Non-Monotonic Reasoning, 1987, ISBN 91-7870-183-X.
- No 170 **Zebo Peng:** A Formal Methodology for Automated Synthesis of VLSI Systems, 1987, ISBN 91-7870-225-9.
- No 174 **Johan Fagerström:** A Paradigm and System for Design of Distributed Systems, 1988, ISBN 91-7870-301-8.
- No 192 **Dimitar Driankov:** Towards a Many Valued Logic of Quantified Belief, 1988, ISBN 91-7870-374-3.
- No 213 **Lin Padgham:** Non-Monotonic Inheritance for an Object Oriented Knowledge Base, 1989, ISBN 91-7870-485-5.
- No 214 **Tony Larsson:** A Formal Hardware Description and Verification Method, 1989, ISBN 91-7870-517-7.
- No 221 **Michael Reinfrank:** Fundamentals and Logical Foundations of Truth Maintenance, 1989, ISBN 91-7870-546-0.
- No 239 **Jonas Löwgren:** Knowledge-Based Design Support and Discourse Management in User Interface Management Systems, 1991, ISBN 91-7870-720-X.
- No 244 **Henrik Eriksson:** Meta-Tool Support for Knowledge Acquisition, 1991, ISBN 91-7870-746-3.
- No 252 **Peter Eklund:** An Epistemic Approach to Interactive Design in Multiple Inheritance Hierarchies, 1991, ISBN 91-7870-784-6.
- No 258 **Patrick Doherty:** NML3 - A Non-Monotonic Formalism with Explicit Defaults, 1991, ISBN 91-7870-816-8.
- No 260 **Nahid Shahmehri:** Generalized Algorithmic Debugging, 1991, ISBN 91-7870-828-1.
- No 264 **Nils Dahlbäck:** Representation of Discourse-Cognitive and Computational Aspects, 1992, ISBN 91-7870-850-8.
- No 265 **Ulf Nilsson:** Abstract Interpretations and Abstract Machines: Contributions to a Methodology for the Implementation of Logic Programs, 1992, ISBN 91-7870-858-3.
- No 270 **Ralph Rönnquist:** Theory and Practice of Tense-bound Object References, 1992, ISBN 91-7870-873-7.
- No 273 **Björn Fjellborg:** Pipeline Extraction for VLSI Data Path Synthesis, 1992, ISBN 91-7870-880-X.
- No 276 **Staffan Bonnier:** A Formal Basis for Horn Clause Logic with External Polymorphic Functions, 1992, ISBN 91-7870-896-6.
- No 277 **Kristian Sandahl:** Developing Knowledge Management Systems with an Active Expert Methodology, 1992, ISBN 91-7870-897-4.
- No 281 **Christer Bäckström:** Computational Complexity of Reasoning about Plans, 1992, ISBN 91-7870-979-2.
- No 292 **Mats Wirén:** Studies in Incremental Natural Language Analysis, 1992, ISBN 91-7871-027-8.
- No 297 **Mariam Kamkar:** Interprocedural Dynamic Slicing with Applications to Debugging and Testing, 1993, ISBN 91-7871-065-0.
- No 302 **Tingting Zhang:** A Study in Diagnosis Using Classification and Defaults, 1993, ISBN 91-7871-078-2.
- No 312 **Arne Jönsson:** Dialogue Management for Natural Language Interfaces - An Empirical Approach, 1993, ISBN 91-7871-110-X.
- No 338 **Simin Nadjm-Tehrani:** Reactive Systems in Physical Environments: Compositional Modelling and Framework for Verification, 1994, ISBN 91-7871-237-8.
- No 371 **Bengt Savén:** Business Models for Decision Support and Learning. A Study of Discrete-Event Manufacturing Simulation at Asea/ABB 1968-1993, 1995, ISBN 91-7871-494-X.

- No 375 **Ulf Söderman**: Conceptual Modelling of Mode Switching Physical Systems, 1995, ISBN 91-7871-516-4.
- No 383 **Andreas Kägedal**: Exploiting Groundness in Logic Programs, 1995, ISBN 91-7871-538-5.
- No 396 **George Fodor**: Ontological Control, Description, Identification and Recovery from Problematic Control Situations, 1995, ISBN 91-7871-603-9.
- No 413 **Mikael Pettersson**: Compiling Natural Semantics, 1995, ISBN 91-7871-641-1.
- No 414 **Xinli Gu**: RT Level Testability Improvement by Testability Analysis and Transformations, 1996, ISBN 91-7871-654-3.
- No 416 **Hua Shu**: Distributed Default Reasoning, 1996, ISBN 91-7871-665-9.
- No 429 **Jaime Villegas**: Simulation Supported Industrial Training from an Organisational Learning Perspective - Development and Evaluation of the SSIT Method, 1996, ISBN 91-7871-700-0.
- No 431 **Peter Jonsson**: Studies in Action Planning: Algorithms and Complexity, 1996, ISBN 91-7871-704-3.
- No 437 **Johan Boye**: Directional Types in Logic Programming, 1996, ISBN 91-7871-725-6.
- No 439 **Cecilia Sjöberg**: Activities, Voices and Arenas: Participatory Design in Practice, 1996, ISBN 91-7871-728-0.
- No 448 **Patrick Lambrix**: Part-Whole Reasoning in Description Logics, 1996, ISBN 91-7871-820-1.
- No 452 **Kjell Orsbom**: On Extensible and Object-Relational Database Technology for Finite Element Analysis Applications, 1996, ISBN 91-7871-827-9.
- No 459 **Olof Johansson**: Development Environments for Complex Product Models, 1996, ISBN 91-7871-855-4.
- No 461 **Lena Strömbäck**: User-Defined Constructions in Unification-Based Formalisms, 1997, ISBN 91-7871-857-0.
- No 462 **Lars Degerstedt**: Tabulation-based Logic Programming: A Multi-Level View of Query Answering, 1996, ISBN 91-7871-858-9.
- No 475 **Fredrik Nilsson**: Strategi och ekonomisk styrning - En studie av hur ekonomiska styrsystem utformas och används efter företagsförvärv, 1997, ISBN 91-7871-914-3.
- No 480 **Mikael Lindvall**: An Empirical Study of Requirements-Driven Impact Analysis in Object-Oriented Software Evolution, 1997, ISBN 91-7871-927-5.
- No 485 **Göran Forslund**: Opinion-Based Systems: The Cooperative Perspective on Knowledge-Based Decision Support, 1997, ISBN 91-7871-938-0.
- No 494 **Martin Sköld**: Active Database Management Systems for Monitoring and Control, 1997, ISBN 91-7219-002-7.
- No 495 **Hans Olsén**: Automatic Verification of Petri Nets in a CLP framework, 1997, ISBN 91-7219-011-6.
- No 498 **Thomas Drakengren**: Algorithms and Complexity for Temporal and Spatial Formalisms, 1997, ISBN 91-7219-019-1.
- No 502 **Jakob Axelsson**: Analysis and Synthesis of Heterogeneous Real-Time Systems, 1997, ISBN 91-7219-035-3.
- No 503 **Johan Ringström**: Compiler Generation for Data-Parallel Programming Languages from Two-Level Semantics Specifications, 1997, ISBN 91-7219-045-0.
- No 512 **Anna Moberg**: Närhet och distans - Studier av kommunikationsmönster i satellitkontor och flexibla kontor, 1997, ISBN 91-7219-119-8.
- No 520 **Mikael Ronström**: Design and Modelling of a Parallel Data Server for Telecom Applications, 1998, ISBN 91-7219-169-4.
- No 522 **Niclas Ohlsson**: Towards Effective Fault Prevention - An Empirical Study in Software Engineering, 1998, ISBN 91-7219-176-7.
- No 526 **Joachim Karlsson**: A Systematic Approach for Prioritizing Software Requirements, 1998, ISBN 91-7219-184-8.
- No 530 **Henrik Nilsson**: Declarative Debugging for Lazy Functional Languages, 1998, ISBN 91-7219-197-x.
- No 555 **Jonas Hallberg**: Timing Issues in High-Level Synthesis, 1998, ISBN 91-7219-369-7.
- No 561 **Ling Lin**: Management of 1-D Sequence Data - From Discrete to Continuous, 1999, ISBN 91-7219-402-2.
- No 563 **Eva L. Ragnemalm**: Student Modelling based on Collaborative Dialogue with a Learning Companion, 1999, ISBN 91-7219-412-X.
- No 567 **Jörgen Lindström**: Does Distance matter? On geographical dispersion in organisations, 1999, ISBN 91-7219-439-1.
- No 582 **Vanja Josifovski**: Design, Implementation and Evaluation of a Distributed Mediator System for Data Integration, 1999, ISBN 91-7219-482-0.
- No 589 **Rita Kovordányi**: Modeling and Simulating Inhibitory Mechanisms in Mental Image Reinterpretation - Towards Cooperative Human-Computer Creativity, 1999, ISBN 91-7219-506-1.
- No 592 **Mikael Ericsson**: Supporting the Use of Design Knowledge - An Assessment of Commenting Agents, 1999, ISBN 91-7219-532-0.
- No 593 **Lars Karlsson**: Actions, Interactions and Narratives, 1999, ISBN 91-7219-534-7.
- No 594 **C. G. Mikael Johansson**: Social and Organizational Aspects of Requirements Engineering Methods - A practice-oriented approach, 1999, ISBN 91-7219-541-X.
- No 595 **Jörgen Hansson**: Value-Driven Multi-Class Overload Management in Real-Time Database Systems, 1999, ISBN 91-7219-542-8.
- No 596 **Niklas Hallberg**: Incorporating User Values in the Design of Information Systems and Services in the Public Sector: A Methods Approach, 1999, ISBN 91-7219-543-6.
- No 597 **Vivian Vimarlund**: An Economic Perspective on the Analysis of Impacts of Information Technology: From Case Studies in Health-Care towards General Models and Theories, 1999, ISBN 91-7219-544-4.
- No 598 **Johan Jenvald**: Methods and Tools in Computer-Supported Taskforce Training, 1999, ISBN 91-7219-547-9.
- No 607 **Magnus Merkel**: Understanding and enhancing translation by parallel text processing, 1999, ISBN 91-7219-614-9.
- No 611 **Silvia Coradeschi**: Anchoring symbols to sensory data, 1999, ISBN 91-7219-623-8.
- No 613 **Man Lin**: Analysis and Synthesis of Reactive Systems: A Generic Layered Architecture Perspective, 1999, ISBN 91-7219-630-0.
- No 618 **Jimmy Tjäder**: Systemimplementering i praktiken - En studie av logiker i fyra projekt, 1999, ISBN 91-7219-657-2.
- No 627 **Vadim Engelson**: Tools for Design, Interactive Simulation, and Visualization of Object-Oriented Models in Scientific Computing, 2000, ISBN 91-7219-709-9.

- No 637 **Esa Falkenroth:** Database Technology for Control and Simulation, 2000, ISBN 91-7219-766-8.
- No 639 **Per-Arne Persson:** Bringing Power and Knowledge Together: Information Systems Design for Autonomy and Control in Command Work, 2000, ISBN 91-7219-796-X.
- No 660 **Erik Larsson:** An Integrated System-Level Design for Testability Methodology, 2000, ISBN 91-7219-890-7.
- No 688 **Marcus Bjärelund:** Model-based Execution Monitoring, 2001, ISBN 91-7373-016-5.
- No 689 **Joakim Gustafsson:** Extending Temporal Action Logic, 2001, ISBN 91-7373-017-3.
- No 720 **Carl-Johan Petri:** Organizational Information Provision - Managing Mandatory and Discretionary Use of Information Technology, 2001, ISBN-91-7373-126-9.
- No 724 **Paul Scerri:** Designing Agents for Systems with Adjustable Autonomy, 2001, ISBN 91 7373 207 9.
- No 725 **Tim Heyer:** Semantic Inspection of Software Artifacts: From Theory to Practice, 2001, ISBN 91 7373 208 7.
- No 726 **Pär Carlshamre:** A Usability Perspective on Requirements Engineering - From Methodology to Product Development, 2001, ISBN 91 7373 212 5.
- No 732 **Juha Takkinen:** From Information Management to Task Management in Electronic Mail, 2002, ISBN 91 7373 258 3.
- No 745 **Johan Åberg:** Live Help Systems: An Approach to Intelligent Help for Web Information Systems, 2002, ISBN 91-7373-311-3.
- No 746 **Rego Granlund:** Monitoring Distributed Teamwork Training, 2002, ISBN 91-7373-312-1.
- No 757 **Henrik André-Jönsson:** Indexing Strategies for Time Series Data, 2002, ISBN 917373-346-6.
- No 747 **Anneli Hagdahl:** Development of IT-supported Interorganizational Collaboration - A Case Study in the Swedish Public Sector, 2002, ISBN 91-7373-314-8.
- No 749 **Sofie Pilemalm:** Information Technology for Non-Profit Organisations - Extended Participatory Design of an Information System for Trade Union Shop Stewards, 2002, ISBN 91-7373-318-0.
- No 765 **Stefan Holmlid:** Adapting users: Towards a theory of use quality, 2002, ISBN 91-7373-397-0.
- No 771 **Magnus Morin:** Multimedia Representations of Distributed Tactical Operations, 2002, ISBN 91-7373-421-7.
- No 772 **Pawel Pietrzak:** A Type-Based Framework for Locating Errors in Constraint Logic Programs, 2002, ISBN 91-7373-422-5.
- No 758 **Erik Berglund:** Library Communication Among Programmers Worldwide, 2002, ISBN 91-7373-349-0.
- No 774 **Choong-ho Yi:** Modelling Object-Oriented Dynamic Systems Using a Logic-Based Framework, 2002, ISBN 91-7373-424-1.
- No 779 **Mathias Broxvall:** A Study in the Computational Complexity of Temporal Reasoning, 2002, ISBN 91-7373-440-3.
- No 793 **Asmus Pandikow:** A Generic Principle for Enabling Interoperability of Structured and Object-Oriented Analysis and Design Tools, 2002, ISBN 91-7373-479-9.
- No 785 **Lars Hult:** Publika Informationsjänster. En studie av den Internetbaserade encyklopedins bruksegenskap, 2003, ISBN 91-7373-461-6.
- No 800 **Lars Taxén:** A Framework for the Coordination of Complex Systems' Development, 2003, ISBN 91-7373-604-X.
- No 808 **Klas Gäre:** Tre perspektiv på förväntningar och förändringar i samband med införande av informationssystem, 2003, ISBN 91-7373-618-X.
- No 821 **Mikael Kindborg:** Concurrent Comics - programming of social agents by children, 2003, ISBN 91-7373-651-1.
- No 823 **Christina Ölvingson:** On Development of Information Systems with GIS Functionality in Public Health Informatics: A Requirements Engineering Approach, 2003, ISBN 91-7373-656-2.
- No 828 **Tobias Ritzau:** Memory Efficient Hard Real-Time Garbage Collection, 2003, ISBN 91-7373-666-X.
- No 833 **Paul Pop:** Analysis and Synthesis of Communication-Intensive Heterogeneous Real-Time Systems, 2003, ISBN 91-7373-683-X.
- No 852 **Johan Moe:** Observing the Dynamic Behaviour of Large Distributed Systems to Improve Development and Testing - An Empirical Study in Software Engineering, 2003, ISBN 91-7373-779-8.
- No 867 **Erik Herzog:** An Approach to Systems Engineering Tool Data Representation and Exchange, 2004, ISBN 91-7373-929-4.
- No 872 **Aseel Berglund:** Augmenting the Remote Control: Studies in Complex Information Navigation for Digital TV, 2004, ISBN 91-7373-940-5.
- No 869 **Jo Skåmedal:** Telecommuting's Implications on Travel and Travel Patterns, 2004, ISBN 91-7373-935-9.
- No 870 **Linda Askenäs:** The Roles of IT - Studies of Organising when Implementing and Using Enterprise Systems, 2004, ISBN 91-7373-936-7.
- No 874 **Annika Flycht-Eriksson:** Design and Use of Ontologies in Information-Providing Dialogue Systems, 2004, ISBN 91-7373-947-2.
- No 873 **Peter Bunus:** Debugging Techniques for Equation-Based Languages, 2004, ISBN 91-7373-941-3.
- No 876 **Jonas Mellin:** Resource-Predictable and Efficient Monitoring of Events, 2004, ISBN 91-7373-956-1.
- No 883 **Magnus Bång:** Computing at the Speed of Paper: Ubiquitous Computing Environments for Healthcare Professionals, 2004, ISBN 91-7373-971-5.
- No 882 **Robert Eklund:** Disfluency in Swedish human-human and human-machine travel booking dialogues, 2004, ISBN 91-7373-966-9.
- No 887 **Anders Lindström:** English and other Foreign Linguistic Elements in Spoken Swedish. Studies of Productive Processes and their Modelling using Finite-State Tools, 2004, ISBN 91-7373-981-2.
- No 889 **Zhiping Wang:** Capacity-Constrained Production-inventory systems - Modelling and Analysis in both a traditional and an e-business context, 2004, ISBN 91-85295-08-6.
- No 893 **Pernilla Qvarfordt:** Eyes on Multimodal Interaction, 2004, ISBN 91-85295-30-2.
- No 910 **Magnus Kald:** In the Borderland between Strategy and Management Control - Theoretical Framework and Empirical Evidence, 2004, ISBN 91-85295-82-5.
- No 918 **Jonas Lundberg:** Shaping Electronic News: Genre Perspectives on Interaction Design, 2004, ISBN 91-85297-14-3.
- No 900 **Mattias Arvola:** Shades of use: The dynamics of interaction design for sociable use, 2004, ISBN 91-85295-42-6.
- No 920 **Luis Alejandro Cortés:** Verification and Scheduling Techniques for Real-Time Embedded Systems, 2004, ISBN 91-85297-21-6.
- No 929 **Diana Szentivanyi:** Performance Studies of Fault-Tolerant Middleware, 2005, ISBN 91-85297-58-5.

- No 933 **Mikael Cäker:** Management Accounting as Constructing and Opposing Customer Focus: Three Case Studies on Management Accounting and Customer Relations, 2005, ISBN 91-85297-64-X.
- No 937 **Jonas Kvarnström:** TALplanner and Other Extensions to Temporal Action Logic, 2005, ISBN 91-85297-75-5.
- No 938 **Bourhane Kadmiry:** Fuzzy Gain-Scheduled Visual Servoing for Unmanned Helicopter, 2005, ISBN 91-85297-76-3.
- No 945 **Gert Jervan:** Hybrid Built-In Self-Test and Test Generation Techniques for Digital Systems, 2005, ISBN: 91-85297-97-6.
- No 946 **Anders Arpteg:** Intelligent Semi-Structured Information Extraction, 2005, ISBN 91-85297-98-4.
- No 947 **Ola Angelmark:** Constructing Algorithms for Constraint Satisfaction and Related Problems - Methods and Applications, 2005, ISBN 91-85297-99-2.
- No 963 **Calin Curescu:** Utility-based Optimisation of Resource Allocation for Wireless Networks, 2005, ISBN 91-85457-07-8.
- No 972 **Björn Johansson:** Joint Control in Dynamic Situations, 2005, ISBN 91-85457-31-0.
- No 974 **Dan Lawesson:** An Approach to Diagnosability Analysis for Interacting Finite State Systems, 2005, ISBN 91-85457-39-6.
- No 979 **Claudiu Duma:** Security and Trust Mechanisms for Groups in Distributed Services, 2005, ISBN 91-85457-54-X.
- No 983 **Sorin Manolache:** Analysis and Optimisation of Real-Time Systems with Stochastic Behaviour, 2005, ISBN 91-85457-60-4.
- No 986 **Yuxiao Zhao:** Standards-Based Application Integration for Business-to-Business Communications, 2005, ISBN 91-85457-66-3.
- No 1004 **Patrik Haslum:** Admissible Heuristics for Automated Planning, 2006, ISBN 91-85497-28-2.
- No 1005 **Aleksandra Tešanovic:** Developing Reusable and Reconfigurable Real-Time Software using Aspects and Components, 2006, ISBN 91-85497-29-0.
- No 1008 **David Dinka:** Role, Identity and Work: Extending the design and development agenda, 2006, ISBN 91-85497-42-8.
- No 1009 **Iakov Nakhimovski:** Contributions to the Modeling and Simulation of Mechanical Systems with Detailed Contact Analysis, 2006, ISBN 91-85497-43-X.
- No 1013 **Wilhelm Dahllöf:** Exact Algorithms for Exact Satisfiability Problems, 2006, ISBN 91-85523-97-6.
- No 1016 **Levon Saldamli:** PDEModelica - A High-Level Language for Modeling with Partial Differential Equations, 2006, ISBN 91-85523-84-4.
- No 1017 **Daniel Karlsson:** Verification of Component-based Embedded System Designs, 2006, ISBN 91-85523-79-8
- No 1018 **Ioan Chisalita:** Communication and Networking Techniques for Traffic Safety Systems, 2006, ISBN 91-85523-77-1.
- No 1019 **Tarja Susi:** The Puzzle of Social Activity - The Significance of Tools in Cognition and Cooperation, 2006, ISBN 91-85523-71-2.
- No 1021 **Andrzej Bednarski:** Integrated Optimal Code Generation for Digital Signal Processors, 2006, ISBN 91-85523-69-0.
- No 1022 **Peter Aronsson:** Automatic Parallelization of Equation-Based Simulation Programs, 2006, ISBN 91-85523-68-2.
- No 1030 **Robert Nilsson:** A Mutation-based Framework for Automated Testing of Timeliness, 2006, ISBN 91-85523-35-6.
- No 1034 **Jon Edvardsson:** Techniques for Automatic Generation of Tests from Programs and Specifications, 2006, ISBN 91-85523-31-3.
- No 1035 **Vaida Jakoniene:** Integration of Biological Data, 2006, ISBN 91-85523-28-3.
- No 1045 **Genevieve Gorrell:** Generalized Hebbian Algorithms for Dimensionality Reduction in Natural Language Processing, 2006, ISBN 91-85643-88-2.
- No 1051 **Yu-Hsing Huang:** Having a New Pair of Glasses - Applying Systemic Accident Models on Road Safety, 2006, ISBN 91-85643-64-5.
- No 1054 **Åsa Hedenskog:** Perceive those things which cannot be seen - A Cognitive Systems Engineering perspective on requirements management, 2006, ISBN 91-85643-57-2.
- No 1061 **Cécile Åberg:** An Evaluation Platform for Semantic Web Technology, 2007, ISBN 91-85643-31-9.
- No 1073 **Mats Gröndal:** Handling Combinatorial Explosion in Software Testing, 2007, ISBN 978-91-85715-74-9.
- No 1075 **Almut Herzog:** Usable Security Policies for Runtime Environments, 2007, ISBN 978-91-85715-65-7.
- No 1079 **Magnus Wahlström:** Algorithms, measures, and upper bounds for Satisfiability and related problems, 2007, ISBN 978-91-85715-55-8.
- No 1083 **Jesper Andersson:** Dynamic Software Architectures, 2007, ISBN 978-91-85715-46-6.
- No 1086 **Ulf Johansson:** Obtaining Accurate and Comprehensible Data Mining Models - An Evolutionary Approach, 2007, ISBN 978-91-85715-34-3.
- No 1089 **Traian Pop:** Analysis and Optimisation of Distributed Embedded Systems with Heterogeneous Scheduling Policies, 2007, ISBN 978-91-85715-27-5.
- No 1091 **Gustav Nordh:** Complexity Dichotomies for CSP-related Problems, 2007, ISBN 978-91-85715-20-6.
- No 1106 **Per Ola Kristensson:** Discrete and Continuous Shape Writing for Text Entry and Control, 2007, ISBN 978-91-85831-77-7.
- No 1110 **He Tan:** Aligning Biomedical Ontologies, 2007, ISBN 978-91-85831-56-2.
- No 1112 **Jessica Lindblom:** Minding the body - Interacting socially through embodied action, 2007, ISBN 978-91-85831-48-7.
- No 1113 **Pontus Wärmestål:** Dialogue Behavior Management in Conversational Recommender Systems, 2007, ISBN 978-91-85831-47-0.
- No 1120 **Thomas Gustafsson:** Management of Real-Time Data Consistency and Transient Overloads in Embedded Systems, 2007, ISBN 978-91-85831-33-3.
- No 1127 **Alexandru Andrei:** Energy Efficient and Predictable Design of Real-time Embedded Systems, 2007, ISBN 978-91-85831-06-7.
- No 1139 **Per Wikberg:** Eliciting Knowledge from Experts in Modeling of Complex Systems: Managing Variation and Interactions, 2007, ISBN 978-91-85895-66-3.
- No 1143 **Mehdi Amirjoo:** QoS Control of Real-Time Data Services under Uncertain Workload, 2007, ISBN 978-91-85895-49-6.
- No 1150 **Sanny Syberfeldt:** Optimistic Replication with Forward Conflict Resolution in Distributed Real-Time Databases, 2007, ISBN 978-91-85895-27-4.
- No 1155 **Beatrice Alenljung:** Envisioning a Future Decision Support System for Requirements Engineering - A Holistic and Human-centred Perspective, 2008, ISBN 978-91-85895-11-3.

- No 1156 **Artur Wilk:** Types for XML with Application to Xcerpt, 2008, ISBN 978-91-85895-08-3.
- No 1183 **Adrian Pop:** Integrated Model-Driven Development Environments for Equation-Based Object-Oriented Languages, 2008, ISBN 978-91-7393-895-2.
- No 1185 **Jörgen Skågeby:** Gifting Technologies - Ethnographic Studies of End-users and Social Media Sharing, 2008, ISBN 978-91-7393-892-1.
- No 1187 **Imad-Eldin Ali Abugessaisa:** Analytical tools and information-sharing methods supporting road safety organizations, 2008, ISBN 978-91-7393-887-7.
- No 1204 **H. Joe Steinhauer:** A Representation Scheme for Description and Reconstruction of Object Configurations Based on Qualitative Relations, 2008, ISBN 978-91-7393-823-5.
- No 1222 **Anders Larsson:** Test Optimization for Core-based System-on-Chip, 2008, ISBN 978-91-7393-768-9.
- No 1238 **Andreas Borg:** Processes and Models for Capacity Requirements in Telecommunication Systems, 2009, ISBN 978-91-7393-700-9.
- No 1240 **Fredrik Heintz:** DyKnow: A Stream-Based Knowledge Processing Middleware Framework, 2009, ISBN 978-91-7393-696-5.
- No 1241 **Birgitta Lindström:** Testability of Dynamic Real-Time Systems, 2009, ISBN 978-91-7393-695-8.
- No 1244 **Eva Blomqvist:** Semi-automatic Ontology Construction based on Patterns, 2009, ISBN 978-91-7393-683-5.
- No 1249 **Rogier Woltjer:** Functional Modeling of Constraint Management in Aviation Safety and Command and Control, 2009, ISBN 978-91-7393-659-0.
- No 1260 **Gianpaolo Conte:** Vision-Based Localization and Guidance for Unmanned Aerial Vehicles, 2009, ISBN 978-91-7393-603-3.
- No 1262 **AnnMarie Ericsson:** Enabling Tool Support for Formal Analysis of ECA Rules, 2009, ISBN 978-91-7393-598-2.
- No 1266 **Jiri Trnka:** Exploring Tactical Command and Control: A Role-Playing Simulation Approach, 2009, ISBN 978-91-7393-571-5.
- No 1268 **Bahlol Rahimi:** Supporting Collaborative Work through ICT - How End-users Think of and Adopt Integrated Health Information Systems, 2009, ISBN 978-91-7393-550-0.
- No 1274 **Fredrik Kuivinen:** Algorithms and Hardness Results for Some Valued CSPs, 2009, ISBN 978-91-7393-525-8.
- No 1281 **Gunnar Mathiason:** Virtual Full Replication for Scalable Distributed Real-Time Databases, 2009, ISBN 978-91-7393-503-6.
- No 1290 **Viacheslav Izosimov:** Scheduling and Optimization of Fault-Tolerant Distributed Embedded Systems, 2009, ISBN 978-91-7393-482-4.
- No 1294 **Johan Thapper:** Aspects of a Constraint Optimisation Problem, 2010, ISBN 978-91-7393-464-0.
- No 1306 **Susanna Nilsson:** Augmentation in the Wild: User Centered Development and Evaluation of Augmented Reality Applications, 2010, ISBN 978-91-7393-416-9.
- No 1313 **Christer Thörn:** On the Quality of Feature Models, 2010, ISBN 978-91-7393-394-0.
- No 1321 **Zhiyuan He:** Temperature Aware and Defect-Probability Driven Test Scheduling for System-on-Chip, 2010, ISBN 978-91-7393-378-0.
- No 1333 **David Broman:** Meta-Languages and Semantics for Equation-Based Modeling and Simulation, 2010, ISBN 978-91-7393-335-3.
- No 1337 **Alexander Siemers:** Contributions to Modelling and Visualisation of Multibody Systems Simulations with Detailed Contact Analysis, 2010, ISBN 978-91-7393-317-9.
- No 1354 **Mikael Asplund:** Disconnected Discoveries: Availability Studies in Partitioned Networks, 2010, ISBN 978-91-7393-278-3.
- No 1359 **Jana Rambusch:** Mind Games Extended: Understanding Gameplay as Situated Activity, 2010, ISBN 978-91-7393-252-3.
- No 1373 **Sonia Sangari:** Head Movement Correlates to Focus Assignment in Swedish, 2011, ISBN 978-91-7393-154-0.
- No 1374 **Jan-Erik Källhammer:** Using False Alarms when Developing Automotive Active Safety Systems, 2011, ISBN 978-91-7393-153-3.
- No 1375 **Mattias Eriksson:** Integrated Code Generation, 2011, ISBN 978-91-7393-147-2.
- No 1381 **Ola Leiffer:** Affordances and Constraints of Intelligent Decision Support for Military Command and Control - Three Case Studies of Support Systems, 2011, ISBN 978-91-7393-133-5.
- No 1386 **Soheil Samii:** Quality-Driven Synthesis and Optimization of Embedded Control Systems, 2011, ISBN 978-91-7393-102-1.
- No 1419 **Erik Kuiper:** Geographic Routing in Intermittently-connected Mobile Ad Hoc Networks: Algorithms and Performance Models, 2012, ISBN 978-91-7519-981-8.
- No 1451 **Sara Stymne:** Text Harmonization Strategies for Phrase-Based Statistical Machine Translation, 2012, ISBN 978-91-7519-887-3.
- No 1455 **Alberto Montebelli:** Modeling the Role of Energy Management in Embodied Cognition, 2012, ISBN 978-91-7519-882-8.
- No 1465 **Mohammad Saifullah:** Biologically-Based Interactive Neural Network Models for Visual Attention and Object Recognition, 2012, ISBN 978-91-7519-838-5.
- No 1490 **Tomas Bengtsson:** Testing and Logic Optimization Techniques for Systems on Chip, 2012, ISBN 978-91-7519-742-5.
- No 1481 **David Byers:** Improving Software Security by Preventing Known Vulnerabilities, 2012, ISBN 978-91-7519-784-5.
- No 1496 **Tommy Färnqvist:** Exploiting Structure in CSP-related Problems, 2013, ISBN 978-91-7519-711-1.
- No 1503 **John Wilander:** Contributions to Specification, Implementation, and Execution of Secure Software, 2013, ISBN 978-91-7519-681-7.
- No 1506 **Magnus Ingmarsson:** Creating and Enabling the Useful Service Discovery Experience, 2013, ISBN 978-91-7519-662-6.
- No 1547 **Wladimir Schamai:** Model-Based Verification of Dynamic System Behavior against Requirements: Method, Language, and Tool, 2013, ISBN 978-91-7519-505-6.
- No 1551 **Henrik Svensson:** Simulations, 2013, ISBN 978-91-7519-491-2.
- No 1559 **Sergiu Rafiliu:** Stability of Adaptive Distributed Real-Time Systems with Dynamic Resource Management, 2013, ISBN 978-91-7519-471-4.
- No 1581 **Usman Dastgeer:** Performance-aware Component Composition for GPU-based Systems, 2014, ISBN 978-91-7519-383-0.
- No 1602 **Cai Li:** Reinforcement Learning of Locomotion based on Central Pattern Generators, 2014, ISBN 978-91-7519-313-7.
- No 1652 **Roland Samlaus:** An Integrated Development Environment with Enhanced Domain-Specific

- Interactive Model Validation, 2015, ISBN 978-91-7519-090-7.
- No 1663 **Hannes Uppman**: On Some Combinatorial Optimization Problems: Algorithms and Complexity, 2015, ISBN 978-91-7519-072-3.
- No 1664 **Martin Sjölund**: Tools and Methods for Analysis, Debugging, and Performance Improvement of Equation-Based Models, 2015, ISBN 978-91-7519-071-6.
- No 1666 **Kristian Stavåker**: Contributions to Simulation of Modelica Models on Data-Parallel Multi-Core Architectures, 2015, ISBN 978-91-7519-068-6.
- No 1680 **Adrian Lifa**: Hardware/ Software Codesign of Embedded Systems with Reconfigurable and Heterogeneous Platforms, 2015, ISBN 978-91-7519-040-2.
- No 1685 **Bogdan Tanasa**: Timing Analysis of Distributed Embedded Systems with Stochastic Workload and Reliability Constraints, 2015, ISBN 978-91-7519-022-8.
- No 1691 **Håkan Warnquist**: Troubleshooting Trucks – Automated Planning and Diagnosis, 2015, ISBN 978-91-7685-993-3.
- No 1702 **Nima Aghaee**: Thermal Issues in Testing of Advanced Systems on Chip, 2015, ISBN 978-91-7685-949-0.
- No 1715 **Maria Vasilevska**: Security in Embedded Systems: A Model-Based Approach with Risk Metrics, 2015, ISBN 978-91-7685-917-9.
- No 1729 **Ke Jiang**: Security-Driven Design of Real-Time Embedded System, 2016, ISBN 978-91-7685-884-4.
- No 1733 **Victor Lagerkvist**: Strong Partial Clones and the Complexity of Constraint Satisfaction Problems: Limitations and Applications, 2016, ISBN 978-91-7685-856-1.
- No 1734 **Chandan Roy**: An Informed System Development Approach to Tropical Cyclone Track and Intensity Forecasting, 2016, ISBN 978-91-7685-854-7.
- No 1746 **Amir Aminifar**: Analysis, Design, and Optimization of Embedded Control Systems, 2016, ISBN 978-91-7685-826-4.
- No 1747 **Ekhioz Vergara**: Energy Modelling and Fairness for Efficient Mobile Communication, 2016, ISBN 978-91-7685-822-6.
- No 1748 **Dag Sonntag**: Chain Graphs – Interpretations, Expressiveness and Learning Algorithms, 2016, ISBN 978-91-7685-818-9.

Linköping Studies in Arts and Science

- No 504 **Ing-Marie Jonsson**: Social and Emotional Characteristics of Speech-based In-Vehicle Information Systems: Impact on Attitude and Driving Behaviour, 2009, ISBN 978-91-7393-478-7.
- No 586 **Fabian Segelström**: Stakeholder Engagement for Service Design: How service designers identify and communicate insights, 2013, ISBN 978-91-7519-554-4.
- No 618 **Johan Blomkvist**: Representing Future Situations of Service: Prototyping in Service Design, 2014, ISBN 978-91-7519-343-4.
- No 620 **Marcus Mast**: Human-Robot Interaction for Semi-Autonomous Assistive Robots, 2014, ISBN 978-91-7519-319-9.
- No 677 **Peter Berggren**: Assessing Shared Strategic Understanding, 2016, ISBN 978-91-7685-786-1.

Linköping Studies in Statistics

- No 9 **Davood Shahsavani**: Computer Experiments Designed to Explore and Approximate Complex Deterministic Models, 2008, ISBN 978-91-7393-976-8.
- No 10 **Karl Wahlin**: Roadmap for Trend Detection and Assessment of Data Quality, 2008, ISBN 978-91-7393-792-4.
- No 11 **Oleg Sysoev**: Monotonic regression for large multivariate datasets, 2010, ISBN 978-91-7393-412-1.
- No 13 **Agné Burauskaite-Harju**: Characterizing Temporal Change and Inter-Site Correlations in Daily and Sub-daily Precipitation Extremes, 2011, ISBN 978-91-7393-110-6.

Linköping Studies in Information Science

- No 1 **Karin Axelsson**: Metodisk systemstrukturering - att skapa samstämmighet mellan informationssystemarkitektur och verksamhet, 1998, ISBN 91-72-19-296-8.
- No 2 **Stefan Cronholm**: Metodverktyg och användbarhet - en studie av datorstödd metodbaserad systemutveckling, 1998, ISBN 91-72-19-299-2.
- No 3 **Anders Avdic**: Användare och utvecklare - om anveckling med kalkylprogram, 1999, ISBN 91-72-19-606-8.
- No 4 **Owen Eriksson**: Kommunikationskvalitet hos informationssystem och affärsprocesser, 2000, ISBN 91-72-19-811-7.
- No 5 **Mikael Lind**: Från system till process - kriterier för processbestämning vid verksamhetsanalys, 2001, ISBN 91-73-73-067-X.
- No 6 **Ulf Melin**: Koordination och informationssystem i företag och nätverk, 2002, ISBN 91-73-73-278-8.
- No 7 **Pär J. Ågerfalk**: Information Systems Actability - Understanding Information Technology as a Tool for Business Action and Communication, 2003, ISBN 91-73-73-628-7.
- No 8 **Ulf Seigerroth**: Att förstå och förändra systemutvecklingsverksamheter - en taxonomi för metautveckling, 2003, ISBN 91-73-73-736-4.
- No 9 **Karin Hedström**: Spår av datoriseringens värden - Effekter av IT i äldreomsorg, 2004, ISBN 91-73-73-963-4.
- No 10 **Ewa Braf**: Knowledge Demanded for Action - Studies on Knowledge Mediation in Organisations, 2004, ISBN 91-85295-47-7.
- No 11 **Fredrik Karlsson**: Method Configuration method and computerized tool support, 2005, ISBN 91-85297-48-8.
- No 12 **Malin Nordström**: Styrbar systemförvaltning - Att organisera systemförvaltningsverksamhet med hjälp av effektiva förvaltningsobjekt, 2005, ISBN 91-85297-60-7.
- No 13 **Stefan Holgersson**: Yrke: POLIS - Yrkeskunskap, motivation, IT-system och andra förutsättningar för polisarbete, 2005, ISBN 91-85299-43-X.
- No 14 **Benneth Christiansson, Marie-Therese Christiansson**: Mötet mellan process och komponent - mot ett ramverk för en verksamhetsnära kravspecifikation vid anskaffning av komponentbaserade informationssystem, 2006, ISBN 91-85643-22-X.