Linköping Studies in Science and Technology
Dissertation No. 1132

Differences in Situational Awareness and How to Manage Them in Development of Complex Systems

Jens Alfredson

LiU-2007
Graduate School for Human-Machine Interaction
Division of Industrial Ergonomics
Department of Management and Engineering
Linköping University, SE-581 83 Linköping, Sweden

Linköping 2007

Linköping University
Differences in situational awareness and how to manage them in development of complex systems

ISBN: **978-91-85895-87-8**
ISSN: **0345-7524**
Dissertation No. 1132
Linköping 2007
Printed by LiU-Tryck, Linköping, Sweden

Abstract

What's up, Doc? Situational awareness (SA) is about being aware of what is going on. Already when a complex system is developed there is an opportunity to help a future user of the system to form a better SA. Let us make the best out of this opportunity! When assessing SA, differences in SA will sometimes appear. This dissertation is about SA, and how to manage differences in SA in development of complex systems. This topic is highly valid for development of a variety of complex systems, although most examples in this dissertation are from the aviation domain.

Framed by state of the art literature, suggestions are made on theoretical improvements of SA theory, with a focus on differences. The difference between what you are required to be aware of and what you are aware of is suggested as a SA-indicator. Also, the difference between what you are aware of and what you think you are aware of is suggested as another SA-indicator. Further, differences within a team such as variations in degree of agreement could be used for team SA assessment.

Also, the term situation management (SM) is suggested, with a proposed wider meaning than SA, including SA *and* every part of the perception action cycle, the management of mental resources, and external means of managing the situation. SM is a suitable term when developing complex systems due to the focus on the situation and how that could be managed, instead of only focusing on what is perceived by an individual or team.

Assessing differences in SA and to differentiate between various types of differences are recognised as important prerequisites to effectively manage differences in SA in development of complex systems. Several assessment techniques are reviewed and especially advantages and disadvantages of the use of eye movements for SA assessment are described. With reference to the literature as well as to the appended papers differences in SA due to a) design alternatives, b) roles in the design-use process, c) context, and d) level of analysis, are described. Differences in SA are suggested to be regarded as both quantitative (i.e. high or low SA) and qualitative (e.g. various aspects of a situation are regarded).

Approaches such as, SM, on-line evaluation of SA, simulator based design, as well as measuring and analysing SA on multiple levels simultaneously, are suggested as means to manage differences in SA in the development of complex systems.

Sammanfattning

Situationsmedvetenhet (Eng. Situational Awareness), (SA), handlar om att ha koll på läget och vara medveten om vad som händer. Redan då ett komplext system utvecklas får vi en möjlighet att påverka vilken SA en framtida användare av systemet kan komma att få. Det gäller att ta tillvara på detta tillfälle! Ibland uppträder skillnader i SA, beroende på en rad olika orsaker. Denna avhandling handlar om SA och hur man kan använda de skillnaderna vid utveckling av komplexa system. Detta är relevant vid utveckling av en rad olika typer av komplexa system, även om de flesta exempel i denna avhandling kommer från flygdomänen.

Avhandlingen innehåller beskrivningar hämtade från litteratur inom området och förslag på utveckling av SA-teori utifrån fokus på just skillnader. Skillnaden mellan vad du behöver vara medveten om och vad du verkligen är medveten om föreslås ge en indikation om individens SA. Vidare föreslås skillnaden mellan vad du är medveten om och vad du tror dig vara medveten om också ge en indikation om individens SA. SA kan skattas för en grupp av människor som arbetar tillsammans, genom variationerna i hur samstämmiga deras uppfattningar är.

Termen situationshantering (Eng. Situation Management), (SM), föreslås med en vidare mening än SA, inkluderande SA, men också varje del av perceptionscykeln, hantering av mentala resurser och hantering av situationen genom extern påverkan. SM är en väl lämpad term vid utveckling av komplexa system då fokus här är på situationen och hur den kan hanteras, snarare än fokus på vad en individ eller en grupp uppfattar.

Att skatta skillnader i SA och att kunna särskilja olika typer av skillnader är viktiga förutsättningar för att kunna hantera skillnader i SA vid utveckling av komplexa system på ett bra sätt. I avhandlingen gås flera sätt att skatta sådana skillnader igenom och speciellt tas för- och nackdelar med ögonrörelsemätning upp. Med referens till litteraturen och till de bilagda artiklarna beskrivs skillnader i SA beroende på a) designalternativ, b) roller i processen från utveckling till användning c) kontext och d) analysnivå. Skillnaderna i SA föreslås ses som både kvantitativa (dvs. hög eller låg SA) och kvalitativa (tex. olika aspekter av en situation).

Ansatser såsom SM, realtidsvärdering, mätning och analys av SA på flera nivåer samtidigt samt simulatorbaserad design föreslås för att hantera skillnader i SA vid utveckling av komplexa system.

Acknowledgements

First, I would like to thank my academic supervisors, Kjell Ohlsson, Gunilla Derefeldt, Erland Svensson, and Martin Helander for their guidance and support.

Thanks also to my co-authors and colleagues who have worked with me and helped me in this research.

Finally, I would like to thank my current manager, Rikard Andersson, and my former manager, Jan Westlund, for the time and resources provided to me as a PhD student working for Saab.

Appended Papers

1) Derefeldt, G., Skinnars, Ö., Alfredson, J., Eriksson, L., Andersson, P., Westlund, J., Berggrund, U., Holmberg, J., & Santesson, R. (1999). Improvement of tactical situation awareness with colour-coded horizontal-situation displays in combat aircraft. *Displays 20*(4), 171-184.

2) Alfredson, J., & Ohlsson, K. (2000). Pilots' understanding of situational awareness. In P. Wright, E. Hollnagel, & S. Dekker (Eds.), *Proceedings of the 10th European Conference on Cognitive Ergonomics* (pp. 172-181). Linköping, Sweden: University of Linköping.

3) Björklund, C., Alfredson, J., & Dekker, S. (2006). Mode monitoring and call-outs: An eye-tracking study of 2-crew automated flight deck operations. *The International Journal of Aviation Psychology, 16* (3), 263-275.

4) Alfredson, J., & Andersson, J. (Submitted). Effect of task complexity and decision maker separation on shared situational awareness in command and control in network centric warfare. *Military Psychology*.

5) Alfredson, J. (2002). Individual differences in visual behaviour in simulated flight. In B. E. Rogowitz & T. N. Pappas (Eds.), *Proceedings of IS&T/SPIE's 14th Annual Symposium: Vol. 4662. Human Vision and Electronic Imaging VII* (pp. 494-502). Bellingham, WA: SPIE Press.

6) Alm, T., Alfredson, J., & Ohlsson, K. (2007). Business process reengineering in the automotive area by simulator-based design. In A. El Sheikh, A. T. Al Ajeeli, & E. M. Abu-Taieh (Eds.), *Simulation and Modeling: Current Technologies and Applications* (pp. 337-358). Hershey, PA: IGI Global.

1 Introduction

Situational Awareness (SA) is a frequently used term in Human Factors (HF) literature, especially in the aviation domain. It is a term used by pilots, researchers, developers and more. What SA is about is to know what is going on in the past, present and the future.

SA is a central concept for complex systems with human in the loop, partly due to its positive correlation to performance. Therefore, it is important to regard SA already in the development of complex systems, to understand what it is, what dimensions it has, and how to manage those insights for development purposes. If we learn how to manage differences in SA in the development of complex systems, the value of the systems in use will increase. Complex systems appear in many sectors of our society and the applicability of knowledge on their development therefore has great potential.

Although this dissertation provides most examples from the aviation domain, examples from other domains, the automotive domain and the area of command and control are provided. In the aviation domain SA has been extensively investigated, and more so than in any other domain. However, the applicability of results to development of complex systems has not always been in focus.

As the focus of this dissertation is on SA related to development of complex system rather than a general SA research interest, the technical design is of special interest and also the effects of design on the future use of the system, in terms of SA. Therefore, differences in design alternatives, roles in the design-use process, context, and level of analysis is will all be addressed. A user of a product developed now, active in a future situation, will experience variations in SA depending on how that product is designed.

The aircraft cockpit, with its design, represents one part of the interface between the aircraft system and the pilot. The other part of the interface is the parts of the pilot that interact with the cockpit. Since the possibilities of making design changes in the human part of the interface are very limited, the natural focus of the interface design is on the cockpit.

The cockpit design is important, since a good cockpit design is one way to limit the number of human errors. Human error is a major contributor to accidents and incidents (Nagel, 1988), which holds as an argument even if it is vital to be sceptical to error-counting in general which has been explained by, for instance, Dekker (2003; 2005; 2007). Since aircraft accidents can be costly, both economically and in other ways, it is important to minimise the risk of an accident, by trying to optimise the cockpit design. By improving the man-machine interaction, so that the system performance is enhanced by, for instance, better decision making, the risks of an accident or other undesirable events are minimised.

The design of a cockpit and the displays within it very much influences the pilot's ability to execute his/her tasks in the right way. This holds true both for a military and civil aircraft, so the design of various aircraft with different purposes might benefit from some common ideas. The question of positive and negative effects of a conjunction of civil and military research in aircraft HF has been raised by Angelborg-Thanderz (1997). However, not much research has been conducted on SA for general aviation pilots (Shook & Garland, 2000), although there are a few examples of studies on general aviation pilot SA in the literature (e.g. Bolstad, Endsley, & Howell, 2002). The tasks that are to be carried out in a cockpit are very varying, depending on what type of aircraft the cockpit is a part of and the purpose of that aircraft. Even so, there are activities of common nature even between a civil and a military aircraft. A pilot has to aviate, navigate and communicate.

Also, interaction between other users than pilots and other complex systems than aircraft could benefit from the same understanding of SA and its applicability to the development of a complex system. By addressing the issue of SA for the development of complex systems and at the same time keeping in mind that every context is unique, the technical development might be enhanced to better support the management of future situations, where the complex systems are to be used.

To create good design, information about the design criteria is needed. To achieve good design it is useful to focus on a set of sub-goals in the process towards the good design. Examples of such sub-goals are, to reach a good level of pilot mental workload or to help forming good SA. Good SA is the foundation for effective decision making (Endsley, 1995b). Endsley, Bolstad, Jones, and Riley (2003) as well as Endsley, Bolté, and Jones (2003) describe SA oriented design as comprised by three components: SA requirements analysis, SA-oriented design principles, and SA measurement and validation.

It is essential to know what individuals are aware of in a situation, and to understand why they are acting the way they are. Design is about making choices and it is therefore particularly relevant to focus on the differences in SA, since the differences could guide the design choices.

1.1 Problem and Objectives

Complex systems are in themselves costly and their rational is that they provide high values in the settings that they are used in. If the system fails to function or functions suboptimal the consequences for this could be costly, severe or even devastating. A specific problem is when a situation could not be managed, due to lack of SA. This dissertation is focused on that problem and how the development of complex systems could support the management of such a situation. In particular, the dissertation addresses the problem of how to manage differences in SA in development of complex systems.

This problem is present in development of a variety of complex systems, and in application areas where complex systems are used. Professionals that have an interest in the development process of complex systems either as researchers, managers or developers of such a process, or professionals active in the process, have all been anticipated readers of this dissertation. The reader will be provided with a report on how to regard the concept of SA in development of complex systems, and provided with suggestions on how to manage differences in SA as means of coping with the declared problem, according to the dissertation objectives.

The objectives of this dissertation have been to:

- Review and explore the concept of SA, with relevance to the development of complex systems.
- Examine and describe differences in SA in theory and through empirical evaluations.
- Analyse and suggest how differences in SA could be managed in development of complex systems.

The effort to review and explore the concept of SA, with relevance to development of complex systems is concentrated to the chapter *situational awareness*.

The effort to examine and describe differences is SA in theory and through empirical evaluations is concentrated to the chapter *differences in situational awareness*.

The effort to analyse and suggest how differences in SA could be managed in development of complex systems is concentrated to the chapter *managing differences in situational awareness in the development of complex systems*.

1.2 Overview of the Dissertation

SA can differ both quantitatively, that is, more or less SA, and qualitatively, that is, various aspects of SA (Alfredson, 2001a). This dissertation focuses on differences due to design alternatives, due to roles in the design-use process, due to context, and due to level of analysis. The dissertation also aims at giving guidance on how to manage these differences in the development of complex systems. By recognizing the differences and not only focus on one absolute value, the differences in themselves can be used as valuable information in a process of developing complex systems.

The differences in SA due to design alternatives, is one of the most commonly used differences, for guidance in development processes. For instance, if SA is assessed for two candidate design concepts, quantitative differences could guide the design decision towards the concept generating the highest SA, and qualitative differences could give valuable guidance for further development of refined concepts in an iterative design process.

The differences in SA due to roles in the design-use process, refers to the notions of SA that is formed by actors in various roles, such as the user, the designer, the researcher, the manager, and so forth. The roles could due to different information provided to them, different experience and training background, different tasks and other role-specific perspectives differ in their notion of both the concept of SA, as such, and the notion of a specific person's, team's or role's (such as a group of users) SA.

The differences in SA due to context, appears due to that SA, by definition, is situation dependent and contextual fluctuations over time, between domains and settings in which it is studied, and so forth, will affect the SA of the person experiencing it as well as a person trying to assess the SA. To know how to interpret and best make use of our assessments of SA we need to understand the differences due to context, to be able to constructively use them in the development process.

The differences in SA due to level of analysis are not differences in the actual SA, but instead differences that appears for the person assessing SA dependent of how detailed the analysis is made. Since SA is not, and perhaps cannot be, measured in an absolute sense, but instead relative to something, such as, between persons, conditions, before-after, compared to some stipulated norm, such as defined objects of interest to be aware of, and so forth, is it very important to be aware of the differences appearing dependent on the level of analysis, since if the analysis is only performed on one level, the actions made in the development process based on that analysis may be misleading. In other words, it is hard to know the relevance and significance of the outcome of the analysis, if it is only based on relative assessment and only performed on one level. If you would only be able to perform detailed enough analysis you would always find differences in SA, even how unimportant they may be, and on the other hand, with a very high level analysis you may not detect differences in SA important for design decisions.

The dissertation takes its starting-point in the concept of SA, by addressing how SA, as a concept, can be perceived, comprehended and projected for the future. Parts of the text in this dissertation have been presented previously, partially rephrased, in Alfredson (2001a). The following chapter is explicitly focusing on differences in SA, how to assess them and provides examples of various types of differences. The next chapter is about how to manage differences in SA in the development of complex systems, also based on the findings of the appended papers together with knowledge from the literature. Approaches that are discussed in this chapter is, for instance, SM, on-line evaluation of SA, simulator based design, as well as measuring and analysing SA on multiple levels simultaneously. Last, some final conclusions are proposed.

The appended papers reports from various contexts, such as, military/commercial aviation, command and control, and the automotive area, using various methods (e.g. various types of questionnaires, literature review eye-tracking etc.) in simulator settings as well as real world settings. What is common for theses papers is that they all contribute to the understanding of how differences in SA can be managed in the development of complex systems. Below, a short summary follows on how the content in the appended papers links into the rest of the dissertation.

Paper 1, contributes to the section on differences in SA due to design alternatives with data from a simulation where three design alternatives of colour schemes for combat aircraft horizontal-situation display were evaluated. The tested dichrome and polychrome colour schemes received higher SA ratings by the pilots in the simulation, than the monochrome alternative. The colour design, thereby, could be regarded as a factor for tactical SA under the studied conditions. The paper also contributes to the chapter on managing differences in SA in the development of complex systems as an example of how simulator-based design could be used to guide the design towards solutions promoting good SA, which also Paper 6 describes. Since the colour schemes were specified in a standardized colour appearance system, and because of the adequate fidelity in the simulator, the findings in this study thereby were possible to use for further design of fighter aircraft horizontal-situation displays.

Paper 2, contributes to the section on differences in SA due to roles in the design-use process with data on the pilots' view on SA, related to the views reflected in literature. The paper presents results of a questionnaire distributed to a number of, British military pilots and Swedish private pilots. No major differences among pilots on how to interpret the concept of SA were found. However, the deviations between the view of the pilots and the view reflected in literature were found. This implies that there might be differences in SA due to roles in the design process, that is, between the pilots and the producers of the view in research literature (such as researchers) and thereby possibly also between pilots and consumers of that literature (such as designers), even though that aspect was not explicitly studied in this paper. The paper also contributes to the section on differences in SA due to context with data on how SA among pilots is experienced as ***situation dependent***.

Paper 3, contributes to the section on differences in SA due to roles in the design-use process with data on pilots' low compliance to procedures, possibly reflecting a gap between what the user is prescribed to do and what he/she is really doing. The paper presents results from a simulation where professional pilots engaged in a realistic commercial flight scenario did not comply with manufacturer, or air carrier procedures regarding mode monitoring and call-outs. Probably would a designer's understanding of a pilot's SA based only on formal documentation, such as procedures, be likely to differ from the actual SA of pilots. The paper also contributes to the section on differences in SA due to levels of analysis as an example of when data on a high level of analysis do not demonstrate the diversity that lower level of analysis does. At a high level of analysis the paper reports on normal conditions with low workload and no negative effect on the flight path or safety during the observations, giving no indication of varying SA. However, the more detailed level of analysis revealed a diversity of locally and personally tailored strategies that pilots develop to try to stay ahead of what the automation is doing, possibly indicating varying quantity or quality of SA.

Paper 4, contributes to the section on differences in SA due to context with data on the effect of task complexity and decision maker separation on shared SA in a command and control setting, from two experiments. The paper reports on one experiment in an environment for experimentation of command and control in network centric warfare, where three officers simultaneously where asked to follow a scenario and to prioritize how they believed they should act. Less shared SA, in terms of degree of agreement, were assessed during complex events compared to less complex events. The second experiment reported were performed during an air force exercise, where officials from six positions in the chain of command and control performed a number of air operations and were asked to rate own and others SA. Less shared SA, in terms of degree of agreement between decision makers who were physically separated than between co-located decision makers were reported.

Paper 5, contributes to the section on differences in SA due to levels of analysis, exemplifying with data from two experiments on when a high level of analysis do not demonstrate the diversity that lower level of analysis does. The paper reports on one experiment with commercial aviation pilots in a commercial aviation simulator where the pilots' looking behaviour were similar at one level of analysis and individually different at a lower level of analysis. At the higher level of analysis flight phases could be distinguished from each other by analysing the amount of time the pilot's gaze was spent head-up versus head-down, indicating regularities in the looking behaviour among the pilots, giving few indication of varying SA from irregularities in the looking behaviour. At the lower level of analysis, however, the pilots in this experiment reacted and performed with large individual differences, possibly indicating varying quantity or quality of SA, similar to what was reported in Paper 3. The second experiment was performed in a combat aircraft simulation with military pilots. Also in this experiment a higher level of analysis showed regularities when a lower level of analysis revealed individual differences, this time in the blinking behaviour.

Paper 6, contributes to the chapter on managing differences in SA in the development of complex systems with an analysis of how simulator-based design could help the human-machine interaction (HMI) design for in vehicle systems. By the use of simulator-based design the differences in SA could be detected and used for guiding the design in an iterative development process. The paper focuses human-in-the-loop simulation and provides two examples of simulator use for design issues. The first example is a study, not performed by the author of this dissertation, of two night vision systems, of which one was a novel situation-dependent display. The second example was a study of visual search and SA when using colour-coding as means for target discrimination in a flight simulator, also reported in Paper 1. SA, or the broader situation management, is proposed in the paper, as a useful concept to focus in simulator-based design.

2 Situational Awareness

In this chapter parts of the SA literature are reviewed. The chapter consists of three parts 1) *perception of the concept of SA*, including definitions of the concept, 2) *comprehension of the concept of SA*, reflecting on the essence of the concept, and 3) *projection of the future status for the concept of SA*, presenting some ideas on how the concept could be further enhanced.

The concept of SA has grown in popularity in aviation, and within other information-rich enterprises such as process control, intensive care, telemedicine, stock brokerage and traffic monitoring. This interest in SA is on both empirical and theoretical levels. Increased SA often is correlated with increased performance, which is an incentive for increasing SA.

2.1 Perception of the Concept of Situational Awareness

In the late 80's and early 90's there were a healthy and fruitful debate in the research community about SA, what it was, and how or if it should be defined. Even though not all questions that arose in that debate has been answered, more recent work has to a large extent focused on using the concept in various domains, and less focused on questions on definitions.

In the section below, some extractions of views from the research literature on the perception of the concept of SA are found.

First, just a short reflection on the use of *situation awareness* versus *situation<u>al</u> awareness*. When going through the research literature SA will be found spelled out both as situation awareness and situational awareness, which may be confusing. The most spread view is that the two terms are reflecting the same phenomenon and therefore are totally synonyms and only reflect two alternative ways of spelling. Often a researcher or even a research group tend to prefer one of the spellings and stick to that, as far as possible. For this dissertation I have made the choice to use situational awareness, since I have easier to associate that to the dynamic aspects of SA rather than a static situation.

Almost as confusing as an expression without any meaning, is an expression with a variety of meanings. This problem is solved if one clear definition is agreed upon, but other questions then arise. An example of this is: Can an agreed upon definition be valid in all contexts?

Sarter and Woods (1995), argue that:

> First, extended efforts to develop the ‘right’ definition or a consensus definition of situation (and mode) awareness will probably not be constructive. Rather, the term situation awareness should be viewed as just a label for a variety of cognitive processing activities that are critical to dynamic, event-driven, and multitask fields of practice. (p. 16)

Later in the same text: "Second, it appears to be futile to try to determine the most important contents of situation awareness, because the significance and meaning of any data are dependent on the context in which they appear" (p. 16). The point made by the authors being that finding a general definition for all contexts leads to a vague definition that will not fit the needs of each application area. Sarter and Woods (1991), point out: "any definition accounting for the context-sensitivity of data runs the risk of being too general to really help in understanding the basis of situation awareness" (p. 47).

One important criterion of a definition of SA is that it should support the search for suitable measurement techniques that are operable within the contexts to be measured. To achieve that, various factors ought to be considered. These can be regarded as belonging to two main criteria that relate to two parts of the enhancement process:

1) The first criterion being that a definition should be well founded.

Taylor and Selcon (1990), offer the following: "New solutions often require new paradigms. SA is needed for interface design to proceed beyond workload reduction towards providing solutions to mission problems" (p. 105). In this context, it is important not to let a definition of SA be composed of terms belonging to another paradigm, if the end result is a new paradigm with a wider meaning.

However, if the purpose is to create a term with a wider meaning within the current paradigm, the mentioned type of definition is suitable. Thus, the idea could exist either within or outside a current paradigm, but to be able to enrich the domain it has to fit its purposes, which could be the usefulness within a context of development of complex systems, that refers to the second main criteria (see below).

It is also desirable that a definition is based on both theoretical and practical knowledge. Practical knowledge may be used to counterbalance parts of the theoretical base not optimal for applied settings. Practical knowledge could also adjust a definition so that it is more easily accepted, thereby increasing the usefulness among practitioners. Practical knowledge, in this context, is knowledge acquired in field settings that can underlay and guide practitioner actions. Theoretical knowledge, on the other hand, relates more strictly to intellectual interests; to the *what* and *why* of a phenomenon.

2) The second criterion being that a definition should be useful for the applied context.

A definition of SA needs to foster consensus. To achieve this, information and data that supports the definition and describes typical settings where the concept of SA can improve safety and efficiency will contribute to common understanding of key concepts within the definition.

For the same reason, definitions and descriptions need to be clear and unequivocal. The potentially negative consequence of a concise definition, unlike a general and vague one, is the potential to discourage consensus. However, by being aware of this potential problem and seeking to overcome disagreement, the value of an agreed upon clear definition is greater than when using vague concepts and definitions.

Many attempts have been made to find a suitable definition of SA. A majority is suited to fit the purposes of the military sector, and as this dissertation has a general focus (both military and civil complex systems) they may not be exactly applicable. Some definitions focus on surroundings; such as Whitaker and Klein (1988), who propose that SA is "the pilot's knowledge about his surroundings in light of his mission's goals" (p. 321). Tolk and Keether (1982) define SA as "the ability to envision the current and future disposition of both Red and Blue aircraft and surface threats" (quoted from Fracker, 1988, p. 102). Taylor and Selcon (1991) state that:

> SA is a multi-dimensional construct. It can be modelled in terms of the processes of memory and attention. SA concerns the state of knowledge about the world, or the model of external reality that enables adaptive decisions to be made in uncertainty. (p. 789)

Van DerWeert (1986), states that SA is "knowledge of the current and near-term disposition of both friendly and enemy forces within a volume of space." (quoted from Endsley, 1988, p. 789).

Other definitions take into account more than surroundings. McKinnon (1986) regards knowledge of missing information when explaining SA as:

> The pilot's knowledge of:
> 1. Where both friendly and enemy aircraft are and what they are doing.
> 2. What the pilot's flight knows and what the flight's operations are for offense and defense.
> 3. What other flights know and what their intentions are.
>
> What information from above is missing. (quoted from Fracker, 1988, p. 102).

Phillips (1986) addresses the metacognitive aspects or *the knowing of knowing* with the definition: "knowing where the friendlies are and what they are doing, knowing what my flight knows and options for offense and defense, knowing what other flights know to a lesser extent, and knowing what I don't know directly".

Another description is from Harwood, Barnett, and Wickens (1988), when discussing who is in charge, the pilot or the automated system:

> Where: the pilot's knowledge of the spatial relationships among aircraft and other objects.
> What: the pilot's knowledge of the presence of threats and their objectives, and of his own aircraft system state variables.
> Who: the pilot's knowledge of who is in charge, him or an automated system.
> When: the pilot's knowledge of the evolution of events over time. (quoted from Fracker, 1988, p. 102).

According to Judge (1992):

> The Air Force's operational community has adopted the following definition of SA: A pilot's continuous perception of self and aircraft in relation to the dynamic environment of flight, threats, and mission and the ability to forecast, then execute tasks based on that perception. (p. 40)

The above definition is similar to the definition proposed by Endsley (1988), but the pilot's self perception is brought out; indicating that SA is not only related to the surroundings.

In some definitions and explanations of SA the time aspect is not addressed, while in others it is emphasised. Adams and Pew (1990), do not regard SA as a constant knowledge or ability, but point to the time dynamic aspects of the working knowledge of the pilot:

> We have made a rule of using the term situational awareness only with reference to the actual active or working knowledge of the pilot. Beyond respecting the fact that awareness belongs to people and not the environment, this usage was motivated by research-oriented considerations. Most critically, in stipulating that it is the working knowledge of the pilot that is of interest, we acknowledge that at any moment in time, a person's effective or active knowledge base consists of but a fraction of her or his total knowledge. (p. 520)

Some reference to a dynamically changing situation are also expressed by Sarter and Woods (1991): "situation awareness can be defined as all knowledge that is accessible and can be integrated into a coherent picture, when required, to assess and cope with a situation" (p. 55), and yet more by Siem (1992):

> Situational awareness (SA) is defined by Subject Matter Experts (SMEs) as "a pilot's continuous perception of self and aircraft in relation to the dynamic environment of flight and mission, and the ability to forecast and then execute tasks based upon that perception". (p. 41)

An alternative position is to focus on the effects of SA, thereby proposing an indirect meaning of the term. Regal, Rogers, and Boucek (1988), state that SA means that "the pilot has an integrated understanding of factors that will contribute to the safe flying of the aircraft under normal or non-normal conditions. The broader this knowledge, the greater the degree of situational awareness" (p. 1). Metalis (1993) presents another example by focusing the ability to multitask:

> A pilot who has SA is like an "expert" who can look at a huge array of discrete stimuli and immediately integrate them into "chunks", or meaningful bytes of knowledge upon which he can base appropriate action. An expert musician sees discrete notes and performs music. An expert chess master sees pieces on a board and perceives winning game strategy and tactics. An expert pilot sees the outside view and the cockpit instruments and perceives his human/aircraft system flying with respect to relevant others through space and time. But unlike the other experts, who may focus their attention on only one topic, the pilot must be able to multitask between several different subsystems, and he must do so not at his pace but within the time and priority constraints dictated by the flying environment. This is SA. (p. 115)

For expert pilots spatial long-term working memory (i.e. mechanism for extending working memory that requires skilled use of storage in and retrieval from long term memory) have been found to be a predictor for SA (Sohn & Doane, 2004).

An ecological view is advocated by Flach (1995b), who hesitates to attribute causal explanations to the concept of SA. As a contrast to the definitions focusing on the effects of SA are the definitions focusing on the essence of SA. As an example, Endsley (1990) explained SA by expressing that it "can generally be defined as the pilot's mental model of the world around him" (p. 41). Endsley's (1988), definition of SA seems to have attracted most adherents; "the perception of the elements in the environment within a volume of time and space, the comprehension of their meaning, and the projection of their status in the near future" (p. 792). Another example of an explanation of SA, focusing on the essence of SA is: "More generically, SA is the interaction of a person's knowledge and perception to form an understanding of the current situation" (Vidulich, Stratton, Crabtree, & Wilson, 1994, p. A7).

For practical purposes, the definition problem may be approached by letting methods and techniques that measure SA provide a foundation for a definition of SA. For instance, Shively and Goodman (1994): "While realising that other factors may contribute to SA, we will operationally define it here as system knowledge, as measured by the subtask metrics" (p. 1271). Taylor and Selcon (1990), argue that: "The approach to SA measurement investigated at the IAM involves subjective rating of bipolar SA construct dimensions elicited from aircrew using the Personal Construct Theory/Repertory Grid Technique. Using this approach no a priori SA definition is necessary" (p. 108).

A task-oriented perspective, with the ambition of being consistent with Endsley's approach, have been suggested by Patrick and James (2004) where "SA is the product of performing satisfactory the task of 'achieving and maintaining SA'" (p. 63). SA then is the product of one task among several that a user has to perform. SA in itself may be regarded as including several aspects. A pilot, for instance, has to be aware of many things simultaneously. Nagel (1988, p. 215), expressed this as:

> The pilot must maintain awareness, for example, of where the aircraft is geographically, what its current performance state is, its altitude, power setting, configuration, and relationship to other aircraft. The pilot must (1) be mindful of verbal communications and instructions from air traffic control personnel, (2) be aware of the environment, and (3) systematically monitor for and anticipate changes in weather that might

affect his or her performance and that of the aircraft. Information to maintain an appropriate awareness of the flight situation is obtained visually, through the windows of the aircraft and from instruments, aurally from other crew members within the cockpit and from air traffic controllers over radio channels, and through the vestibular senses. (quoted from Adams & Pew, 1990, pp. 519-520).

To be able to evaluate a definition of SA we need to comprehend what SA really is.

2.2 Comprehension of the Concept of Situational Awareness

To understand what SA really is we need to realise that SA, as a concept, could be a part of a linear as well as a cyclic model. Also, we need to understand what SA consists of in terms of SA components or SA levels.

Accordingly, it is important to understand the relationship between SA and metacognition, SA and teams as well as SA and technology, for instance, in terms of Joint Cognitive System – JCS (Hollnagel & Woods, 1983; 2005; Woods & Hollnagel, 2006), to value the use of SA in the context of development of complex systems. These aspects will be further explained in this section.

Homunculus is Latin for little human. The homunculus refers to the postulated central control mechanism of the brain. The homunculus problem appears when looking at all possible ways of trying to explain cognition, both when trying to explain how people think, within the field of cognitive psychology, and when trying to create new forms of cognition, within the field of artificial intelligence. More precisely, the homunculus problem appears when trying to explain cognition through a linear model, a model that begins in the objective external world and ends in the internal subjective awareness of the individual.

A linear model has a beginning and an end, just as a line has a beginning and an end. In the beginning of the line there is an event that can be detected, through its causal interaction with the physical environment. The information about the event then travels through the line when being detected, and transformed by the human information processing system. In the end of the line the information just drops out in the blue, without being used for creating meaning through conscious awareness, if there is nothing in the end of the line that can catch the information before it disappears.

By using memory as an explanation the line could be extended and the information could be delayed. This is not solving the problem, just pushing the problem in front of you. A tempting solution is to put a homunculus in the end of the line, taking care of all the unsolved problems. This, however, does not solve the fundamental question about how human cognition works, because the question about how the homunculus cognition works appears instead.

An alternative model to the linear model is a cyclic model, where the information-line is replaced by an information-circle, as in Neisser's perceptual cycle (Neisser, 1976), (see Figure 1). Also, Fuster (1989, 1995) supports the idea of a perception-action cycle, based on knowledge of the human neuroanatomy.

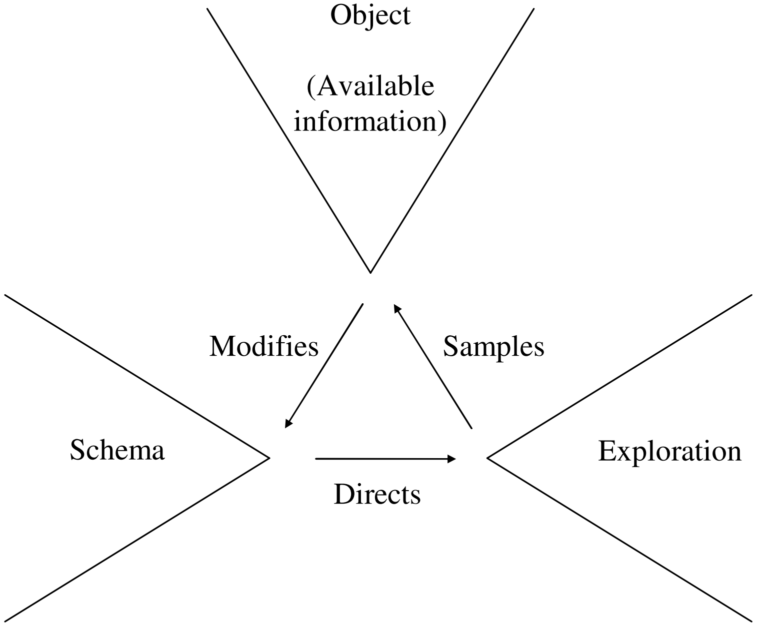

Figure 1. Free from Neisser (1976): The perceptual cycle.

Neisser (1976) explains how perception is a constructive process, but that there is no "inner man" to view mental images of the consciousness. Since a circle does not have an end, like a line has, there is no need for a homunculus taking care of information dropping out of the model in the circular model, as there is in the linear model.

This, however, does not mean that a circular model is better than a linear model. In a circular model the problem is not that the information leaves the model unexplained, but rather that the information stays in the model unexplained!

Without regarding any regulation of the information flow in the information-circle all explanations tend to be extremely context dependent, and in its extreme fatalistic. Things happen because other things happen, involving the whole world and all its history, in the explanation of a simple event. This is the other side of the coin, where we begin to miss our cute little homunculus to be there, whenever we reach the limit of our understanding of cognition.

Since a circle does not have a defined beginning or end, it is also extremely hard to theoretically explain parts of an interesting, isolated, situation, with a cyclic model since everything is part of the same reality. Wherever you chose to put an end to the causal chain, and say that you are satisfied with the information-complexity of the circle, you will find a good friend of our homunculus looking up of the edge of the circle, tempting you to expand the circle even further.

The environment for any one JCS generally contains another JCS that can impose constraints both onto a JCS at the same level and onto a JCS within it, and the constraints shape the actions of a JCS (Woltjer, 2005). For instance, the system boundaries for free flight, may be set to include only the crew, the flight, local airspace, or even traffic flow depending what is interesting to analyse (Hollnagel, 2007). However, it is not easy to set the system boundaries before making the analysis, since the analysis itself may provide information useful for making that choice.

Both types of models have been proved practically useful. Who am I to judge which model is the theoretically correct, when I do not have a total understanding of how cognition works? People have been arguing about linear and circular models, for a long time, without coming to a clear conclusion about the problem. In practise, however, many models are complementary to better understand human behaviour. For instance, a cyclical model of system cognition has been found to describe and explain the brittleness of the introduction of the novel technologies, area navigation (RNAV) and datalink, in a commercial aviation context (Goteman, 2006).

The linear model has been used to explain parts of the human information processing, and thereby reducing the homunculus so that it becomes even smaller, without aiming at taking it away completely. The understandings of parts of the human information processing can very well be used for practical improvements without having a complete understanding of the problem.

For instance, Wickens (1992) explains the searchlight metaphor of attention by the words:

> Attention in perception may be described by the metaphor of a searchlight (Wachtel, 1967). The momentary direction of our attention is a searchlight. Its focus falls on that which is in momentary consciousness. Everything within the beam of light is processed whether wanted (successful focusing) or not (failure to focus). Within the context of this metaphor, two properties of the searchlight are relevant to phenomena of human experience: (1) the "breadth" of the beam and the distinction, if any, between the desired focus, or penumbra (that which we want to process), and the umbra (that which we must process but do not want to), representing the issues of divided and focused attention, respectively; (2) the direction of the searchlight - how it knows when, what, and where in the environment to illuminate - describing the properties of selective attention. (p. 75)

As explained above, the searchlight metaphor does not solve the homunculus problem. It only describes a part of the information flow by using the metaphor. There are lots of hazards in doing so. The homunculus still remains and it is important to be aware of that. It is easy to fall for the temptation of looking at the searchlight metaphor as an explanation of attention, which it is not. It is just a metaphor to help understanding the problem well enough to be able to do something constructive about it.

The same goes for any other linear model within the field of human information processing. For example, as mentioned above, the linear models of how human memory functions, just delays the information, but it still ends at the homunculus. The same is true for the concept of SA, which easily ends up being nothing but a homunculus behind the fancy facade.

Paper 2 indicates that subjective components are essential parts of the notion of SA expressed by pilots. Annett (2002) proposes that the classical objective/subjective dichotomy may be too absolute and that we instead should consider human judgements as varying in terms of degree of intersubjectivity, that is, shared meaning between observers of an event. Wallenius (2005) is referring to the Endsley model of SA (Endsley 1988; 1995b), when suggesting that "presumably the objective part is more prevalent in SA Level 1, dealing with physical elements in the environment, whereas the higher levels of SA, comprehending their meaning, should be more subjective" (p. 45).

A combination of subjective and objective components in the formation of mental concepts like SA is appealing from both a theoretical and practical point of view.

In the following section some reflections are presented on what such components of the SA concept could be, partially rephrased from Alfredson et al. (1996). SA is a concept that has been proved to be useful first, within the aviation domain, and later in a wide area of domains. The wide recognition and frequent use of the concept indicates that it fills a need. SA has been defined and re-defined several times. It is difficult to clarify statements about high or low SA without explicit reference points. Also, the SA concept could be positioned to other concepts to help understand the concept.

First, the issue of SA is related to decision making in dynamic systems, where speed and accuracy of operator response is critical. As automation and task complexity increase, an operator is at greater risk of becoming lost in a system. This tendency is especially common in multimodal systems, where a specific display unit can, at different times, represent quite different physical states of the world. Accordingly, system failures due to mode errors have become more common. The SA concept has proved to be a fruitful framework for categorising operator errors, for instance, pilot errors in civil aviation (Endsley, 1995b, Jones & Endsley, 1996), or errors associated with offshore drill crews (Sneddon, Mearns, & Flin, 2006).

Second, SA is closely related to established concepts in the information-processing paradigm, predominately in human factors research. SA is often regarded as being principally in the cognitive domain (Hartman & Secrist, 1991). One SA model in mainstream HF research that includes information processing comprises three levels (Endsley, 1988; 1995b):

- Level 1, perception of task relevant elements. For instance, a pilot has to perceive task elements such as other aircraft, display meters, enunciators, and radio messages.
- Level 2, interpretation of task elements. On this level, a synthesis is made that go beyond the characteristics of the perceived elements. Information is interpreted in terms of its relation to pertinent goals. A fighter pilot, for instance, should be able to assess the intentions of enemy aircraft on the basis of their number, distance, speed and formation. Equally true is that a civilian pilot must assess intent of other aircraft such as in a busy air corridor.
- Level 3, projection of future states. This represents the highest level of SA, where the operator predicts the unfolding of events which, in turn,

provide a basis for decision making. A fighter pilot realising an enemy attack predicts its speed and direction and then chooses the optimal alternative - counter attack, evasion action or retreat. Heavy air traffic over a civilian airport has comparable prediction requirements.

Each of these levels contains identifiable cognitive processes and attendant performance deficits. Lack of SA at level 1 may be caused by a range of factors that include vigilance decrements, discrimination failures and non-optimal sampling strategies in supervision. Errors at level 2 are related to mismatches between system characteristics and an operator's mental model. Level 3 errors may occur in spite of accurate perception and interpretation of task relevant information, as projection of future states is an extremely complex cognitive task. Taxing on working memory and attentional resources, mental simulations require good mental models. Focusing on errors, however, is not necessarily fruitful for reaching higher level goals, such as increased safety, since it is hard to determine what an error is, how to classify it, how to add its effects to other errors, and what the effects are to safety (Dekker, 2003; 2005; 2007).

Normally we know in behavioural sciences what to measure and how to measure it, since we usually are interested in the assessment of presence of something. In the case of SA, focus is on assessment of absence of it, which is a paradox that complicates measurement. In general, it is an easier task to measure the occurrence or prevalence of something. In measuring the lack of something, the expected outcome might be a result of inappropriate methods, unreliable instruments, invalid procedures, and so forth, as well as a true omission/lack of data. Literature seldom is concerned with perceiving and discussing SA in normal (colloquial) situations. Literature is usually concerned about situations where, for instance, a pilot is lacking SA. To appropriately measure something that is missing it is necessary to identify what should have been present. Normative models of (high) SA are seldom at hand in a particular context. Often a lack of SA is related to a subjective feeling of the constituents of high SA.

The increased interest in SA might be explained both in empirical and theoretical terms. Some empirical circumstances contributing to increased interest are incidents and severe accidents that might be attributable to user faults. Intuitively, we realise that the risk that users become lost in systems increases with the higher degree of automation present in dynamic systems.

Not only the system that the user is attempting to control may change rapidly, but also the representation of the system may change frequently, along with the means for monitoring and steering. With today's multimode systems, information and the means of presentation are further complicated. Along with more sophisticated technology interfaces become more complex. Increased complexity is usually contradictory to user friendliness.

Theoretically, a comprehensive SA model integrates a number of specific cognitive processes, which have been identified as relevant to the tasks to be fulfilled by a user. In this context, user thinking represents an advanced form of information processing. A minimum of five components making up SA were identified in a review of literature, depending of the resolution that is useful for a deeper understanding of SA. These components are descriptions that can be used to understand the concept of SA, by linking it to other theories of human behaviour. The components, however, are not discrete parts of SA in the sense that they divide SA into sub-SA terms to be measured instead of measuring SA. The rationale behind the use of SA is that the phenomenon is not covered by other terms and is to be viewed as a whole. Further, the components cannot be interpreted as a linear information flow. As Neisser (1976) described the perceptual cycle, human perception is far from being linear. Neisser's cycle has been used to spread some light on the dim theoretical surroundings of SA by, for instance, Smith and Hancock (1995) as well as Adams, Tenney, and Pew (1995). Each component is integrated with a variety of feedback loops. At a specific point in time, SA is highly dependent on previously experienced states of SA, reflecting only a snapshot of the ongoing dynamics. Below, five components are identified:

- Attention (Vigilance, decoding, etc.)
  The first component concerns attentional resources that are allocated to the processing of relevant task specific information. Vigilance (user's ability to detect gradual or discrete changes of stimuli), may be necessary for survival in some scenarios. Vigilance is to a large extent determined by physiological prerequisites and less prone to training. Decoding of data depends on both modality specific attributes and prior knowledge. A lack of SA at this level might be about simple misreading of instruments or digital data, weak signals/stimuli that are not detected, non-optimal mental resource allocation, and so forth.
- Perception (Active search, identification, etc.)
  The second component embraces perceptual processes, active search processes, which result in identification of information/objects. Users are sometimes exposed to such a high number of stressors that their perceptual abilities are impeded. A lack of SA on this level implies indulgence, fatigue or stress reactions (for instance, distortion of normal search patterns); deficient selective processes leading to "freezing behaviour" or "confusion". In extreme environments, for instance, in a

fighter aircraft the pilots are sometimes exposed to high G forces, high level of stress and a high speed. The result could be, for instance, a loss of peripheral perception of information or a loss of or detrimental colour perception.

- Memory (Encoding, storage, retrieval, meta-information capabilities, etc.)
  The third component is often divided into encoding, storage and retrieval of information. The encoding process generally requires rehearsal and elaboration of information to be efficient. Short-term storage of information is limited and easily lost among surrounding noise or information. Working memory capacity has been found to be a predictor for flight SA (Sohn & Doane, 2004). Long term storage is more robust, permanent and, in some sense, unlimited. Still the retrieval of information can be the crucial bottleneck in remembering relevant information. Learning sufficient retrieval cues during training is important. Providing users with realistic contextual cues influence retrieval. Another pertinent memory quality may be individual differences in meta-information capabilities. This may be one of the most neglected determinants of SA. A lack of SA at this level might be about a too high mental workload; preventing efficient encoding or elaboration of information. Accordingly, users may lose important information, lose temporal or spatial cues relevant to their tasks, get lost in a multimode environment, and so forth. Obsolete metacognitive structures may not be applicable in a new context, resulting in, for instance, misjudgement of their own ability to solve upcoming problems.
- Interpretation (Mental model, meaning, inference, decision, etc.)
  The fourth component is about interpretation of the attended, perceived and memorised information. Usually the information has meaning defined by task or mission. While primary components, in an aviation context, take care of the detection and identification of aircraft in the vicinity or enemy aircraft, this component could be about interpretation of intentions of surrounding pilots based on behaviour and manoeuvres. Drawing upon previous experiences, pilots make inferences and decisions about current events. Faults at this level may stem from premature or undeveloped mental models of others' behaviour, inappropriate prerequisites for interpretations in a particular context, or wrong decision rules. Humans are also subjected to a wide range of biases in their judgements.
- Prediction (Mental simulation, forecast, prophecy, etc.)
  The fifth component concerns prediction of future events as a ground for decisions. This is a demanding cognitive task that requires a superior ability of mental simulation. Using available information, a pilot must weigh different information and make a short term forecast. Experience increases users' ability to make long term projections about future events

that they may encounter. A lack of SA at this level may have detrimental effects on planning and long term survival. Concrete behaviour might be misdirected and even jeopardising lives.

Failures in these components may affect SA negatively. Endsley, Bolté, and Jones, (2003) described factors that work to undermine SA as SA demons, and exemplified with: Attentional tunneling, Requisite memory trap, WAFOS (Workload, Anxiety, Fatigue, and Other Stressors), Data overload, Misplaced salience, Complexity creep, Errant mental models, and Out-of-the-loop syndrome.

Good SA is not necessarily an ability that the user possesses in the same sense that he/she possesses good calculus ability or a good pitch perception. SA consists of a conglomerate of human abilities and interaction with task, colleagues, organisation, environment and technique. A user's task can easily be changed according to the current situation. Hence, it is possible that a user is in full control one moment with a high level of SA, and in an adjacent sequence of actions is out of control with a drastically reduced SA. This phenomenon reflects the rapid dynamic changes that can occur in a hyper dynamic environment, such as an aircraft.

Certainly perception of data, understanding of the meaning of the data, and the projection of future events, are all important aspects of SA (Endlsey, 1988). However, it is not quite as certain that the levels are really levels, that the model covers all aspects of SA, or that the model explains the essence of SA, related to a specific context.

Dividing SA into levels imply that the situation assessment is a linear process, starting with a raw percept and ending in high level SA, without concern how this process is affected by the current mental state, not even reflecting on how the SA itself affects this process. As Neisser (1976) explains the human perception could also be seen as a perceptual cycle. Human information processing is not necessarily linear. The first level of SA is dependent on the SA in both the other levels. By understanding what has been perceived, the vigilance could be affected, and thereby affecting the situation assessment at level 1 SA. Being able to predict future events provides changes in attention that also affects the situation assessment of level 1 SA. Both the level 2 SA, and the level 3 SA affects the level 1 SA. The situation assessment process has several feedback loops, making it far from linear. The SA at one point is very much dependent on the SA in previous points of time. A model of SA that is not emphasising the dynamics of SA leaves out an important characteristic of the concept.

Another question that has to be answered, before being convinced that the three levels are a correct way of describing SA is: Can SA be divided into parts (levels), be assessed, and then be fused together again to form an understanding of SA, without losing important information when dividing and summing up SA?

It is possible that SA is something more than just the sum of the three levels. If it is not, why use the overall concept of SA in the first place? An obvious risk that appears when dividing SA into three levels is that the real magic is made between the levels and not at the levels itself. When fusing scores from three levels into one overhead SA score a model of how the levels add up to SA will be used, explicitly or implicitly.

Even though the three levels of SA are important aspects of SA, it is not very well founded to postulate these three aspects, as the only aspects of SA. On the contrary, there are at least four good reasons for reflecting on other aspects that are of importance for SA:

- The intuitively feeling that SA is dependent on intelligent minds like the human mind to exist, and related to the conscious awareness of the mind.
- That measures of SA are used to make evaluations of human-machine interface designs.
- That focusing on the three levels simultaneously might be difficult.
- Oscillating SA might not always reflect dynamic changes of the situation.

One of the reasons why the concept of SA has been able to win ground, although missing a well based theoretical foundation, is that it seems that almost everybody has something to relate to as SA. Even though there are a variety of definitions of SA, and the theoretical problems of SA remains, most people (and even researchers) tend to agree that they have something to relate to as SA in their experience.

It is far from easy to reach the same consensus on that the mental concept of SA is applicable on machines. Machines are without the intelligence of the human mind, and without conscious awareness. It could be taken as an incentive to improve a SA model if it is applicable to a machine and thereby might be missing what makes SA characteristic to humans.

In a study by Mulgund, Rinkus, Illgen, Zacharias, and Friskie (1997) the three levels in Endsley's model of SA have been compared to a model of data fusion (DF) processes:

> The Joint Directors of Laboratories (JDL) Data Fusion Subpanel have identified three levels of fusion processing products (Waltz & Llinas, 1990; White & Cohen, 1980):
>
> Level 1 DF: Fused position and identity estimates
> Level 2 DF: Friendly or hostile military situation assessments
> Level 3 DF: Hostile force threat assessments. (p. 2)

The study points out the similarities and even fuses the two models into one four-level model consisting of the first level of DF followed by Endsley's three levels. The model was then the base of a prototype demonstrating how an on-line SA processor could be used (Mulgund et al., 1997). Demonstrating that this is possible, also demonstrates that the SA model used are very mechanistic in its structure. Maybe it is less surprising that a SA model that are influenced by sequential mechanical processes, are used to explain the SA of computers, than it is that such a model is used to understand human cognition.

By applying Endsley's SA levels it is possible to assess a machine's SA, comparing internal representation in a computer with the external truth. The SA of one computerised pilot model has been measured by the use of a reference computerised pilot model (Zacharias, Miao, Illgen, Yara, & Siouris, 1996). The basis of the comparison is Endsley's SA model. As Endsley suggests that SA can be measured by objective questions, for instance, using SAGAT (Situation Awareness Global Assessment Technique) (Endsley, 1988; 1995a; 2000; Jones & Kaber, 2005), testing SA at level 1-3 by comparing a statement with an objective truth, this could be applied to computers as well as humans, although the human subjectivity makes it hard to argue that these kind of judgements are really objective (e.g. Annett, 2002; Walker, Stanton, & Young, 2006).

One of the ideas behind the study by Zacharias et al. (1996), was that one computerised pilot model was not provided with all the information about the external situation, while another pilot model was. The SA of the computerised pilot model could thereby be measured by comparing the objective truth in one of the pilot models, with the incomplete representation in the measured pilot model.

Applying the same SA model on humans, as on groups of humans and on machines, implies that the SA of a person is to be comparable to the SA of a group or a machine. By doing so, the risk is that some aspects of human cognition that are important to grasp the situation are neglected. Also the view proposed in this dissertation could be extended to incorporate machines or groups of agents. However, it is important to regard the importance of the fundamental difference between the human conscious awareness, and a non-conscious awareness.

It might be difficult to focus on the three levels simultaneously because of practical reasons. Which data should be considered to belong to which level? How should the information about the three levels be fused and weighted? There are also theoretical reasons why it might be difficult to focus on the three levels simultaneously. For example, are the levels independent of each other, or how are they related? The dynamic changes of the situation change quickly due to the complex interaction of the elements in the situation, for instance, a pilot and the surrounding context. Is it possible to correctly measure SA without regarding the interaction, by using predefined questions? Is it possible to achieve such a high sampling frequency that the dynamic changes of the situation can be extracted through the data from the points in time when SA is assessed? One of the reasons for creating a good design of a complex system is to enhance the performance of the total system of user and machines, by enhancing the decision process. The use of the SA concept might very well help achieve just that.

However, it is not easy, and might not even be possible, to extract some relevant characteristics of the environment to represent the objectively true SA. The interaction between the subjective impressions of the situation in relation to the dynamics of the context, affects (if not even determines) the outcome of the decisions in each situation. If decision making is affected by context it is important that these aspects are regarded when designing an environment where important decisions are likely to be made.

Below, the author's view on SA (Alfredson, 2001a; 2001b) is presented to help understand the concept. SA is a complex construct with many various requirements associated with, for instance, metacognitive aspects and collaboration aspects (Paper 2). SA may be regarded as dependent on a number of basic components. These components could be categorised into three groups of awareness areas:

- Awareness of intra-human information
- Awareness of inter-human information
- Awareness of non-human information

The intra-human component includes both cognition and metacognition. The inter-human component includes awareness of team communication direct or through communication technology. The non-human component includes information from surrounding objects and artefacts, for instance, instrument information and awareness of the information in an outside view, and so forth. hese components do not remain at the same ratios of importance for SA. As context changes the proportions of the components of SA change as well, (see Figure 2). For a pilot, for instance, awareness of radio communication may be of great importance when being informed on changing plans and metacognition may be of great importance in another situation, such as when deciding whether to delegate a task or not.

However, SA is more than just a construction of its components. It is a term with a specific connotation. Just like a flower in some sense is a construction of water, sunlight, chlorophyll, and so forth, a flower yet also is something more. It is something of its own. Like a flower is dependent of all its components to thrive, SA is very sensitive to lack of its basic components. The flower of SA is continually growing and changing so that at one point of time it is not the same as it was just a moment ago. The changes may be fast, and this makes it difficult to describe, but not impossible, as purpose remains constant.

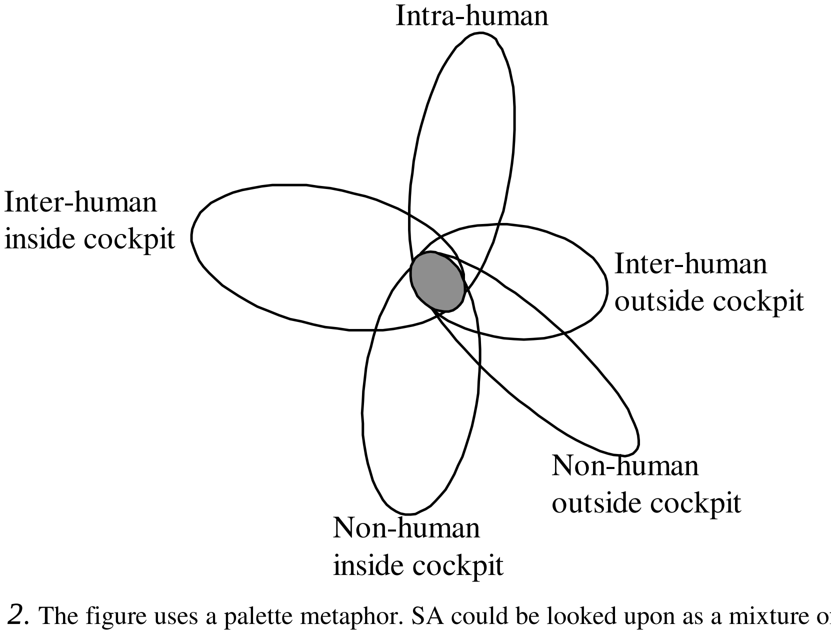

Figure 2. The figure uses a palette metaphor. SA could be looked upon as a mixture of paint on a palette. The components of SA are represented by paint of various colours. As the context is changing, the proportions of the components of SA change and the colour of SA thus also changes.

One way to describe this flower is to compare it with its relatives, like mental workload or attention. Svensson, Angelborg-Thanderz, and Wilson (1999) as well as Svensson and Wilson (2002) have described the relationship between the concepts, performance, SA, difficulty, mental effort, pilot mental workload, mental capacity reduction, motivation, information complexity on displays, and earlier similar analysis have been performed with concepts related to mental workload (Svensson, Angelborg-Thanderz, Olsson, & Sjöberg, 1992).

The relationship between SA and mental workload has been studied and researchers have found that both concepts should be considered for system evaluation (Endsley, 1993; Vidulich, 2000). Also, Wickens (2002) explains how mental workload and SA is intertwined and that, for instance, task management is related to both SA, through task awareness, and to mental workload, since task management uses the limited mental resources of the user.

Jones and Endsley (2001) studied the relationship between schema and SA. SA has even been tried as a measure for quantifying telepresence (Riley, Kaber, & Draper, 2004). Artman (1999) compared SA and coordination for command and control teams and found that the concepts, although interdependent, are to be regarded as different concepts having different meanings.

A strong rationale for using the theoretical concept of SA, is its positive practical implications, such as how useful it is in the process of developing complex systems. Therefore it is also important to regard practical aspects such as take into consideration what is possible to measure, and not be overly abstract.

The definition suggested below therefore focuses on awareness and what is related to that, rather than a purely theoretical abstraction of SA. To illustrate SA and to help the reader with a visual representation, Venn diagrams will be presented in a series of figures. This type of Venn diagrams have been used before to visualise SA (Alfredson, et al., 1996; Alfredson, 2001; Dekker & Lützhöft, 2004). The metacognitive component is stressed in the proposed definition. Considering the criteria of the definition, the following definition is proposed (see also, Alfredson, 2001a; 2001b):

High SA is when someone is fully aware of a high relative awareness

Relative awareness

If someone is not aware of a situation he/she does not have SA, because of an absence of knowing what is happening. Relative awareness is a combination of actual awareness and required awareness (see Figure 3) and thus takes into account that SA relate to performance. Good performance is dependent upon personal awareness of the elements that require awareness. If actual awareness is not as high as required awareness there is a lack of SA (see Figure 4, 5 and 6).

Awareness of relative awareness

If someone is not aware of that he/she has high relative awareness of a situation he/she does not have SA, because, with a bit of bad luck, he/she might just as easily has lost it without knowing. This part of the definition addresses how SA relates to metacognition. If experienced relative awareness is not the same as actual relative awareness there is a lack of SA. For instance, if a subject thinks that his/her relative awareness is high, when it actually is low, or he/she thinks that it is low when it is high then the level of awareness of relative awareness is low (see Table 1).

Table 1 illustrates the relationship between *awareness of relative awareness* (metacognitive ability) and *relative awareness*, by showing the possible outcome of SA depending on four possible combinations of relative awareness and awareness of relative awareness. A high relative awareness accompanying a high awareness of relative awareness describes an ideal scenario with high SA. A high awareness of relative awareness, when relative awareness is low implies low SA. A risk for underestimation of a person's SA is present with low awareness of relative awareness together with high relative awareness. Perhaps the most devastating scenario is when low awareness of relative awareness is combined with low relative awareness.

It is not far fetched that a person that is not aware of the required information in a situation might not either be the best assessor at the meta-level of assessing that there are incorrectness's in that respect. Walker, Stanton, and Young (2006) found, in a study on car drivers, that drivers demonstrated little self-awareness of diminished SA.

Or, as expressed by McGuinness (2004, p. 5): "In terms of situational awareness, a well-calibrated individual is one who has a high level of actual SA *and* correctly perceives this to be the case in his or her perceived SA."

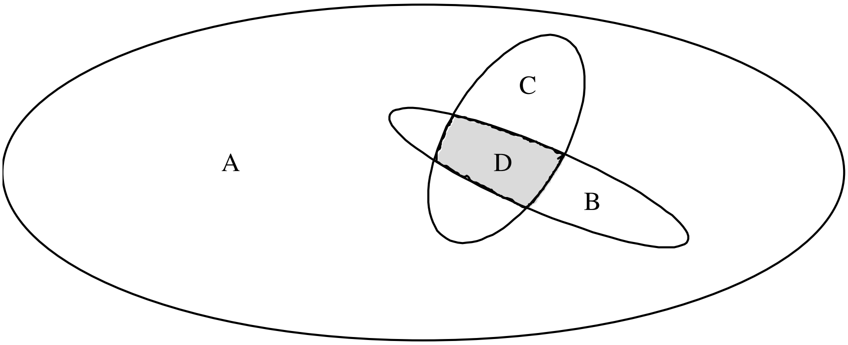

Figure 3. The figure illustrates the relationships between different information areas in a dynamic situation. Each information area contains information about current state, past states, future states and relationships between states of a dynamic situation.
A = *Situation*
This area symbolises all information that concerns a dynamic situation.
B = *Awareness*
This area symbolises information about a dynamic situation of which a person is aware.
C = *Required awareness*
This area symbolises information about a dynamic situation that is required to be fully aware of what is relevant in a situation.
D = *Actual awareness*
This area symbolises information about a dynamic situation that is both in a person's awareness and also required for the person to be aware of what is relevant in a situation.
D/C = *Relative awareness*
The relative awareness is not an information area, but a quotient of two areas; the area of actual awareness and the area of required awareness.

Table 1

The Relationship Between "Awareness of Relative Awareness" and "Relative Awareness"

		Awareness of relative awareness	
		Low	High
Relative awareness	High	- (Underestimation)	+
	Low	-- (Overestimation)	-

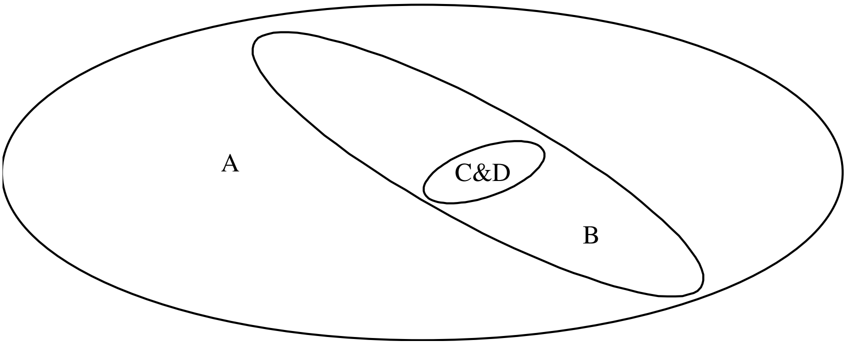

Figure 4. High SA. When a person has high SA the areas C and D become the same area, since the person is aware of all that is important in the situation. B is larger than C and incorporates C.

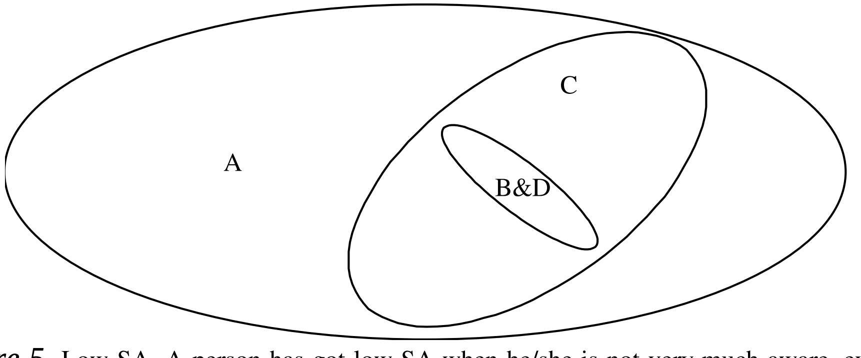

Figure 5. Low SA. A person has got low SA when he/she is not very much aware, even if he/she is aware of the right things. B is smaller than C. This could also be the case when the situation is hard to grasp, because of high demands, that is, C is too large to be aware of. Now B and D becomes the same area in the diagram.

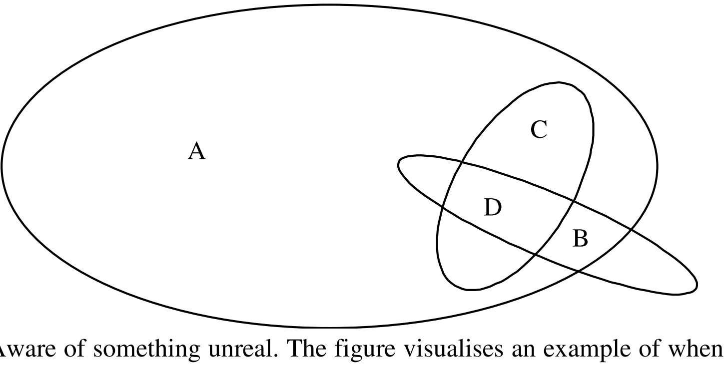

Figure 6. Aware of something unreal. The figure visualises an example of when a person is aware of something that is not a part of the situation at hand. This could be the case when someone is misperceiving, remember the wrong thing, or even hallucinating. Another possibility for this configuration to occur is that the situation A has become smaller. Since the situation is dynamically changing, it cannot only undergo qualitative changes, but also quantitative changes.

However, a task may be well performed even if the individual cannot express how in it is done. Rasmussen (1983, 1986) distinguishes between skill based, rule based and knowledge based behaviour. A very familiar task that has become a skill is often performed without conscious reflection.

The awareness of relative awareness is a metacognitive aspect of SA. Flavell (1976) expresses metacognition as "one's knowledge concerning one's own cognitive processes and products or anything related to them" (p. 232). Metacognition refers to one's knowledge and active regulation of one's own cognitive processes and products, or, in less technical terms; thinking about thinking (Brown, 1987; Flavell, 1976).

Metacognition is an important aspect of cognition, in understanding how people form their understanding of the world, or fails to form a correct understanding, which we are still learning to understand (Chambres, Izaute, & Marescaux, 2002). Perkins (1989) highlights the "metacognitive shortfall", in his theory of shortcomings in everyday reasoning, and suggests various contributing factors: People may not be aware of how badly biased their situation model is; They may not be aware of how easy it is to extend a situation model. They may not reconsider a conclusion after obtaining new information about the situation.

Experts however, are often regarded as superior decision-makers in the literature, partly because they possess better metacognitive skills. For instance, as expressed by Proctor and Dutta (1995) "One attribute of skilled problem solvers is that they are better able to monitor their own cognitive processing" (p. 216), or Means, Salas, Crandall, and Jacobs (1993) "*Metacognitive skills* may well constitute the best candidates for generalizable skills that will aid decision making across domains" (p. 324) and later in the same text "Better performers employ metacognitive strategies, and teaching the strategies to poor performers leads to improvement on the task" (p. 324).

It is not hard to find examples of situations where a user's understanding of the situation is totally dependent on his/her metacognition. For instance, if a user is to decide whether to engage a certain automatic system or not, may depend on how the user judges his/her own current ability to manually perform the corresponding task, considering his/her own understanding of the situation and the own workload, and relating the own abilities to the automations abilities.

Since modern systems continuously are becoming more complex and more highly automated, this problem increases in relevance. Interaction between human and the technical system is not only about interaction on the surface, but about a deep interaction between the agents. That is, interaction is not only about how to perceive elements and understand what they are going to do, but also being aware of the level of awareness that the agents possesses respectively, including yourself.

Other situations where the user has to use metacognition, is in collaboration and communication with other persons. It is not only machine agents that the user has to relate his/her own understanding of the situation to. In fact the user continuously has to assess his/her own mental state, as part of the overall situation assessment.

Since metacognition is important to form an understanding of what is going on in the world, it is easy to see its relevance for building SA. McGuinness (1995) adds two components to the three levels of SA that were proposed by Endsley (1988). The components are *metacognition* and *response selection*. McGuinness (1995) points to metacognition as relevant for both *SA knowledge contents* and *SA-related processes*.

By comparing assessments on SA obtained by SAGAT with confidence ratings about the response, the metacognitive aspect of how well calibrated a subject is in his/her SA assessment, has been investigated in military command and control (Lichacz & Farrell, 2005) and air traffic control (Lichacz, Cain, & Patel, 2003).

When Stanton et al. (2006) proposed a theory for distributed SA (DSA) the system level was described as a meta level for individual SA: ”Moreover, there are then two aspects of SA at any given node: individual SA of one’s own task, and a ‘meta-SA’ of the whole system’s DSA” (p. 1308).

Disregarding the metacognitive aspects of SA, will lead to less valuable interpretations of data collected as measure of SA. Often data from various kinds of SA measures differs in their interpretation. If the data itself is valid, the differences in interpretation might be due to weaknesses in the model of SA. For instance, it is possible that some of the differences between pilots’ rating of own SA and observers ratings of pilot’s SA could be explained in terms of metacognition. Berggren (1999) has debated such an approach, in a civil aviation simulator context. Walker, Stanton, and Young (2006) reported that car drivers’ self-awareness of SA was very low in a study on vehicle feedback. In a study of experienced road cyclists Knez and Ham (2006) found that fatigued participants underestimated their ability to maintain SA.

So far, this chapter has been focused on what SA is but is SA a useful concept when developing a complex system, if used in the right way? It is a question of finding ways of using the concept of SA when it is beneficial to the overall goals, and to complement it when necessary, for instance, by asking questions such as ”tell me what is going on in this situation now” and ”tell me how the current situation differs from the preceding one” (Cooke, 1997). By detecting the cognitive task requirements, these might be used as the driving factors in a design effort, enabling system developers to achieve a much stronger effectiveness at a relatively low cost (Klinger & Gomes, 1993), and identifying specific SA requirements has been pointed out as an important step in designing technological systems that optimize work performance (Endsley, Bolstad, Jones, & Riley, 2003; Endsley, Bolté, & Jones, 2003; Matthews, Strater, & Endsley, 2004).

Also the selection of assessment technique, affects the usefulness of the SA concept in the development process. When the simulation is frozen in the SAGAT method (Endsley, 1988) it is not only the simulator that freezes, but also the evolving understanding of the concept of SA. The constructor of the questions cements his/her own current notion of SA, manifested in the selection of and formulation of the questions.

Trying to receive information about the situation directly from the subject instead of trying to understand the situation via answers to objective questions about the situation, is a more direct way of extracting information about the situation. By the more direct approach the information is less likely to be biased by the interpretations of the researcher, when trying to interpret the answers. On the other hand is it a higher risk that the subject biases the information if the subject is to assess the situation all by himself/herself. It is also difficult to detect when that is the case.

By using both approaches simultaneously they might complement each other, trying to use elicitation of SA in addition to assessment of SA. Neglecting the elicitation part might lead to less understanding about the situation, since if the subject has reached some level of SA, the contents, and not only the amount, of that SA might be important to be aware of when designing artefacts. Applying the traditional aspects of SA together with knowledge elicitation, on the development of complex systems, opens up promising possibilities to enhance the design of complex systems.

Together, the user and a technical system form a JCS, were they act as cognitive agents through artefacts (Hollnagel & Woods, 1983; 2005; Woods & Hollnagel, 2006). By using the concept of JCS as a transition platform for the use of the two approaches the information from each of the approaches could be integrated into one understanding of the situation under study.

An agent-based systemic situational awareness theory has been proposed by Walker, Salmon, Green, Stanton, and Baber (2004). Later also, Stanton et al. (2006) describes distributed SA as system oriented, rather than individual oriented, with SA distributed between agents (human or non-human). Accordingly, it could be that one agent is aware of the perception, that is, level 1 SA, and another agent is aware of the comprehension, that is, level 2 SA, and a third agent is aware of the projection, that is, level 3 SA, and SA thereby distributed. The individual SA of each agent thereby does not have to be complete, for the system SA to be complete. However, agents have to have an awareness of who is likely to have useful information within the system (Stanton et al., 2006).

Adapted from Alfredson (2001a), Figure 7, depicts a similar view on how awareness of own or other agents' awareness could be regarded. This view, that individual SA contains an assessed SA of other team members' SA, is also the reason for focusing on degree of agreement as an important aspect of shared SA as applied in Paper 4.

Being aware of the own SA as well as of other agents' SA is mentally demanding. Even though the long term memory is practically infinite, the conscious awareness may not be aware of all the other agents SA simultaneously because of cognitive limitations, for instance, in short term memory and attention (Miller, 1956). The consciousness has to process information from the external world, through the senses, as well as memories and internal representations, such as mental models.

In this sense, the consciousness is in the centre of action, and our understanding of the world depends on which concepts are elaborated there. The distinction between man and artefact thereby might be questioned, since both are mechanisms that are used to create the concepts, and both are representations or concepts themselves. There might not even be a clear border-line between what is man and what is artefact, and if there are such a border-line it might not be all that important to define. We use our hands as tools, just as we use a pair of glasses. They are just components in a JCS.

Below, Figures 7, 8 and 9 help visualising how the earlier presented model of individual SA (cf. p. 30), may be extended to a description of joint cognitive systems, either human teams or mixtures of machine agents and humans.

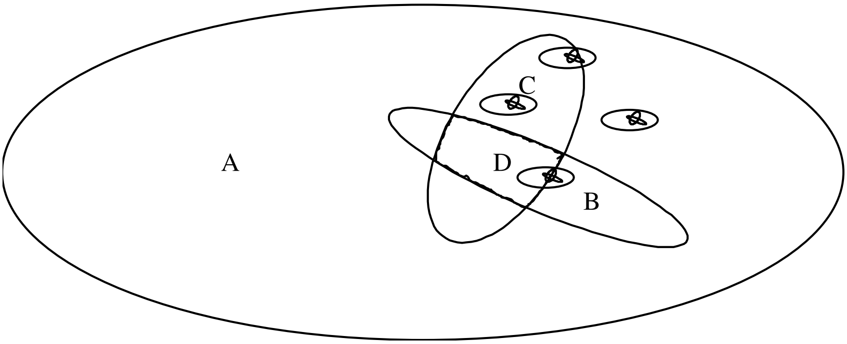

Figure 7. Representations of SA. As described earlier, a person's SA consists of a variety of components (see Figure 2). For instance, a person might be aware of another persons SA, as well as the own SA. The figure above illustrates how representations of SA might appear in a person's own SA.

We know that the cognitive performance of a team gets reduced by collaboration (Andersson, 1996; 2001), compared to the potential of each individual. However, in command and control settings, the situation often is too complex to be handled by any single individual. The control of the system towards a common goal is distributed between several decision makers who only have a part of the task model each (Brehmer, 1991).

Since the SA has to be shared within a team, shared SA and team SA are useful concepts for describing SA within and between teams. Figure 8, visualises one view of shared SA and Figure 9, visualises a view of team SA.

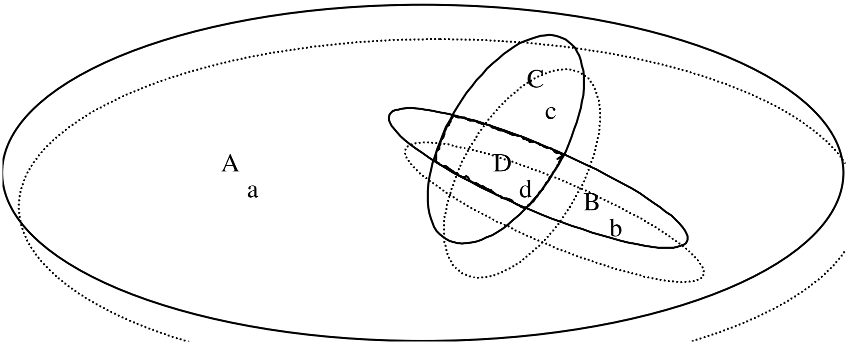

Figure 8. Shared SA. Two individuals, here represented by two overlaying representations, might share some SA. In the figure above one individual's SA is represented by the dotted line and the lower-case letters and another individuals SA by the other ordinary line and capital letters. In the figure above two individuals are represented, but there is no theoretical limit to the number of individuals represented. Also, one or more of the agents represented in the figure above, may be a non-human cognitive system, such as a computer.

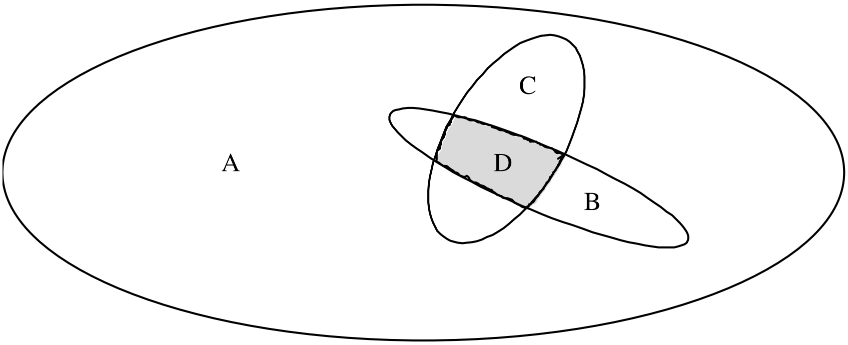

Figure 9. Team SA. A team could be regarded as a unit of analysis. As such, the team awareness is similar to the awareness of an individual. The awareness of the team is something else than the sum of the awareness of the individuals in the group. Here, the parallel to the individual SA is that the individual SA is more than its components (see also Figure 2, the palette metaphor). Apart from the individuals' SA, factors as, social behaviour, communication, organisation, and so forth influence team SA. One major distinction, however, between individual SA and team SA constitute consciousness of the individual, which does not have an obvious parallel in the team case. The figure could just as well represent JCS SA. Just as a team may be the unit of analysis, a joint cognitive system may be the unit of analysis.

Artman and Garbis (1998) define team SA as "The active construction of a model of a situation partly shared and partly distributed between two or more agents, from which one can anticipate important future states in the near future" (p. 2). In consistency with the definition of individual SA, team SA may also be described as:

High team SA is when the team members are fully aware of each other's high individual SA

It is far from trivial to do a precise separation of individual SA and shared SA between individuals or groups, since the individual, as a part of his/her understanding of the situation, has to be aware of other individuals' and groups' prioritizations of what is important and their current state of SA. Sonnenwald and Pierce (2000) use the term "interwoven situational awareness" when addressing individual, intragroup and intergroup shared understanding of the situation. Hauland (2002) concludes that team SA must be measured both at the individual level and at the team level.

Just as the differences between an externally assessed awareness for a person and a subjective self assessment could be used as a part of the assessment of the individual SA, differences within a team, or JCS, could be used for assessing the team SA in an equivalent procedure. It is important to have a reciprocal awareness between decision makers of SA aspects, to form a shared understanding of the situation at hand. *Degree of agreement* is therefore an important aspect of shared SA, as argued in Paper 4, and thereby also of team SA. If agents differ in their understanding of the situation (or task) they have more difficult supporting each other. Further, if agents differ in their understanding of each other's SA or in their understanding of the team SA (or JCS SA), they also have more difficult supporting each other. The equivalent approach could be used also for large organisations (i.e. groups of teams), systems of systems (or JCS of JCS), roles in the design-use process and more.

In a study by Berggren, Alfredson, Andersson, and Modéer (2007) shared SA was assessed in two various environments, and with various methods, including the use of degree of agreement. The environments were the microworld C3Fire that have been used for monitoring SA in previous studies (Granlund, 2003) and the Swedish Armed Forces Developmental Environment. For the experiments in both environments correlations were found between subjectively rated shared SA and degree of agreement, which indicates that both methods assess aspects of the same concept, that is, shared SA. Also, the subjectively rated shared SA was found to correlate to subjectively rated team performance in both occasions, which could indicate that shared SA contributes positively to team performance. Further, Berggren, Alfredson, Andersson, and Modéer (2007) calculated degree of agreement in two different ways, by means of calculating squared differences and rank order correlation and found no differences in the result between the two methods of calculating, but recommend calculating squared discrepancy since that is the easier calculation to perform and understand.

Endsley and Jones (2001) proposed a model of inter- and intrateam SA to explain how team members' individual SA are building SA within and between teams. However, what should be regarded as intrateam and what should be regarded as interteam is not always obvious. For instance, Proctor and Lipinski (2000) as well as Proctor, Panko, and Donovan (2004) regard crews in separate aircraft operating to the same objective as a team.

Dekker (2000) defines crew SA as "the extent of convergence between multiple crew members' continuously evolving assessments of the state and future direction of a process" (p. 51). Even more essential than a convergent assessment of a situation or an event is to have a true understanding of the particular event. A widely spread misjudgement is often worse than opposite views. Also, each individual has specific tasks to perform, and therefore different needs of information. Information presented to a user should therefore, optimally, be adapted both to the current situation and to the user's capabilities and skills (Alfredson, Lindoff, & Rencrantz, 2004).

One of the benefits of being able to adjust the information to each individual is that the information content and/or the information form may be adjusted to the specific individual and his/her needs. In order to optimize SA, an ideal system development should embrace a dynamic adaptation of interfaces to current vehicle status, situational conditions, contextual prerequisites, and individuals' status, operator performance, as well as historical behavioural data. What we know is that the design of the interface will determine the operators action envelope, and thereby the ability to manage the situation.

Sometimes, a conflict will appear between adapting information to each individual, and sharing the information between individuals in a team. When adapting information to specific needs of an individual or to a specific task, the information becomes distorted. The adapted information does not appear in the same manner for different users. When these users then try to communicate the information to each other, they may find that one user is confused by the form of the information that the other user is referring to. This may lead to misunderstandings and thus degradation of SA.

Therefore, the adaptation of information has to be moderate so that it will not make it more difficult for the users to communicate and share information. Also, the concept of SA itself has to be developed and adapted to future conditions to be considered useful, for instance, when developing complex systems.

2.3 Projection of the Future Status for the Concept of Situational Awareness

SA theory is evolving and this section is analysing and suggesting how the SA concept could be further enhanced. Here, the focus is on how the concept could be used in development of complex systems. Below, examples of trends and possible directions for the future development of the SA concept are briefly described. For instance, the question of if SA should be divided into sub-SA parts or regarded as global SA is discussed. Also, the section explains how, every time SA is assessed it is done with a purpose defined by a meta-situation. Finally, the concept of SM is suggested rather than the SA concept, for development of complex systems.

It is tempting to draw a box labelled SA in a figure among other boxes representing human cognition. Also, it is just as tempting to state that there are no such boxes in the head of a real person. However, even if SA does not exist in a little box in the head, it exists as a concept and it is used, and therefore it exists. It is thereby also important to understand how the concept is used and how it could contribute to our ambition to develop better systems.

Also, it is tempting to state that theories and models of SA cannot explain what *really* happens. That is, then SA could not be used to explain what has happened in a situation, since SA is a limited construct regarding essential components of a contextual complex reality, compared to, for instance, a qualitative analysis of a well observed and well documented case. For instance, to regard SA as a difference between actual SA and ideal SA has been argued to mislead analysis (Dekker & Lützhöft, 2004), even though a model of a future user's SA is necessary to be able to consider SA in design. It would be dubious to exclude studies on future situations only because it is easier to figure out what has happened than what will happen, since the value of having an idea of what will happen is larger than the value of knowing what has happened, when designing.

It is only in hindsight that we are able to describe what has really happened. Also, it is only in a very limited number of occasions when a specific episode has been scrutinised by research, in hindsight, that we could obtain such rich descriptions. There is a risk that we in our eager to find out what really happened in reality, instead tends to move our focus further away from a user's future reality. The reality for any user of a complex system is influenced by the design of that system made beforehand and not by an analysis performed in hindsight. Even if we are trying to avoid the hindsight bias, that is, distortion of past events because of knowledge of the event in hindsight, for instance, the tendency to regard events that has actually happened as more likely than alternatives that could just as well have happened, or even because we are trying so intense to avoid that bias, we might get another bias instead. The risk is that our research gets biased towards hindsight more than biased by the hindsight bias. What users need is good design. Designers need useful models to provide good design. SA is one framework among others that can provide such models.

The more different a future situation is from the situations that occur in today's practice, the less useful a description of the current practice is for design purposes. Complex systems often have long development cycles, significant changes between product generations, and long time between when they are introduced at the market and when users have become experts on the systems, so the feedback loops are often long. In the meantime models have to be used, for instance, for simulation purposes, and for an iterative development process to be successful. SA is a useful construct for such models. Even if it might be argued that SA is a construct that does not really have an undisputable place in a description of human cognition, it is a construct used in design, useful for designers. For instance, when developing a fighter aircraft, future technology, future tactics and future user profiles have to be anticipated and regarded to create a more accurate description of the future situations where the product is to be used than an alternative model only using experiences from older generation aircraft.

Even if future situations, per definition, do not exist now, they are relevant for the current development. Future situations are not the only non-existing situations of relevance. For instance, mental images or thoughts, and images retrieved from perception are all parts of the situation, and as such part of our SA. For instance, expectations are real in some sense even if they do not turn out to be correct, and they have real impact on our behaviour.

As indicated in this dissertation it is the awareness of the individual that is the centre of our experiences, with the conscious awareness as a central part. If so, it is questionable to focus too strongly on the interface between the human and its surrounding world. It is not always the most suitable unit of analysis to focus on one human and his/her behaviour, yet this is a common strategy.

The human awareness receives its information from inside the human as well as far away from the actual human machine interface. When we accept that, we can adapt to it and use it for our purposes. For instance, we may use *Wizard of OZ simulations* (enhance the functionality through human intervention) to simulate information in an interface without having the actual functionality implemented.

It is a good idea to validate the simulated environment not only on technical performance, but also according to user reactions. A good example is Magnusson (2002) that reported on similarities and differences in psychophysiological reactions between simulated and real air-to-ground missions.

As the systems we interact with grow to become larger, and more complex, we find ourselves further away from the data sources and the information tend to be more abstract. In both civilian and military systems we can notice this trend towards *system of systems*.

When the human-machine interface takes a less central position, the information on both parts of the interface becomes more central. Thereby, we may expect future systems to be characterised by problems such as misunderstandings between mental models and computer models, and not only interface problems, although some types of interface problems may very well remain or even increase in severity. In a nuclear power plant context Mumaw, Roth, Vicente, and Burns (2000) found that "… what makes monitoring difficult is not the need to identify subtle abnormal indications against a quiescent background, but rather the need to identify and pursue relevant findings against a noisy background" (p. 36).

To reduce misunderstandings between cognitive agents it is important to express intentions and have a reciprocal understanding of each other's cognition. It may be that HMC - Human Machine Cooperation is more suitable as a term of analysing future systems than HCI - Human-Computer Interaction (Hoc, 2000). Future interfaces may thereby be characterised by transferring the reciprocal understanding of the agents' cognition to help the agents build SA, through informative displays of not only the external situation but also about the interacting agents' mental state, implicitly or explicitly. Stock (2003) even foresees humour as a necessity in future interfaces, to enhance the human computer interaction. Boy (2000) expresses "The key issue, that is emerging from current practice with modern technology, is then *awareness* of what other agents are doing and have done so far, why they are doing what they do, and what they will do next" (p. 1.578). It is also possible that the consequences of misunderstandings between agents often are more severe than depicting problems for a single agent, due to non-optimal interface design.

This dissertation has been focused on *global SA* rather than sub-components studied elsewhere, such as, hazard awareness/perception (Horswill & McKenna, 2004; Olmos, Liang, & Wickens, 1995), mode awareness (Björklund, Alfredson, & Dekker, 2003; Funk, Lyall, & Niemczyk, 1997; Paper 3; Sarter & Woods, 1995), terrain awareness (Borst, Suijkerbuijk, Mulder, & Van Paassen, 2006), or spatial orientation that could be regarded as a subset of SA (Eriksson, Tribukait, von Hofsten, & Eiken, 2006; Previc & Ercoline, 2004), as well as temporal SA (Alfredson, 1997; Strömberg & Alfredson, 1997), tactical SA (Derefeldt et al., 1998a; 1998b; Paper 1), or the "multiple elements of awareness" that Pew (1995) identified a need to regard: spatial awareness, mission/goal awareness, system awareness, resource awareness, and crew awareness, similar to the three SA components suggested by Wickens (2002): spatial awareness, system awareness and task awareness.

In the other direction of the dimension of decomposing SA, aggregating the concept with decision making or even with action, is an appealing way of analysing problems relating not only to SA.

In fact, it is appealing to adapt to a pragmatic view of the use of concepts, and use whatever concepts help us solve our problems. However, in the name of pragmatism, let us not forget the importance of really knowing the meaning of the concepts we are using, since knowing the meaning of the concepts makes the analysis so much more interesting and useful!

Depending on the problem at hand, alternative concepts, complementing the concept of SA may be more suitable, for instance, the model including control modes, described by Hollnagel (1993). Similar to SA-models the model describes levels, but rather than being strictly quantitative as, for instance, high SA or low SA the levels are qualitative separated, giving the analysis complementary results to the use of SA.

Another alternative approach to SA that has been suggested is sensemaking that often have been applied for analysis at an organizational level (Weick, 1995; Weick, 2001). Endsley (2004) regards sensemaking as a subset of the SA processes, "forming level 2 SA from level 1 data through effortful processes of gathering and synthesizing information, using story building and mental models to find some formulation, some representation that accounts for and explains the disparate data" (p. 324). However, according to Klein, Moon, and Hoffman (2006) most discussions consider sensemaking more than SA, and than sensemaking rather could be regarded as the process of achieving the knowledge states of SA, related to encountered strategies and barriers. Although the concepts of SA and sensemaking mainly have been used in separated studies, the concepts are complementary and could thereby aid an evaluator with different perspectives.

Accordingly, concepts not only differ in the level of concept, where the higher levels of concepts, in a concept hierarchy, grasp more, as we see in the difference between the concepts of spatial SA and SA, where SA incorporates spatial SA. Also, concepts differ in if they regard quantitative or qualitative aspects, for instance, high or low SA opposed to various control modes. In the extreme qualitative end, we find research such as ethnographic approaches.

Once again, this dissertation has focused on a small part of the spectra of possible use of concepts and found the concept of SA useful, but this does not exclude the use of other concepts. On the contrary, an arsenal of many well-defined concepts may very well complement each other.

For future research, it would be interesting to incorporate on-line measures of SA that could be fed into the system, for implicit control and adaptation. Measures in such a system could be, for instance, psychophysiological measures, performance measures, system information, and semi-dynamic subjective measures.

Performance enhancements have been found for adaptive aiding during high mental workload (Wilson, Lambert, & Russell, 2000). It would be interesting to implement a computational model based on the model proposed in this dissertation, as basis for an adaptive system. Relating the actual SA to the perceived SA has been accomplished in a computational model by Shively, Brickner, and Silbiger (1997) as well as Burdick and Shively (2000), but it does not represent the model proposed in this dissertation.

One way of extending the SA theory is to include the meta-situation in which the situation is present. As visualised in Figure 10, any situation is present in a context or meta-situation, which in turn is present in yet another meta-situation, and so forth. Just as the context for a JCS could constrain another JCS, and shape the actions of a JCS (Woltjer, 2005), constrains of a meta-situation shape the outcome of a situation. There are often several reasons for studying a situation. It might be that the situation is studied as part of a study of user behaviour or interface evaluation, of a complex system. The study might also have several restrictions like timeframes, financial restrictions, and technical expectations. The researchers conducting the study, their background and expectations might even be considered to be included in the meta-situation. The study might be conducted with the ambition to be published in a scientific journal, or perhaps with the ambition to cut down on future development costs of a studied artefact. Both the situation and the meta-situation change dynamically, but perhaps at different paces. Still, both the situation and meta-situation might change drastically, because of an unexpected event. It is therefore necessary for anybody who is studying SA to be aware of the situation, the meta-situation, and their fluctuations, to come to the right conclusions. If they are, they can adapt to the changes. An example of a meta-situation could be the design-use process for a complex system used in the situation. Adapting to the dynamics of the situation and meta-situation requires four steps:

- Monitoring the changes.
- Understanding the meaning of the changes.
- Predicting the future impact on the study.
- Deciding and taking action.

The reader can now recognise the parallel between the first three steps and the three levels of SA proposed by Endsley (1988). What is referred to is the researcher's SA. Flach (1995a) stated that:

> It might also be argued that *science* is a complex, dynamic work environment (although the time constants are much longer than even process control). If this is true, then it might be useful to recursively apply the question of situation awareness to the science of situation awareness. (p. 25)

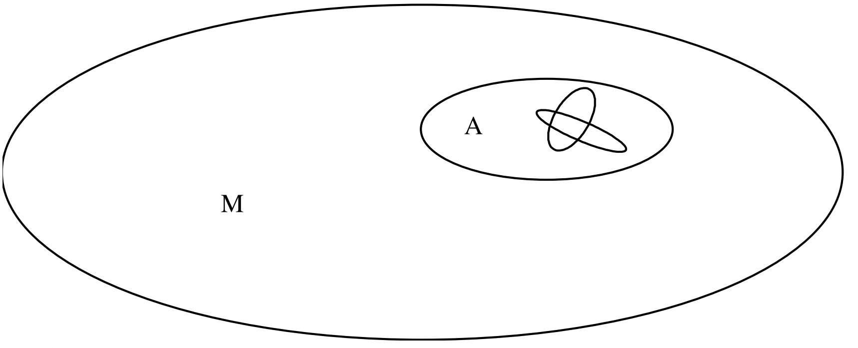

Figure 10. Meta-situation. When studying a situation (A), it always exists in a context or meta-situation (M).

The meta-situation is a relevant part of the researcher's SA. As the meta-situation changes the decisions and actions that are the most suitable also change. The methods and models used might then also change.

There are two fundamentally different ways of using information from an observed event. One is to compare the information to a predefined hypothesis to evaluate if the hypothesis might be correct. The other is to use the information to create or modify a hypothesis. Both ways could be used in one method, but often a method is biased to one of the uses. For instance, the SAGAT is more suitable for evaluating a hypothesis, but some knowledge elicitation method or even ethnographical methods might be more intended to create and modify the hypothesis.

A method is more useful if it provides means to adapt to the changes of the meta-situation. Hereby, it might raise the internal validity, since what is valid might change over time. A risk however, is that the external validity suffers from a too strong adaptation to the studied context, making it hard to generalise the findings. These are all aspects of the meta-situation. The meta-situation, including the researchers SA, is important to regard as well as the subjects SA. All roles in a design-use process have their SA and they all have to cope with their situation.

In conclusion, a wider view on SA than most often used is proposed here, which might rather be regarded as *situation management,* (SM) than situational awareness.

Situation management is about coping with the situation using all available means of awareness and action

The term SM springs from a focus on the situation and how to manage the situation, rather than a focus on what is perceived by an individual or team. Therefore, it is broader than SA, in that SM includes:

- *The SA of an individual or team.* A person's (or team's) SA is a part of the situation and thereby a part of the SM. For instance, if a person is lacking SA the situation that has to be managed is *a situation with a person with low SA.*
- *Every part of the perception-action cycle.* A person could by physical manoeuvring/manipulation or direction of attention influence the perception of the situation and thereby SA. By good SM, regarding the whole perception-action cycle, the SA thereby can be enhanced. This implies that correlations between SA and performance not only could be explained by the causal effect that high situational awareness will lead to good performance, but also the opposite causal direction, that is, good performance will lead to high situational awareness. Most studies on SA and performance have reported positive correlation between SA and performance, indicating that those situations were situations where the SA were too low to manage the situation and therefore an increase in SA would result in better performance. However, the correlation between SA and performance is not perfect, indicating that SA and performance is separate concepts. A view that seems to be opposite to this is the view that SA is part of the operative performance, which have been put forward based on non-causal modelling of empirical results from the aviation domain (Angelborg-Thanderz, 1990; Svensson, Angelborg-Thanderz, & Wilson, 1999; Svensson & Wilson, 2002). This view is highly reasonable in settings where the user is also a decision maker, and the performance *is decisions* or *is highly dependent on* decisions where SA is central. However, the views are not opposite if SM is recognized as the concept related to performance. That is, SM as part of the performance of a decision maker affects the situation and thereby the SA.
- *The management of mental resources.* Within the person or team, the mental resources could be managed to cope with, for instance, mental workload and optimization of SA. As an example, O'brien and O'hare (2007) found that cognitive management training including, for instance, strategies for using divided and focused attention, seems to help persons with lower SA ability to form SA while performing a complex task. For a team in a specific situation the situation might be managed by having parts of the team focused on achieving high SA, whilst other parts of the team could accept low SA and focus mental effort on performing some specific task. Also, for an individual it is not always optimal with as high SA as possible. For instance, a pilot who is trying to handle an emergency

situation will not perform better if he/she is putting all effort into being aware of *exactly where* and *exactly when* the aircraft is going to crash into the ground, which in its extreme only would result in hitting the ground fully aware. This pilot would perform worse if he/she would mange the mental resources to try to increase SA at all means, instead of acting. A similar phenomenon has been proposed for low level communication for distributed asynchronous cognitive systems where an observability horizon, at which attempts to increase control by increasing the amount of observations, actually would cause a loss of control (Lindh, 2006). It might also be that there in a specific situation are various aspects of the SA (qualitatively) that are lacking, rather than that the SA is generally too low in a quantitative sense. The SM might then be to change the qualitative focus of SA, rather than trying to quantitatively increase it to a higher level. It might even be better to have a lower SA, but with the relevant qualitative components, than higher SA. This part of the situation management is about *situational awareness management*, that is, managing the situational awareness of a situation.

- *External means of managing the situation.* There are also other means to manage the SA than those of the person or group experiencing the SA. For instance, a designer of a system used in the situation or a manager for the person, influences the possibilities for the person to perceive, understand and predict, which strongly affects the situational awareness. By good SM regarding design and management, the SA thereby can be enhanced. The affecting factor might be distant in time and space, but still affecting the situation and control of those factors could contribute to the management of the situation. SM therefore includes both the end users at the "sharp end" (Reason, 1990), right where the action is with all sorts of constraints, such as time constraints, as well as persons at the "blunt end". Reason (1997) explains how *organizational accidents* have become an increasing concern in modern complex systems where people at different levels of an organisation interact with technology, forming a system. Also, the safety culture of the organisation, reflecting individuals' attitudes, beliefs, perceptions and values in the organisation is important to manage safe operations, which is an important aspect varying between organisations (Ek, 2006). To analyse, understand and prevent organizational accidents it is not enough to focus only on the SA of an individual or team, although those concerns may be important contributing factors. At each level, within and outside, an organisation there are individuals and teams trying to cope with a situation and forming SA. SM recognises that the levels affect each other, as the goal shifts form optimising SA or striving to share SA, to managing the situation.

The focus on the situation, meta-situation, and SM is particularly relevant from the view of developing complex systems, since SM includes efforts in design and development. It is important to be aware of what will affect the situation to be able to perform an adequate analysis of how to, if not control the situation, so at least influence the situation with acts from various agents. By assessing the impact from each agent and potentially from any act of any agent and comparing those to each other, differences will appear that could influence concrete design efforts, as well as, broader considerations regarding, for instance, development efforts. Examples of such differences in SA relevant for development of complex systems are presented in the next chapter.

3 Differences in Situational Awareness

This chapter focuses on differences in SA, and how to assess them. Differences in SA are described and analysed. There are a variety of differences in SA that could be distinguished. Here, however, the differences studied are limited to 1) differences due to design alternatives, 2) roles in the design-use process, 3) context, and 4) level of analysis. These aspects are all important factors to regard in the meta-situation of development of complex systems.

1) Differences due to design alternatives are important to regard when developing complex systems since the design alternatives, in practice, often are the only components a designer have to chose between, when making a design decision. However, the differences due to design alternatives could also help direct further design efforts into a design supporting SA.

2) Differences due to roles in the design-use process are important to regard when developing complex systems since if the roles differs in the views of what SA is or in the notion of someone's SA, this would risk leading to poor coordination between the roles in the design-role process, affecting the user's SA negatively, directly or indirectly.

3) Differences due to context are important to regard when developing complex systems since SA is strongly context dependent and since the context and how it is perceived are affected by design. An aspect of special importance for development of complex systems is complexity, since complex systems are complex and often also used in a complex context.

4) Differences due to level of analysis are important to regard when developing complex systems since the level of analysis often could be actively chosen, and since the assessed differences in SA could direct the design efforts, affecting the user's SA, directly or indirectly.

Differences in SA can be both quantitative, and qualitative, for instance, there can be more or less SA, and also various aspects of SA (Alfredson, 2001a). Assessing the differences in SA and the ability to differentiate between various types of differences is important prerequisites to effectively managing the differences in SA in development of complex systems. To regard the differences is adding to the work of regarding the SA itself. Since the studies assessing SA often are costly is it important to extract the most out of them. A central issue for being able to regard the differences in SA is assessing SA adequately.

3.1 Assessing Differences in Situational Awareness

In this section several measurement techniques are reviewed. Physiological measures, performance measures as well as questionnaires and subjective measures are reviewed. The use of eye movements for SA assessment is described, in particular, and how that appears to be a promising approach for SA assessment, especially for visually demanding contexts. The review includes measurement technique criteria for SA assessment as well as actual assessment techniques.

To be able to consider differences in SA they first have to be detected. A fundamental criterion therefore is that the assessment technique should be able to detect SA. An additional criterion for assessing SA for the sake of development of complex systems is that that process is adequately aided by the assessment. Therefore, the assessment has to be both practically useful as well as theoretically valid, and give guidance on both actual design issues as well as on the design process.

There are various methods available to assess SA that could be used stand alone or together in clusters of methods. Using one general measure of SA would constitute an ideal strategy. The comparison between measurements made at different times and in different contexts would thereby be possible. However, as stated previously, SA is dependent on the overall context, or meta-situation, which indicates that it is likely that one measurement technique, is more effective in measuring SA in one meta-situation than in another.

For instance, the level of intrusiveness could be a critical factor. In a situation with high workload an intrusive measurement technique is likely to induce lower SA for the subject, but in a situation with low workload the effect could be diminishing or even lead to higher SA for the subject, since the measurement technique may serve as a secondary task, possibly enhancing subject vigilance.

Further, one meta-situation may demand a measurement technique with high diagnostic validity, whilst another meta-situation may call for a measurement technique which is especially easy to use. Accordingly, the ideal situation should allow the use of various measurement techniques, each adjusted to a specific meta-situation.

It is not easy to develop one measurement technique that fully covers all aspects of a complex construct like SA and thereby makes it possible to fulfil all purposes of the evaluation, or of the construct itself. Therefore, it is often necessary to use more than a single measurement to cover various aspects. This is, if it is not enough to measure some of the aspects of SA, for the purposes in question. As there is no method that can be identified as a standard method, it is advisable to combine methods in order to enhance the usefulness, but as stated above, the usefulness becomes lower if the number of measurement techniques distracts a subject. This might be the case if the measurements are used in field settings, for instance, in training or evaluation in a development process.

There are various criteria to be regarded when selecting combinations of measurement techniques. Validity is of course very important, and it is important all the way through the development process, as explained in Paper 6. Assuring that the method is based on a theoretical foundation will make it possible to obtain construct validity. To obtain concurrent validity the method has to be related to other measurement techniques. By relating the method to empirical experiences it is possible to obtain predictive validity. To have face validity, it should be agreed that the method addresses SA as this is generally understood and/or as it is defined especially for the regarded purposes.

Reliability is also important, both internal reliability and reliability over time. Drifts in the measured values over time may be hard to detect and even hard to correctly adjust if detected. If reliable data could not be produced by a measurement technique, a side effect is worse resolution in measured SA. That is, there is a risk that two occasions could not be determined relative to each other if there are too high amounts of uncertainties in the measured values.

If the intrusiveness is too high, reliable data may be difficult to extract. The method should not affect what it is supposed to measure, but if the difference is not too large and if it is in some way systematic, the method could be considered useful, although it might still suffer from some lack of sensitivity and thereby also diagnosticity.

Sensitivity here refers to the capability of a method to discriminate significant variations in the measures and diagnosticity refers to the capability of a method to discriminate causes of these differences and the ability to forecast the course. The importance of diagnosticity is highly dependent upon the purpose of the measurement method.

The purposes of the method provide additional criteria. A method should not only allow the deliverance of valuable data, but should also do this in a practical manner. This is the criterion *ease of use*. It has to be worth doing and should be done in the most suitable way with respect to the purposes. For instance, it might be pertinent that the method is not too difficult to create or too difficult to use. Furthermore, user acceptance, for instance, among operators and administrators is beneficial.

The equipment demanded for the measure should not be too expensive, bulky or inconvenient. The criterion bandwidth is relevant, since it might be necessary to measure SA within a wide spectrum. With dynamic measures the dynamic changes of a situation may be easier to find.

The possibility to receive real-time feedback from a measure should be considered. This might be an important criterion for implementation of on-line situation assessment, for instance, for development of decision support and on-line documentation regarding SA. Compliance to international standards is desirable as well.

There is a wide spectrum of existing methods for measuring SA. Many attempts have been made to categorise these methods in a suitable way to create a clear and logical structure for the purpose of making a rational comparison between methods. Such a comparison is necessary in order to determine which method should be used in a specific case. Endsley (1995a) used the following categories to divide techniques into:

- Physiological techniques, referring to techniques like P300 and other electroencephalographic measurements and eye-tracking devices.
- Performance measures, subdivided into: global measures, external task measures and imbedded task measures.
- Subjective techniques, subdivided into: self-rating and observer-rating.
- Questionnaires, subdivided into: post test, online and freeze technique.

A related example is outlined by Vidulich, Stratton, Crabtree, and Wilson (1994):

- Measures of effectiveness, referring to techniques that focus on the outcome.
- Subjective SA ratings, referring to techniques like SART (Situational Awareness Rating Technique; Taylor, 1990).
- Objective SA measures, referring to techniques like explicit memory probes and implicit signal detection probes.
- Physiological measures, referring to techniques like heart activity, eye blink activity and EEG (electro-encephalogram) activity measures.

Fracker and Vidulich (1991) preferred another variant of categorisation:

- Explicit measures, referring to techniques like retrospective event recall and concurrent memory probes.
- Implicit measures, referring to techniques like signal detection theory (SDT).
- Subjective rating measures, referring to techniques like Likert scales, multiple scales and SWORD - the Subjective WORkload Dominance technique (Vidulich, 1989; Vidulich, Ward, & Schueren, 1991).

Adams, Tenney, and Pew (1995), categorised methods into:

- On-line indices, where studies can be made on, depth of processing, ease of processing and of monitoring information-gathering behaviours.
- Indirect probes of SA, where studies can be made on, comprehension of displays or queries.
- Model-based approaches, both with and without computers.

Below, examples of possible measures of SA is presented, and commented. The measures in this text should be regarded as examples of possible measures of SA, and not regarded as the only possible measures of SA. Measures have been categorised into one of three groups and measures within each group have been studied with respect to the selection criteria presented above. The three groups are 1) questionnaires and subjective measures, 2) performance measures, and 3) physiological measures.

Questionnaires and subjective measures can be used to determine level of SA or to identify questions related to SA. Advantages of subjective SA assessment techniques are ease of implementation, ease of administration, low cost, and non-intrusiveness, as well as an ability to use them in real-world settings (Jones, 2000).

Since subjects formulate answers either to a rating of SA or to a question regarding SA, this kind of measure is dependent on subjects understanding of a situation. If the notion of what SA is differs between the user answering the questionnaire, and the person evaluating the response, there will be a problem with interpretation. Paper 2 reports on differences between the notion of SA found in research literature and pilots' notion of the concept, based on a questionnaire about SA distributed to a number of British military pilots and Swedish private pilots and a literature review.

However, the fact that the measures are dependent on a subject's understanding of a situation is not necessarily a disadvantage, as SA, by its very nature, is dependent upon the subjects understanding of a situation. The question is how this information may be collected. Below, some methods are reviewed.

SART (Taylor, 1990) is an example of a technique that uses a subject's estimation of the experienced situation to measure SA. Selcon and Taylor (1990), describe SART: "SART provides subjective estimates of attentional Demand and Supply, and ratings of Understanding, which are postulated to be the three primary components of situational awareness" (p. 5.1). There is a 3-dimentional SART scale and a 10-dimentional SART scale that both use the subject's ratings to measure SA. In the same text, Selcon and Taylor write:

> The advantage of SART is that, since the dimensions were elicited from the knowledge of aircrew, then they are likely to have high ecological validity. This is likely to be beneficial in applying the scale in the design of aircrew systems, particularly in comparison to more theoretically derived approaches to SA measurement. (p. 5.1)

The stated high ecological validity is an important aspect, but since the dimensions used in SART only partly cover what needs to be known to estimate SA, according to the suggested definition of SA in this dissertation (cf. p. 28), SART does have its limitations.

For instance, it appears to be very difficult to estimate subject's awareness of relative awareness, by using only subjective ratings, as it is likely that subject's estimation of the awareness of relative awareness is inaccurate when the subject has low awareness of relative awareness. That is, a person that have not noticed the own decreased SA, is not the most reliable source of information for detecting that decrease in SA.

Later, CC-SART has been developed, also regarding the cognitive compatibility (Taylor, 1995a; 1995b), which was used in Paper 1.

With SA-SWORD, a subject makes comparisons in pairs of competing design concepts, to judge which is the most SA enhancing (Jones, 2000; Vidulich & Hughes, 1991). Vidulich and Hughes (1991), describes SA-SWORD: "The SA-SWORD technique was an SA adaptation of a subjective workload assessment tool [called the Subjective WORkload Dominance (SWORD) technique]" (p. 1307). SWORD (Vidulich, 1989; Vidulich, Ward, & Schueren, 1991) uses a series of relative judgments to determine workload.

Referring to a study by Hughes, Hassoun, and Ward (1990), Endsley (1995a), expresses:

> Not surprisingly, the display that was subjectively rated as providing the best SA using SWORD was the display for which pilots expressed a strong preference. It is difficult to ascertain whether subjective preference led to higher SWORD ratings of SA, or vice-versa. (p. 69)

The Crew Awareness Rating Scale (CARS) consists of two sets of four questions (McGuinness & Foy, 2000). The first three in each set corresponds to Ensley's (1988) levels of SA, perception, comprehension, and projection, whilst the fourth question is about integration and goal identification. Also, the first set of four questions is assessing SA *content*, whilst the second set of questions is assessing SA *processing*.

Specially designed for the use in the military services Mission Awareness Rating Scale (MARS) have been developed based on the CARS methodology (Matthews & Beal, 2002).

Later, probe assessments have been combined with self-ratings using signal detection theory in the QUASA technique – Quantitative Analysis of Situational Awareness (Edgar & Edgar, 2007; Edgar, Edgar, & Curry, 2003; Edgar, Smith, Stone, Beetham, & Pritchard, 2000; McGuinness 2004; 2007). For instance, may a subject be presented a probe statement about the position of an enemy unit, a statement that could be true or false, and the subject's task is to state if the statement is true or not (Edgar, Edgar, & Curry, 2003).

Situation Awareness Global Assessment Technique (SAGAT) is a technique that uses a questionnaire to compare parts of a subject's experienced conditions with corresponding description of real conditions (Endsley, 1988; 1995a; 2000; Jones & Kaber, 2005). Endsley (1988), describes SAGAT in seven steps, where a pilot is given a simulated scenario, and then it is randomly stopped. The pilot is asked randomly selected questions that compare the real and perceived situation.

A composite SAGAT score is determined and the process is repeated. Endsley continues:

> The primary limitation of SAGAT is that the simulation must be halted to collect the data. (If the data were collected while the simulation was running, it would probably interfere with the task, and would allow the pilot to "cheat" by being able to immediately observe whatever was asked for.)
>
> The advantages of SAGAT are that:
>
> 1. It provides a current "snapshot" of a pilot's mental model of the situation, thus reducing the effects of collecting data after the fact.
>
> 2. It takes an inclusive look at all of a pilot's SA requirements, thus providing a global measure of SA.
>
> 3. It directly measures a pilot's knowledge of the situation, SA, and is not dependent on inferences from less direct sources.
>
> 4. It can be objectively collected and, most often, also objectively evaluated.
>
> 5. The measure possesses direct "face validity". (p. 101)

Endsley (1995b) points out that the SA concept "...rests on various information-processing constructs that have been discussed for some years. The value added by the SA concept is as a means of integrating these constructs in terms of operator's overall goals and decision behavior" (p. 51). Endsley (1995a) reports that stopping the simulation, and blanking the screens does not affect the persons SA, so that the internal validity should be significantly decreased.

Many researchers, however, do not agree, for instance, Patrick, James, Ahmed, and Halliday (2006) state that one of the major advantages with observer methods is that it "obviates the need for successive 'freezes' of the scenario, which are likely to interfere with the very phenomenon being assessed" (pp. 414-415).

Some researchers try to improve the understanding of the concept of SA by placing it in a circular model of human information processing, in contrast to the work by Endsley that is based on a linear model of human information processing. Tenney, Adams, Pew, Huggins, and Rogers (1992), as well as Smith and Hancock (1995) uses Neisser's perceptual cycle (Neisser, 1976), as the foundation of their work on SA. By using this model the dynamic features of situation awareness is being expressed in a more explicit way.

The risk we are taking when we are stopping a scenario and blanking the screens is that the measuring of SA changes the subject's SA. This becomes obvious when reflecting on Neisser's perceptual cycle, where constant feedback from the environment surrounding the individual is used to build the dynamics of perception.

In what way the subject's SA is affected is not obvious. It could be that the subject's SA is increased by making the subject change his/her interest to the questions in the questionnaire, but it could equally well result in a decrease in SA, since the subject may forget what was important because of what is asked in the questions or because the interruption itself affects SA, as a dynamic concept. These are all threats to the internal validity, which should be regarded when deciding on measurement techniques of SA. McGowan and Banbury (2004) reported on an experiment where an interruption-based measure of SA had impact on performance. Also, it is not obvious how various behaviours in terms of skill based, rule based or knowledge based behaviours (Rasmussen (1983, 1986) are affected by an interruption in itself or the value of an assessment made during the interruption.

Another threat against the internal validity is if the memory quiz really measures SA. Neisser's perceptual cycle (see Figure 1), stresses the constant information flow between the person and the person's environment. Information from schema, inside the person, directs the exploration of the outside world, so that samples create objects of available information that can modify the schema, and so forth. In this light it is far from obvious that whatever is available in the working memory by that time is the best representation of what SA really is.

Endsley (1995a) has reported that there are no effects of the SAGAT interruptions. However, just as there is a problem measuring the absence of SA, there is a problem of measuring the absence of negative effects of the interruptions. Why were there no effects reported? Was it because there were no effects, or because the effects were not found?

If there really were no effects on stopping the simulator, it would be tempting to question if the simulator really supports SA compared to a real flight situation. If you stop a simulation that does not support SA, you would probably not miss it as much as if the simulation was really good.

With a methodology derived from SAGAT Jones and Endsley (2004) has made efforts to develop real-time probes for measuring SA. In contrast to SAGAT the screens are not blanked and the subject instead is periodically presented a single verbal query and is required to verbally respond.

Also, SPAM (Situation Present Assessment Method) does not include blanking of the screens and relies on user response time as a measure of SA (Durso et al., 1995; Durso, Bleckley, & Dattel, 2006; Durso & Dattel, 2004).

In an amalgamation of SAGAT and SPAM researchers at the FAA Technical Centre has developed SAVANT - Situation Awareness Verification Analysis Tool (Jeannot, Kelly, & Thompson, 2003).

Also, based on SPAM on online probing technique SASHA_L – SA for SHAPE online, where SHAPE is a project acronym for Solutions for Human-Automation Partnerships in European ATM, has been developed to be used together with an offline, ten questions questionnaire SASHA_Q (Jeannot, Kelly, & Thompson, 2003).

SACRI (Situation Awareness Control Room Inventory) is another attempt to modify SAGAT, which have been used in both nuclear power plant control rooms (Hogg, Follesø, Strand-Volden, & Torralba, 1995) and for car design (Walker, Stanton, & Young, 2006).

For measurement of SA in air traffic control the method SALSA (Situation Awareness bei Lotsen der Streckenflugkontrolle im context von Automatisierung – situation awareness of en-route air traffic controllers in the context of automation) has been developed taking into account that air traffic controllers use an event based mental representation of the air traffic (Hauss & Eyferth, 2003).

The methods above are not redundant, but complementary. They focus on various aspects of SA. For instance, SART and SAGAT have been found not to be correlated (Boag, Neale, & Neal, 1999; Endsley, Selcon, Hardiman, & Croft, 1998) and Snow and Reising (2000) found that SA-SWORD and SAGAT were not significantly correlated in a study where the metrics were used in a simulated low-level ingress scenario.

Verbal protocols have been used to assess SA (e.g. Durso et al., 1995; Walker, Stanton, & Young, 2007). When verbal protocols are used, the subject is instructed to verbalise or think aloud throughout the mission. This method, which is extensively used in problem solving research, purports to provide information regarding the subject's thinking processes which would be of great potential value in assessing SA. Unfortunately, the method is also very intrusive, mainly because:

- Not all thoughts are simple to formulate. The subject has to come up with answers to several questions, for instance: What am I currently thinking? Which of the thoughts should I choose to verbalise? (The thoughts appear more frequent than the subject is able to verbalise) How should I express the thoughts? and so forth. These questions increase the subject's mental workload, which in turn could affect SA.
- Overt verbal responses tend to interfere with users' task performance. Aspects of the psycho-motoric responses and speech perception are especially vulnerable.
- The volume of data collected with this method is typically large and extremely time-consuming to analyse.

Using self-ratings, subjects could rate their own level of SA, (for instance, on a Likert scale), making it possible to estimate SA. This is, in some sense, a direct measure, since the experience of SA takes place in the subject's mind.

However, according to the definition of SA suggested in this dissertation, SA is dependent, not only on relative awareness, but also on awareness of relative awareness. Or, as expressed by Helander (2006) "One cannot just ask people if they have situation awareness; they wouldn't know if there are aspects that they are unaware of" (p. 86). If awareness of relative awareness is good then a subjective rating of SA is more likely to be adequate, but poor if awareness is low.

There are also other means to use subjects' ability to make self-ratings. For instance, it is possible to let a subject rate something connected to SA, but not SA itself. The rating is then used in a model of SA to estimate SA. One of the advantages of self-rating is that it could be a direct measurement, whereas the latter measurement is considered as more indirect. Due to the subjective component of self-ratings, the lack of objectivity could be considered as an argument for using an objective measure instead, or as a complement.

The FASA (Factors Affecting Situation Awareness) questionnaire, developed for high diagnostic assessment of aircrews' acquisition and maintenance of SA, consists of 30 questions equally divided into five sub-scales to be answered on a five point Likert scale by commercial airline pilots (Banbury, Dudfield, Hoermann, & Soll, 2007).

SA could also be assessed by subject's responses to questions. In this way the knowledge from specialists (the advantage of observer rating) can be combined with a quite direct measure (the advantage of subjective rating) and still be objective, since answers can be compared to a reality. There are, however, problems with this approach. One problem is to know which questions are to be asked and another is when and how to ask them. Generally it could be stressed that accuracy of data is especially dependent on the theoretical model used to describe SA.

Another example of a subjective measure is PSAQ - participant situation awareness questionnaire that uses a five-point scale (Matthews, Pleban, Endsley, & Strater, 2000).

Also, a rating scale has been developed by the Swedish Defence Research Agency (Berggren, 1999; Alfredson et al., 1997) based on a description of SA (Alfredson et al., 1996).

Later similar rating scales have been used, for instance, China Lake Situational Awareness (CLSA) scale (Adams, 1998).

By combining questionnaires about SA with questions about motivation, expected performance and mood a better description is obtained (Angelborg-Thanderz, 1990; Svensson, Angelborg-Thanderz, & Sjöberg, 1993; Svensson, Angelborg-Thanderz, Sjöberg, & Gillberg, 1988; Svensson, Angelborg-Thanderz, Sjöberg, & Olsson, 1997).

Using observer rating, where trained observers watch subject behaviour and then assess level of SA, is another method to estimate SA. This approach has the advantage that the knowledge of the trained observer is used to understand what is happening, something that a subject with a low level of SA cannot do.

This method could be regarded as more objective than the self-rating because the trained observer is not experiencing SA directly, but rather indirectly through perception of what is happening. The subjective part of the measurement is that the trained observer evaluates what subjectively is perceived. The measurement is dependent on the ability of the observer. This dependence makes it difficult to know for sure if the data from the measurement is reliable. The criteria, which the observer uses, might not be clear and a subjective opinion of the observer might influence judgement. The method may still be useful, especially in contexts where observer ratings can be structured.

Questionnaire data and physiological data could be correlated with these ratings and hence provide a good estimation of the SA. Also, combination of observer rating and objective questions on flight knowledge, similar to SAGAT, have been found to be correlated as measurement of team SA in low experience level aviators (Prince, Ellis, Brannick, & Salas, 2007), and could be included in a combined set of assessment methods.

SARS (Situational Awareness Rating Scale) is used both as an observer rating, but also for self assessment and peer assessment using a 6-point Linkert scale for 31 elements (Bell & Lyon, 2000; Bell & Waag, 1995; Waag & Houck, 1994), and have, for instance, been used for assessing SA in a police shooting simulator (Saus, et al., 2006).

SALIANT – Situation Awareness Linked Indicators Adapted to Novel Tasks (Muniz, Stout, Bowers, & Salas, 1998), is an event based SA assessment method using previously known SA behaviours for the development of scenarios for observation of SA behaviour, which has been used in both low fidelity simulations as well as in operational context (Milham, Barnett, & Oser, 2000).

A similar approach is used in SABARS – Situation Awareness Behaviorally Anchored Rating Scale (Matthews & Beal, 2002; Matthews, Pleban, Endsley, & Strater, 2000), where an expert observer uses a five-point scale for 28 rating items specially designed to assess the SA of a platoon leader, also used in the police shooting simulator study by Saus et al. (2006).

Similar to an observer rating, peers could assess each other's SA. A disadvantage compared to observer rating is that the peer could not fully attend to the assessment of SA among his/her peers, since the user also have task related work to attend to. An advantage however, is that the rating could not only be used for assessing the SA of the rated, but also the SA of the peer making the rating, in contexts where the assessment of other users' SA could be regarded as a part of the own SA. Similar, reciprocal peer ratings can be used for assessment of shared SA, when degree of agreement could be considered to be a part of the shared SA as the approach used in Paper 4.

The strengths of performance measures are that they assess final system performance, record user's actions, and that sufficiency of SA could be inferred in some situations (Pritchett & Hansman, 2000). It is possible that performance measures could indicate SA levels, since high SA enhances performance. However, SA does not necessary have a linear relation to performance. Thus, it is not always the case that low SA leads to poor performance. As discussed previously, one of the main reasons for using the term SA is that it is something else than just simply a synonym for another concept such as performance, mental workload or attention. If performance measures are used, complementary measures are necessary.

When an aircraft flies from point A to point B, this is accomplished by means of various courses of actions the performance is dependent on how it is done. Consequently, various measures of performance provide variations in the information to be valued. The problem is determining which measure of performance that provides the best indicator of SA and how the information from that measure should be connected to information from other measures to assess level of SA.

Performance measures have been used as implicit measures of SA (Gugerty & Tirre, 1995; Vidulich, 1995), for instance, the Global Implicit Measure (GIM) approach, which is based on "the use of objective measures of the status of the pilot-aircraft system taken with respect to previously defined rules of engagement." (Brickman et al., 1995, p. 339).

One of the advantages of performance measures is that they often could be assessed in real time, providing an opportunity for feedback to the system, and on-line assessment of SA as stand alone techniques or combined with other techniques.

Since SA is closely related to performance, certain aspects of SA have sometimes been regarded as part of the operative pilot performance (Angelborg-Thanderz, 1990; Svensson, Angelborg-Thanderz, & Wilson, 1999; Svensson & Wilson, 2002).

Also, communication is sometimes an important part of both SA and performance. There have been some attempts to analyse SA through communication analysis. In some contexts communication is part of the performance, as an important task to perform for the users, whilst in other contexts it is rather to be considered as a coordination activity directly related to SA and only indirectly to performance. By categorizing communication data from an air combat simulator Svensson and Andersson (2006) contributed to how communication in a fighter aircraft links to performance.

Communication has also been used as a measure of team situation awareness in the process based measure CAST – Coordinated Awareness of Situation by Teams (Gorman, Cooke, Pederson, Connor, & DeJoode, 2005; Gorman, Cooke, & Winner, 2006).

A wide range of physiological and psychophysiological measures could indicate various aspects of SA direct or indirect, by providing information of a person's mental state. However, very few experiments have studied psychophysiological measures of SA (Wilson, 2000). Candidate techniques includes, positron emission tomography (PET) scanning, electro-encephalogram (EEG), near infrared spectroscopy (NIRS), evoked potentials, heart rate (HR) and heart rate variability (HRV), blood pressure, galvanic skin response (GSR), catecholamines, electromyography (EMG), temperature biofeedback, and so forth.

The links between the techniques above and SA is not yet understood, and they now rather complement other techniques in specific aspects of interest. However, there are indications of that EEG could help assessing SA at level 2 (i.e. recognition of significance) as increased frontal lobe activity in the theta and gamma bands (French, Clarke, Pomeroy, Seymour, & Clark, 2007). Koester (2007) found that EEG could be used to detect high demands for SA at level 3, through high levels of stress and mental activity in a study conducted in maritime simulators. Saus et al. (2006) reported on individual differences in HRV associated with SA, in that persons with higher resting HRV had higher SA in a study in a police shooting simulator. Promising techniques are techniques for measuring Eye-Point-Of-Gaze (EPOG) and other eye related data, even if those techniques also are under further research and development for assessment of SA. Below, aspects of eye movement for assessment of SA in visually demanding environments such as cockpits are discussed in relation to interface design, and development of complex systems.

With sophisticated EPOG equipment, a battery of measures becomes available, for instance, objects of fixation, fixation dwell time, the frequency of fixations, inter fixation interval, the visual scanning patterns, entropy, eye blink frequency, saccadic eye movement, and pupil dilation, and so forth.

Computer interfaces has been evaluated with eye movement data in several studies. Graf and Krueger (1989) used eye movement data as a performance measure, which served as an evaluation of an alphanumeric display, as well as a measure of the cognitive load of the user. Lankford et al. (1997) have used eye gaze tracking in the design of telecommunications software, and arguing that:

> Graphical interfaces can benefit substantially if eye tracking data on a user's visual interaction with the software is considered as a part of the design process. This eye tracking data would supplement the more traditional human performance data (e.g. timing and errors gathered during usability testing). (p. 1371)

Further, Goldberg (2000) reviewed eye movement-based evaluation of interface usability, and found that eye movement-based analysis was useful, if the right evaluation criteria were used.

In addition to aiding the design process, eye movement data can be used for training purposes. That an interface is good is not only determined by that the user is performing well for the moment, but also that the user is learning, and becoming better over time. By using the interface experience will be gained and hopefully lead to that the user is better prepared for new tasks. It has been pointed out that a good decision support should also be good for training an ability to handle decisions in the right way (Zachary & Ryder, 1997). This could very well be equally true for certain aspects of a display design.

Even if the use of eye movement data cannot ensure that a display makes the pilot learn quicker, some types of recording of eye movement data could very well be used in learning situations. The eye movement data can serve as feedback to the learning user, about his/her behaviour, for instance, scanning behaviour. Just to be able to see what was really looked at and to study the own behaviour could be a motivating and educating moment. For instance, the ability to watch in real-time exactly where a practising pilot is looking while flying could be useful to an instructor at certain times. EOG and other psychophysiological measures have been used to draw conclusions on how to improve basic aviation training (Dahlström, 2007), which in turn may enhance the SA in a future aviation situation through the trained pilots.

Another application for SA assessment through eye movement data is on-line monitoring, as part of the system controlled by the user. With on-line monitoring of the eye movements novel functions will become possible (Alfredson & Nählinder, 2002). A simple feature, available in various commercial systems, would be to detect when the eyes are closed. By differentiating between various kinds of blinks, information about the user's mental state might become available (Paper 5). This information could, for instance, be used to decide when to activate automated functions by the system, when the pilot is assumed to be less capable than necessary. The presentation on the displays could therefore also be adjusted to where the pilot currently is looking. For instance, a display that appears in the peripheral part of the field of view could stop visualizing detailed information that the pilot cannot extract anyway, and instead show salient information on computer screens in the peripheral parts of the field of view, if needed.

However, it is not obvious that SA may be measured through psychophysiological measures. Wilson (1995) expresses that "...there may be settings in which psychophysiological measures can provide information relevant to determining if an operator is or is not aware of certain types of situations and whether or not the operator is actively seeking information" (p. 141), and Byrne (1995) stated that psychophysiological measures can yield viable estimates of essential components of SA.

The relationship between eye movements and SA is not obvious and the relationship between eye movement and SA do not necessary has to go over the concept of mental workload. The relationship between eye movements and memory has been studied for long, as well as the relationship between eye movements and cognitive models (Stark & Ellis, 1981).

An example of an attempt to measure SA with eye movement data, for instance, fixation rate, is the study about chess players' SA (Durso et al., 1995). In this study, five SA methodologies were compared, and what was found was that eye movements "seemed to be the most complicated and yielded the fewest insights" (p. 301). However chess playing is an activity that is very different from flying an aircraft. It might be that in other areas it is easier to find relationships between eye movements and SA.

Smolensky (1993) stated that "Evaluation of eye movements might be used to assess certain contributing constructs to situation awareness (SA)" (p. 41), and Hauland (2002) used eye movement data to measure team SA in training of en route air traffic control. Paper 3 used eye-point-of-gaze recordings together with recordings of communication and automation state to assess mode awareness. Later also Sarter, Mumaw, and Wickens, (2007) found failures to verify manual mode selections and notice automatic mode changes, through analysis of eye-tracking in combination to behavioural data and mental models.

To be able to judge whether eye movement data is useful for assessing SA, it is useful to distinguish between different kinds of eye movements, and some ways to formalise the movements into measures. Let us first take a brief look at some types of eye movements.

A person uses both foveal vision and peripheral vision to gather visual information. With the foveal vision, which is about two degrees of visual angle, the person can perceive a detail at a certain point, the eye point of gaze.

The movements of the eye that maintain fixations are called fixational eye movements. To make the image of constant moving objects staying in the fovea, pursuit movements are made by the eye. Movements of the eye that are jumping from one point to the other are called saccadic movements. Vistibuloocular eye movements are compensating for the movements of the own head. The optokinetic eye movements are used when a large part of the visual field is rotating. Finally, the vergence or disjunctive eye movements are used, for instance, to follow a target that is appearing closer or further away, by moving the eyes in opposite directions to one another.

The time that the eye is not moving, but staying in the same position, is called fixation time or dwell time. Fixation times are used to evaluate computer displays or other interfaces. It is not always obvious how to interpret the meaning of a long fixation time. For instance, important information can make a user fixate at some information for a long time, as well as badly displayed and confusing information can lead to long fixations. However, by comparing the fixation times to other data it can sometimes be possible to draw valuable conclusions from the data.

Certain instruments or certain parts of a computer display in an aircraft are more important than others are. For instance, the importance of the attitude director indicator has been indicated by several studies using eye movement data (Fitts, Jones, & Milton, 1950; Harris & Christhilf, 1980; Itoh, Hayashi, Tsukui, & Saito, 1990). To know the relative importance of different parts of a computer display is needed to be able to make the right design decisions, when conflicting interests occur in the design.

Because the movements of the eye can follow a moving object, it is possible to follow a motionless image upon the retina, but still perceive the object to be in motion, which is explained from an evolutionary approach by Walls (1962). This allows us to keep an object fixed on the retina, which is necessary to be able to see the details in a moving object. In an aircraft cockpit, for instance, moving objects can appear in the outside view as well as inside the cockpit. Because of the dynamic nature of some of the displays in a military aircraft, moving objects will be quite frequent in some of the displays.

For a pilot it is not enough to follow an object with his/her gaze and to perceive high resolution image of that single object. Information has to be searched for. The search time is very much dependent on the number of elements to search (Mocharnuk, 1978). To search for a target the pilot is using visual scan patterns. Scan patterns can carry a lot of information about the interaction between the pilot and the cockpit displays. Often individual differences appear when scan patterns are mapped (Dick, 1980; Fitts, Jones, & Milton, 1950; Marshall, 1996). Dick (1980) indicated the existence of a number of mini-scan patterns, when studying approach/landing sequences with a simulated Boeing 737. The mini-scan patterns were not only different between pilots, but also between parts of the sequence of flight. Sometimes the scanning behaviour has been different between experts and novices, as for car drivers (Mourant & Rockwell, 1972). Eye movements while scanning pictures has been studied (Gould, 1976). Further, fixation times and scanning behaviour have been found related to pilot mental workload (Itoh, Hayashi, Tsukui, & Saito, 1990; Kennedy, Braun, & Massey, 1995; May, Kennedy, Williams, Dunlap, & Brannan, 1990; Svensson, Angelborg-Thanderz, Sjöberg, & Olsson, 1997).

Eye blink activity is sometimes detected together with recording of other eye movements. It can be used without or together with other eye movement data, for instance, to assess mental workload as described in Paper 5. Blink rate, blink duration, blink latency, and blink amplitude are examples of eye blink activity.

Eye blink activity has been and is used as an indicator of mental workload. For instance, Wilson, Purvis, Skelly, Fullenkamp, and Davis (1987) found that pilots tend to blink less during moments with important information. It has been reported that cognitive processing suppresses eye blinks (Boehm-Davis, Gray, & Schoelles, 2000; Veltman & Gaillard, 1998; Wilson, Fullenkamp, & Davis, 1994), even though some research indicates the opposite (Wood & Hassett, 1983). This is explained by the pilot's tasks that were so visually demanding that the blinks during the runs were suppressed. Stern (1988) refers to blinks as "mental punctuation", because blinks tend to occur when cognitive processing is fulfilled.

Also, it has been found that the combination of eye blink activity and cardiac measures are sensitive to changes in task demands, and classifying mission segments (Svensson & Wilson, 2002; Wilson, Badeau, & Gawron, 1993; Wilson & Fisher, 1991).

To know where the pilot is looking, detecting the pilot's eye point of gaze, it is not enough to know his/her eye movements. If the head is moving, which it is from time to time, for instance, on a non motion-restricted pilot, the head movements also influences the eye point of gaze. It is not all methods that allow as high quality on the eye movement data, while measuring eye point of gaze, as when only measuring eye movements. The extra criterion, that the head movements also should be detected can sometimes motivate the collection of other than pure eye movement data. Less time consumption and with the ability of separating eye movements from head movements, are considered as the two driving forces to accomplish a good method to measure eye movements in operational conditions without being obtrusive. Several methods exists that measure eye point of gaze.

However, the eye point of regard does not have to be the same as the eye point of gaze. Apart from the foveal vision the pilot can use the peripheral vision to extract visual information. What information is used from the peripheral vision is harder to track than the information from the foveal vision, because it covers all information in the field of view, even if the information is not used.

Fitts, Jones and Milton (1950) demonstrated the benefits of recording eye movement data in a cockpit, and explained that ”If we know where a pilot is looking we do not necessary know what he is thinking, but we know something of what he is thinking about” (p. 24). It is hard to find a reason to why this should not be still true, even if the cockpit itself does not look the same now as when the statement was made.

In the study by Fitts, Jones and Milton (1950) a camera was installed behind and to the side of the pilot, and photographs were made of his eyes by the reflection of a mirror. By doing so, the pilot’s normal fixation and movement patterns were not disturbed. Schieber, Harms, Berkhont, and Spangler (1997) as well as Svensson, Angelborg-Thanderz, Sjöberg, and Olsson (1997) used a similar approach by video taping the pilot’s eye movements. By using a video camera it is possible to record where the subject is looking, if both the head movements and eye movements are recorded.

By using a video recording, not only is the eye point of gaze possible to extract, but also the videotape could be used to analyse other behaviour, like oral communication. If the videotape is analysed by hand it could very well be quite time consuming. Since the recording techniques has followed the technical evolution over many years between the studies above, for instance, the use of a video camera, the prospects of collecting good eye movement data, in an easy and effective way has improved.

Now, there are several technical solutions on the market that provides automated analysis of camera views, for instance, by image recognition. On the other hand other factors may make it more difficult. A modern cockpit has less traditional instrumentation and more computer screens. The information on a computer screen is often dense and cluttered compared to traditional instrumentation, and the displays are often changing in more rapid dynamic presentations, especially displays in military aircraft.

Eye point of gaze has also been recorded in operational aircraft using helmet mounted camera that records both the scene in front of the pilot and his eye movements within the scene (Thomas, 1963). Alfredson, Nählinder, and Castor (2005) compared five techniques for measuring eye movements in applied psychological research. The report describes the techniques, GazeTracker, EOG, Jazz, Smart Eye and Video based, with which the authors have experience of data collection in applied military settings. The advantages and disadvantages of the different techniques are discussed and examples of their uses are presented.

The EPOG measures contribute to understanding the level of a subject's SA, since a person's awareness frequently is dependent on what that person sees. There are however some limitations in the usefulness of these measures:

- The measures rely on visual information, and human awareness is affected by other modalities such as auditory information (for instance, radio communication). Thus, EPOG gives an incomplete picture of the human awareness. The proportions of the components of SA, as stated above, also differ according to contextual demands. This means that the influence of visual perception on SA varies; changes that EPOG measures cannot identify.
- Within the area of visual information, measures have imperfection. For instance, a subject can be aware of an event or an object (but only large and familiar objects, such as the sun or the ground) in the field of vision even if it has never been focused or directly tracked.
- Focusing on or tracking an object does not necessarily demonstrate a high level of awareness. Even if a person looks at an object, the interpretation of how that percept affects awareness and thereby SA is not measured. Additionally, it is not surprising that a subject is focusing on something. Accordingly, the risk is that, a post-hoc explanation of the visual search pattern in this case will be misleading.
- Another limitation, which is a limitation of the implementation of the measures rather than of the measures themselves, is what information should be included or excluded. For instance, the question of how long the shortest relevant time to analyse should be. Apart from individual differences, this can vary substantially from time to time. Sometimes, only a very short glimpse is enough to make the subject aware of something and sometimes it is not. Only an absolute minimum can be given. What remains is the variable fixation duration.
- The fifth limitation is time related. As stated earlier (see Figure 3) SA is dependent on information encompassing current state, past states and future states. Most of the information gained from EPOG measures is related to the current state and only a fraction can be related to past states and future states, and there are difficulties in judging which information that is related to which state. This limitation does not have to be reason enough to exclude EPOG measures as a way to indicate SA, especially if

the user task is visually oriented. In Endsley (1995b) a taxonomy for classifying and describing errors in SA is presented and the result of a study of incidents in North American civil aviation reports that the main part of the failures is related to perceptual processes (Jones & Endsley, 1996). Also, Sneddon, Mearns, and Flin (2006) found most errors in SA occurring at the perceptual level, in a study of offshore drill crews. Failures in perceiving a situation reduce the amount of correct information of the current situation more than they affect the information of future events. This could be used as an argument for the use of EPOG measures at least within civil aviation, since the information gained from EPOG measures is mainly within the critical field of current state information.

Combining subjective and objective measures can help the comprehension of their meaning. By comparing pilot comments to dwell times on displays, Harris and Christhilf (1980) came to the conclusion that pilots tend to say that they look at a display with their periphery, while, quite possibly what they really meant was that they look at it with a quick glance. In Paper 3, eye movements where collected and compared with vocal statements such as call-outs studying pilot behaviour. Lorenz et al. (2006) combined subjective SA assessment with EPOG to study effects of pilots receiving onboard taxi navigation support.

By combining eye movement data with other measures the eye movement data can be put into a context. The context of a cockpit display is the cockpit and the tasks carried out in the cockpit, by the pilot or pilots. Since a lot of the information flow from the displays to the pilot is passing the pilot's eyes, it is of interest where the pilot is looking. By knowing where the pilot is looking it is possible to draw some conclusions about the computer interface that the pilot is using, and how to change the interfaces, so that they better serve their purposes. The combination of eye movement data to other data sources, to provide a better assessment of SA and a more complete picture of what is happening, is really an interesting area that is still improving. Integrating data on where a person is looking into a representation of a virtual or real environment, opens up for powerful interaction design. Sennersten et al. (2007) verified the accuracy in the use of automated logging of dynamic 3D-objects of gaze attention in a game environment, with potential benefits in the design process as well as integrated into products.

From the section above it could be concluded that differences in SA could be assessment by several measurement techniques, with different strengths. By various combinations of measurement techniques there are great potentials for aiding a development process for complex systems. Techniques can complement each other in various ways. Two techniques assessing the same aspect of SA may, for instance, provide different qualities in data to provide a richer view of that aspect. Also, two techniques could complement each other in assessing different aspects of SA, so that the following analysis could provide a good estimate of the global SA. When combining measurement techniques a straightforward strategy is suggested that allocates one or more techniques to every aspect of SA that is of interest (Alfredson, 2001a).

For instance, if the difference between what you are required to be aware of and what you are aware of is accepted as a SA-indicator, two important aspects to assess are *required awareness* and the current *awareness*. Also, if the difference between what you are aware of and what you think you are aware of are accepted as another SA-indicator, two other important aspects to assess are *relative awareness* and *awareness of relative awareness*. Further, if differences within a team such as variations in degree of agreement are accepted as indicators of team SA, important aspects to assess are *awareness of individual and team SA between the team-members*, as well as *each individual's awareness of the situation*.

To illustrate with an example (see Figure 11), the required awareness could be assessed by task analysis including cognitive task analysis, and the current awareness could be assessed by EPOG, forming an indication of the relative awareness.

At the same time an observer rating may be used as a complementary assessment of the relative awareness, through observations of the actual awareness. In an analysis the fused indication of relative awareness, assessed with more than one technique, could be compared with an assessment of the awareness of relative awareness obtained by, for instance, a self rating. Further, to be able to analyse degree of agreement as an indicator of team SA the awareness of individual and team SA between the team-members could, for instance, be assessed with self ratings and peer-ratings, and each individual's awareness of the situation assessed with a questionnaire of relevant aspects of the situation.

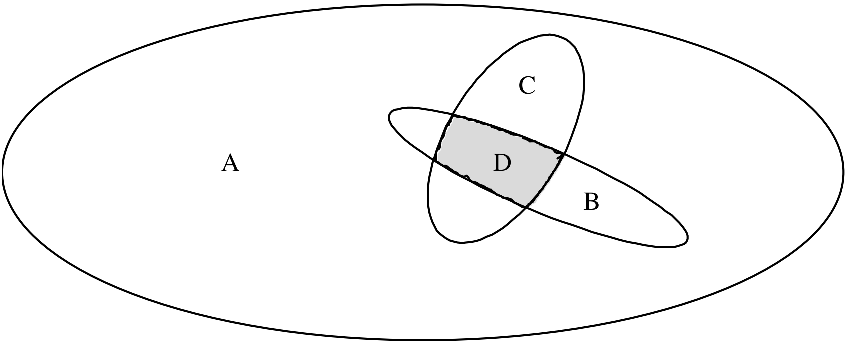

Figure 11. The figure illustrates how various aspects of SA could be assessed with one or more measurement techniques, to provide an indication of SA through an analysis of a set of example measurement techniques. The example measurement techniques could be exchanged with other techniques, depending on the purpose of the SA assessment given by the meta-situation.
A = *Situation*
The situation is assessed by the researcher from constraints in meta-situation, for instance, by applying a scenario to a simulation, and subjectively assessed by the persons SA, during the simulation.
B = *Awareness*
The awareness is assessed with EPOG.
C = *Required awareness*
The required awareness is assessed with task analysis including cognitive task analysis.
D = *Actual awareness*
The actual awareness is assessed by observer ratings. An observer assesses a user with respect to awareness vis-à-vis required awareness. Additionally, an overall estimation of the level of SA may be conducted by the observer in comparison with the calculated level of SA.
D/C = *Relative awareness*
The relative awareness is obtained, based on the differences between the EPOG and the expected behaviour from the task analysis as well as through the observer ratings.

When developing complex systems, the assessment of SA often aims at aiding design choices. Assessing, and analysing differences in SA due to design alternatives are attended in the following section.

3.2 Differences in Situational Awareness Due to Design Alternatives

This section describes how differences in SA due to design alternatives could be both qualitative as well as quantitative, in its nature. Further, methodological concerns regarding the evaluation process itself, and the evaluation process as part of the development process is described for the sake of relating SA aspects due to design alternatives to a broader applicative perspective.

When there are two or more candidate design solutions taking part of an evaluation, information about the assessed differences are interesting, both the quantitative and qualitative differences. The quantitative differences could be used for making choices between design alternatives, for instance, choosing a design alternative that promotes SA, rather than a design alternative that does not. The qualitative differences could be used for improving the design in an iterative design cycle, or to give input for new design solutions.

Paper 1, exemplifies on when differences in SA is due to design alternatives. In this case the colour setting of a display in a simulated fighter aircraft resulted in significant differences in SA for fighter pilots measured with CC-SART, Cognitive Compatibility Situation Awareness Rating Technique (Taylor 1995a; 1995b). The quantitative differences guided the design into a choice of colour coded displays, and the qualitative differences aided later iterations of the colour settings design by a simulator based design process. For further reading, results from the study including colour aspects, have been presented and reported earlier (Derefeldt et al., 1998a; 1998b)

In this study the effect of colour was positive, which is not always the case. However, the differences in SA due to design alternatives often provide valuable input regardless of the directions of the effects. For instance, Cummings and Tsonis (2006) found a negative effect (i.e. increased complexity) as an effect of too much use of colour on displays in an air traffic management task.

In both of the cases above the evaluations demonstrated potential to guide future design. To use assessment of SA for making design choices and to improve preliminary designs are traditional and common ways of making use of situation assessments.

Another example is, for instance, Walker, Stanton and Young (2006) that used various SA assessments to find out which modality to use for vehicle feedback in car design, or Ma and Kaber (2007) that used a version of SAGAT to find support for the use of in-vehicle navigation aiding during normal conditions for enhanced driver SA.

Yet another example is the development of a threat display for aircraft (Alfredson, Rosenblad, & Skinnars, 2002). When the principal design of the threat display was determined two experiments were conducted to validate some aspects of threat representation focusing on SA. The first experiment aimed at studying three candidate visual concepts for representing threats for aerial vehicles (Strömberg & Alfredson, 1997). The second experiment aimed at studying a novel visual concept for representing threats for aerial vehicles compared to a more traditional approach (Linde, Strömberg, Pettersson, Andersson, & Alfredson, 2001).

The differences in SA between design alternatives could be assessed in various manners. Comparing two or more alternatives in a controlled experimental setting often is a useful approach in a development context, since design is about making choices and an experimental setting aids such a choice in that it provides a structured framework for the evaluation.

However, a strictly experimental approach, comparing two or more design alternatives to find out a winning candidate to bring further also has drawbacks to the process of developing complex systems. Dekker and Nyce (2004) point to the problem that experimental steps into the future are narrow, risking low generalizability, and that the mapping between experiment and future reality may miss important factors. However, a similar problem is the problem of generalizing from a case study.

To avoid these problems it is beneficial to have a broader understanding of the methodology used and of the applicable theory, in this case SA. Artman (1999) demonstrated benefits on understanding SA by the use of both experiments and field observations as complementary methods. The methodological understanding is important to correctly assess the differences in SA, and the understanding of SA theory to make valid conclusions and considerations to the development process. Having this in mind, SA could complement other evaluations to guide design in a couple of ways.

Firstly, traditional design approaches in HF emphasise task analysis; methods adapted to clearly-defined tasks divided into a series of steps, but less suited to dynamic tasks where goals can quickly shift and even appear conflicting. The SA model takes shifting and conflicting goals into account.

Secondly, SA provides a framework for evaluation of interaction between different system components where certain elements can be highlighted as more important. Most information processing design guidelines address issues at the level of a single component, while the real challenge for designers is how to find an optimal integration of all components.

This integration can very well be evaluated in terms of improved SA. Thus, SA offers a normative model for analysis of trade-offs that may confront a designer.

Also, SA assessment could be performed over time, giving a feel for long term trends, for instance, technological trends or contextual change that affect SA. When comparing SA assessments made at various points in time it is highly beneficial to use the same assessment methodology at all the occasions to ensure valid comparisons. Also, old technical solutions could be compared to new. For instance, Walker, Stanton, and Young (2007) found that drivers of 15 years older cars with higher levels of vehicle feedback had better SA than drivers of modern cars.

One of the major reasons for using the concept of SA in the development of complex systems is pragmatic: Applying the concept to the design might help creating good design. Since SA is used to evaluate human-machine interfaces, the concept should enhance the design process, and increase the benefits of the result of the improved design process. The concept of SA can be used for at least three purposes in the design process:

- Enhancing the design process, by providing rationale sub-goals created from an understanding of SA.
- Enhancing the actual design by helping the designer to understand the design criteria, and there relative importance.
- Ensuring that the design is satisfying, by testing the design in relation to SA.

Design recommendations, based on Endsley's three level model of SA (Endsley, 1988) have been presented to help designers to make a better design (Endsley, 1995b; Endsley, Bolté, & Jones, 2003; Endsley, Bolstad, Jones, & Riley, 2003). Jones, Endsley, Bolstad and Estes (2004) even developed a designer's situation awareness toolkit (DeSAT) to provide support to the designer through tutorials and application specific tools.

SA could very well enhance both the design process as well as the design itself. However, by broadening the use of the SA concept within the field of design it might be possible to find even better applications to enhance the design and the design process, for instance, making the design more sensitive to the characteristics of a specific context, and explicitly regarding the differences in SA due to roles in the design-use process.

3.3 Differences in Situational Awareness Due to Roles in the Design-Use Process

This section describes how SA can differ between roles in the design-use process. After an example of roles in the design role process and a discussion of the role of design in the process, this section addresses the topic of function allocation, due to its effects on SA and its relevance for development of complex systems. A need for function allocation is described both at the level of agents in a JCS, but also at a higher level, between roles in the design-use process (see Table 2). In both cases, the function allocation could be conducted, as a static part in advance, as well as dynamically during the events or process.

There could be differences in SA between individuals or groups having various roles in the design-use process. The differences could be both on what the concept of SA really is, and also differences on the notion of a specific person's, team's, or role's (such as a group of users) SA.

Table 2

The Roles of Function Allocation Affecting SA in Development of Complex Systems

	Static	vs. Dynamic
Between agents	Static function allocation between agents in a JCS	Dynamic function allocation between agents in a JCS
vs.		
Between roles	Static function allocation between roles in the design-use process	Dynamic function allocation between roles in the design-use process

Paper 2 reports on differences in the notion of what SA really is, measured by a questionnaire among British, military pilots and Swedish private pilots related to descriptions of the concept in research literature. A designer who receives information about what SA is from the research literature, direct or indirect, would thereby risk a different view in contrast to the pilot's view. Another example of inconsistency between roles in the design-use process is presented in Paper 3, where major differences between the actual behaviour and the behaviour described in procedures measured in a simulated commercial aircraft setting is present. If a designer got the impression that, pilots behave according to what is written in the procedures, the designer would risk forming inaccurate mental models with the risk of misjudging the SA of pilots. Goteman (2006), found differences between the descriptions used for design and the real situations when studying the introduction of the novel techniques, head up display (HUD) and RNAV, in a commercial aviation context.

There are many roles in the design-use process. One possible relationship between roles in a fictive example of a design-use process would be that one researcher is developing design methodology, another develops SA theory and assessment methodology and a third is perhaps evaluating end user behaviour. There might be several designers adopting the research findings, in their work after that they, for instance, have red a book or attended a course. The designers are then interacting with managers, marketing people, production people and more, trying to find an acceptable design solution. Even though the designers are working at different companies and with different subsystems, it might be that their design efforts are consumed by several end users, trying to manage a current task, in their organisation. During the life-cycle of a complex system various types of maintenance personal will also interact with and modify the system, and they in turn will utilize user information as part of the support system (Candell, 2004).

All these people might have too little time to optimally perform their tasks, and are all trying to obtain an acceptable result of their efforts. For instance, many decisions are to be made under extreme time pressure, in a hyper dynamic setting with high stakes, which all are characteristics of a typical context, where naturalistic decision making is applicable (Klein, Orasanu, Calderwood, & Zsambok, 1993). Although this fictive example is an extreme simplification of a real equivalent, it yet gives a sense of the multi-facetted situation of the chain of a development process.

To achieve good SA for a single end user or team, there has to be valuable contributions from many parts of this chain. A researcher might contribute with design-oriented research, and a developer with research-oriented design (Albinsson, 2004; Fällman, 2003). It is not enough for an end user to have perfect individual properties, to form good SA, there has to be assistance from other roles in the design-use process. For instance, the design will affect the SA. Attempts have been made to influence the SA by providing design principles (i.e. principles of design for SA) to be used by designers (Endsley, Bolté, & Jones, 2003).

It is not only design of hard artefacts, but also soft concepts that potentially affects a user's SA at the end, such as rules, culture, organisation or language. For instance, concepts and metaphors used may be adapted to better assist the design process. Löwgren (1993) explains:

> Metaphors actually say too much and too little at the same time. They say too much because they can activate too much background knowledge in the user's mind... ...At the same time, metaphors of this type say too little because they do not help the user find the services that are particular to the computer system. (p. 71)

How explicit and detailed should the metaphors be? Perhaps, the technical evolution will bring us useful metaphors after a long period of elaborating with a variety of suggestions. The risk is however, that in the meantime, the metaphors will be too powerful for a user or even a designer to be able to design optimally. Papantonopoulos (2004) describe how metaphors for a function allocation design task “have a formative impact on design thinking by shaping the designers’ conceptualization of the tasks to be allocated, the operating context, and the human cognition” (p. 1542). For instance, subjects used computer metaphors for human cognition, resulting in a biased understanding of the human cognition as information processing activities.

Together, the user and a technical system form a JCS. The interdependence between the two parts of the JCS, to be able to fulfil the tasks of the JCS, is often prominent. By an accurate interaction between the two parts of the JCS, information can be exchanged in order to make the JCS act as one unit, with the desired behaviour.

However, also the designer could be regarded as a part of the JCS, even though the interaction from the designer (i.e. the design) is located earlier in time, than the user’s interaction. For instance, Riley, Endsley, Bolstad and Cuevas (2006) described system design concepts for collaborative planning in military command and control, aiming at enhancing user SA. Reason (1988) addresses the role of design regarding errors by explaining how passive failures, such as errors in design, create latent conditions, or *resident pathogens*. The latent failures are therefore failures that are made far from the point in time and space where it occurs, so it is not the person who is handling the situation at the time of the incident/accident that are the main reason, or at least not the only reason, to the incident/accident.

There are many persons that are involved with complex systems during a long time, and therefore it is vary hard, if even possible, to link the accident/incident to only one person’s behaviour during a specific moment. Yet this is often done, even if it not always constructive.

Determining who will do what in the JCS is one constructive effort to reduce the risk of misunderstandings between the agents. It is important to have determined a strategy for function allocation, to avoid having different sub-systems using different strategies. With a good strategy for function allocation the resources, human as well as technical, are used efficiently. Already Fitts (1951) used principles to allocate functions between man and machine. Some tasks are better suited for humans to perform, while some tasks are better suited for technical systems. Several lists have been constructed that separates what humans are better at and what computers do better, which could be useful for designers of systems to use for guidance. This issue was brought up by Dreyfus (1979).

However, good strategies for function allocation also takes into account the actual interaction aspects and not only divisions of functions. Hollnagel (1999) shifts the focus from substitution to cooperation when elaborating the term function congruence. If synergies between human and system components can be found, the efficiency of the total system becomes better, which have been discussed by Alfredson, Derefeldt, and Jungert (2004).

The function allocation strategy does not have to be statically fixed, even though the overall framework has to be decided. Combining theoretical knowledge with rapid technical evolution may create very powerful adaptive systems in the future. Adaptive function allocation has been tested based on operator performance and mission relevant variables (Scallen & Hancock, 2001). Improved route navigation was found with an adaptive interface compared to a traditional interface (Bennett et al, 2001). Performance enhancements for adaptive aiding during high mental workload have been reported (Wilson, Lambert, & Russell, 2000).

It is not only the function allocation between the agents in the real-time environment that has to be performed, but also the distribution of tasks between the roles in the design-use process has to be function allocated. Otherwise, the risk is that tasks will not be performed, or be badly performed, which may affect the SA at the sharp end.

For instance, whichever strategy for function allocation is selected, it is far from obvious that all designers use it. Papantonopoulos (2004) found some variations between strategies for the design of cognitive task allocation between ten studied system designers. In the same study, a common design strategy for task allocation was found: To apply a selected automation technology before or concurrent to task analysis, which is not the sequence that would be found in typical design literature (Papantonopoulos, 2004). A person reading the design literature would certainly risk getting the wrong impression of how designers work compared to the reported empirical findings. Just as there are differences between the descriptions and the reality for end users, as reported in Paper 2, and Paper 3, there are differences between descriptions and reality for designers.

To be able to understand differences in SA due to roles in the design-use process, we have to understand both the described tasks/views of each role (to understand how that role is regarded by a person in another role) as well as having an idea of the actual tasks/views of the corresponding roles, similar to the SA of a person that is dependent on what to be aware of and actually being aware of it, as described earlier in this dissertation (cf. p. 28). An alternative approach would be to assess each role's perception of the design role process, to investigate if there are differences in that respect.

The function allocation between the roles in the design-use process do not have to be static, but could be allocated dynamically as well as in the case of dynamic function allocation on the lower level of the real-time environment. For instance, the user could perform some design efforts at times.

In E-commerce the roles could be described by three subsystems, store environment, customer, and web technology (Helander 2000; Helander & Khalid, 2000). Khalid (1999) explains that the Do-It Yourself Design (DIYD) concept has been associated with self-assembly products, but states that in many cases the roles are reversed: "A customer designs while the company assembles. Hence, the customer, as the designer needs to understand his/her own motivation for design. Giving customers opportunities to express their needs directly via a Web-based system may enhance customer satisfaction" (p. 344). Who knows better than the user what the user wants? A reason for using DIYD is that some researchers claim that design ability is possessed by everyone (Cross, 1995).

One possible solution to the designers' dilemma of having too much to do in too little time is then to delegate some of the design decisions to the user in an interactive environment. Another possible solution is to ask the user and implement the explicit user needs.

However, these possible solutions, cutting off the designer from the design loop, have the drawback of not taking benefit of the designers care about what user remarks mean and put it in a future work context, as pointed out by Dekker and Nyce (2004). Smallman and StJohn (2005) described a case of naïve realism, where marine officers prefer analogue (photorealistic) symbols of marine vessels instead of artificial symbols, although experimental results demonstrated better performance in recognition test (item identification), of abstract symbols. Also Alm (2001), demonstrated a similar effect with superior threat identification using artificial aircraft representations compared to analogue representations.

Woods (1995) explains how representation design could become the practice of designing marks in a medium to accurately portray relevant properties in the world. Stressing what is relevant rather than what is visually present in a situation is one practical consideration that is worth stressing when developing complex systems (Albinsson & Alfredson, 2002; Howard, 2002).

The role of the designer is important and valuable in the development of complex systems. Even if an end user could express a need, such as the need of knowing the directions to other objects, it is the task of the designer to create a design solution, whether it is a visual display (e.g. Alfredson & Romare, 2002), a combination of visual and tactile displays (e.g. Eriksson et al., 2006) or tactile and 3D-audio displays (e.g. Carlander & Eriksson, 2006).

It is a simple solution, to let the user make all the choices, but is it the best solution? Does the user really want to make all the choices? Maybe there are too many choices? Do you really want to make the recipe yourself when you go to a restaurant? Is it not better if the professional food-designers in the kitchen make some of the choices? It may also be that the efforts of making the choices are too heavy for the user's preferences.

All choices take mental effort to do. It is possible that we will have to develop a better understanding of the cost and benefit of choice, to be able to find the optimal level of choice. Even a trained designer does not possess all the HF related knowledge that may be beneficial, how is it then possible to expect the user to know that? Maybe the user knows what he/she wants, on a low level, but not on a higher level. Designing for the user demands design knowledge and especially design knowledge regarding HF. By just transferring design work to the user, the risk is that the user could not perform the design due to lack of competence or lack of time to make the effort, and the risk is also that other tasks of the user become worse performed due to the added workload imposed by added design tasks.

Also, performance and team SA will risk negative effects in an unrestricted environment with a lot of options. In a micro world setting with geographically distributed decision-makers acting as a team to control a dynamic system Brehmer and Svenmarck (1995) demonstrated better performance for decision-makers in a restricted hierarchical information structure than for those who, were fully connected so that they all could communicate with each other. Also, in a study of military command and control at battalion level Artman (2000) demonstrated how a technological implementation with fewer restrictions on information spreading seams to increase the efforts needed for the team to establish team SA.

Personalization of functions and user interfaces is no guarantee for implementing true user requirements. Sometimes the effectiveness of an organization will become superordinated personal preferences and individual attitudes. Additionally, personal preferences are not always considering optimal design solutions from an information presentation point of view. Interview data are not always reflecting users' desires, but could well be influenced by commercials and technical capabilities. Sometimes it is a large discrepancy between what people are claming they are doing and what they actually are doing. Therefore, the outcome of a feasibility analysis provides measures of people's opinions, but does not necessarily reflect their future behaviour.

In many modern contexts the development has not ended when the product leaves the factory where it has been produced. A complementary approach to letting the user make the final design is to build in capability into the technical systems to assess the situation and adapt its behaviour accordingly. Fredriksson (2004) presents a method and approach toward establishment of open computational systems called online engineering, that focuses on the continuous refinement of behaviour and qualities, according to temporary concerns and individual aspects of cognitive agent, both human and machines. The process of creating SA is present in the running system, as well as in the preceding technical development process. To be able to aid SA the roles in the design process have to complement each other and therefore have to, at some level, share an understanding of each other's roles and contributions to this process.

As reported in Paper 2, the notions of SA that pilots have is not the same as the notions of SA used in the literature. Perhaps we could expect yet another notion among the design community? Adding to the problem is that the designers might not be provided with the time that they need to regard SA even if there would be consensus on what SA is (which has not been shown). For instance, Willén (1995) found through interviews of 14 Swedish engineers that most of them did not even use 5% of their time on user aspects.

Users are often considered as one of the subsystems of the global device by designer, and occasionally modelled through basic design principles or elements of an imagined scenario, dependent on what type of meetings are held through the design process (Darses & Wolff, 2006). Not only the type of meetings, but also the design perspectives adopted by the designers in the design work affects the design outcome (Hult, Irestig, & Lundberg, 2006).

The designers' contexts shape their understanding of their situation and thereby their SA. Although the end user is at the sharp end, designers and researchers are at the sharp end too, just another end. Who is not? What human being is not constrained by reality? SA is a matter for everybody. We have to manage our situation, whatever the situational context, or meta-situation, may be. However, by assessing SA in various contexts the development process of complex systems will be provided with valuable input.

3.4 Differences in Situational Awareness Due to Context

This section provides a description on how context could induce differences in SA. Especially the contextual aspect of complexity is focused, due to its effects on SA and its relevance for development of complex systems. Also, the topic of how context is presented to the user is addressed.

SA strongly depends on characteristics of the experienced situation and not, for instance, only on individual differences. SA therefore is strongly context dependent. Different contexts will lead to different SA for the person experiencing the situation as well as for anybody trying to assess the SA.

Paper 2 reports on differences in SA due to context by providing examples of when SA among English military pilots and Swedish private pilots is experienced as situation dependent. Paper 4 reports on differences in SA due to context by demonstrating that complexity and separated decision makers reduce shared SA measured by degree of agreement among decision makers in two command and control settings. In this study complexity was operationalised as events that contained uncertainty and risk for changing plans.

However, this is only one aspect of complexity. For instance, Cummings and Tsonis (2006) proposed partitioning complexity into *environmental*, *organizational*, and *display* for an air traffic management task and Xiao, Hunter, Mackenzie, Jefferies, and Horst (1996) found four components of task complexity in an emergency medical care setting: *multiple and concurrent tasks, uncertainty, changing plans,* and *compressed work procedures and high workload*. Complexity, in turn, is only one aspect on context. However, Paper 4 demonstrates that the studied parts of the context (i.e. context and decision maker separation) affects shared SA. Complexity is a central aspect of context not only because it affects SA, but also since complexity is an important aspect to regard when developing complex systems. Complexity is built into complex systems and complex systems tend to be used in complex contexts. Due to the technical evolution the number of complex systems are increasing and the complexity of the systems are increasing as well. Due to the modern trend of increasing number of complex systems together with the increasing complexity of modern systems, a need for refinement in distinction and descriptions of systems has grown.

Holistic views of large systems or even of systems of systems do not have a single determination. Hyper complex systems often compose multiple levels of abstractions, making the system view all more fruitful both for engineering the systems, but also as a mental model for the user of the systems. Also systems of systems are constructed in the same trend.

There are special concerns for HF issues in complex systems, and systems of systems since the user often have both important and demanding tasks. For instance, efforts considering HF are undertaken in complex defence systems development, for instance, by the US-army (MANPRINT, 2000), NATO (Geddie et al., 2001), and the European cooperation GARTEUR (Castor et al., 2003).

With increasing complexity, we will also be exposed to new and more devastating threats. With increasing complexity, increasingly vulnerability follows. The vulnerability will be enhanced due to the increasing amount of possible breakdowns of a large, complex system. Also, the vulnerability will be enhanced due to the increasing number of interdependences in a complex system. Not only will the likelihood of breakdown increase, but also the negative consequences of a breakdown will be more severe in a large complex system. As a further complication the breakdown of a complex system often is more difficult to diagnose to be able to establish a suitable prevention strategy for avoidance of future breakdowns.

Analysing shared SA between agents in a complex context is one approach to understand that process. In both civilian and military contexts it is important to strive for shared SA in order to create efficient organisations. The mutual sharing of information is desired, but not sufficient for creating a common understanding of complex tasks. Hyper dynamic and complex systems are impossible for a single individual to control. Only by having some level of shared understanding of the system, the task and the context, the individuals interacting with the system may understand each other and act together towards a common goal.

Differences in SA due to context could also be induced dependent on how the context is perceived by the person, and thereby also on how the context is presented to the person. In many modern systems the user does not interact directly with the reality, or even directly with a virtual reality, but through various conceptual layers provided by technology that affects SA. For instance, pilots, ship officers and air traffic controllers prefer intermediate levels of automation, and in these settings there are concerns that improper technological development will risk that the user looses the directness to the controlled system (Dutarte, 2001). Distancing the user from the controlled reality will risk bias SA through perception of the reality through technology. Adding to the problem is that various subsystems may not represent the reality consistently. Garbis and Artman (2004) found that different systems, such as, electric supply system, mimic panel, and signal system, in an underground line control room represented the status of the line in different manners, which made it more difficult to form team SA. This was recognised as a problem, especially when an undesirable situation occurred, since the systems were designed to support normal operations. Also, the human to human communication may be negatively affected if technology distancing users to each other. For instance, Johansson (2005) reported on command and control settings with the risk of getting information technology filtering away important features of communication, such as spoken interaction.

Also, a context is never exactly the same from one time to another and nor between individuals. This will complicate the generalization of results and the judgement of when an assessed difference in SA due to context is significant. It will also be affected by the level of analysis.

3.5 Differences in Situational Awareness Due to Levels of Analysis

This section addresses that level of analysis will affect the assessed differences in SA. Since the earlier section on assessing differences in SA described the general aspects of SA assessment, the section below will only focus on the effect of level of analysis, due to its effects on development of complex systems.

If you could conduct extremely fine grained analysis you would detect differences in SA between any conditions, even if the differences were not significant from the perspective of design and development. If, however, you would conduct a very low resolution analysis you run the risk of not detecting important differences in SA. Paper 3 reports on how commercial pilots in a simulated typical flight expose similar behaviour, but a lower level of analyses is revealing a wide range of behaviours, possibly indicating differences in quantity or quality of SA. Paper 5 reports on how pilots on one level of analyses expose the same behaviour, but at a lower level of analysis is diversified, with data from two experiments.

Also, the resolution in the time dimension is important for useful analysis of differences in SA. For instance, Patrick, James, and Ahmed (2007) found, in a study of awareness of control room teams that SA scores hardly could be aggregated within and between situations, since SA is not only task specific, but also situation specific, including the characteristics of the team. Therefore, a low-frequent time sampling of SA assessment would risk missing important fluctuations of SA, quantitative or qualitative. A high-frequent time sampling or even a continuous dynamic measure of SA would be theoretically preferred, even if many SA assessment techniques in practise would risk becoming intrusive or inconvenient to use if SA would be assessed too frequently.

However, a dynamic measure of SA would not show its full potential if corresponding measures used in the same analysis, such as, performance or workload, would not be available for the same detailed level of analysis. Also, the difficulties of aggregating qualitative and quantitative aspects of dynamically assessed SA into meaningful chunks, for the level of analysis of the research question, would still remain. For instance, a question of how to design a feature of a complex system may not be answered by examining any specific point in time with high-frequent resolution of SA, but rather by the study of interaction patterns appearing from time to time where SA assessment and analysis could contribute to the design decision.

Lindh (2006) explains how relativistic and uncertainty phenomena, similar to those of modern physics, makes it impossible to precisely assess value metrics associated with high level conceptual models of command and control, based on observations of interaction in open distributed systems. Since this empirical data is not on SA but on lower level of communication between distributed asynchronous cognitive systems, we could only speculate on the effects on SA. It might be that the effects are so small that they are practically neglectable. However, since the theoretical models of communication between cognitive agents are similar on both levels and since SA are influenced by the local, subjective influence of a cognitive agent (such as a person), it would probably be wise to regard these factors, as hypotheses, until the effects have been determined. For instance, two persons interacting together at the same time, are not necessary able to achieve the same SA, even theoretically, and even less in practise. Further, a person's SA could, in that same line of thinking, never be exactly predicted and described in advance, and therefore it would be theoretically impossible to determine the SA-demands in detail, for instance, for the sake of assessing SA. In fact, it would not be possible to give an *exact* value of SA, with *total confidence*. In a wider perspective this would mean that the concept of SA would be hard to validate. For practical purposes however, such as managing differences in SA in the development of complex systems, the ill-defined, ill-structured and ill-validated concept of SA still is highly valuable.

4 Managing Differences in Situational Awareness in Development of Complex Systems

This chapter is about how to manage differences in SA, in the development of complex systems. The differences in SA contain information. If you are able to manage the differences you can use that information in the development of complex systems.

Some of the differences in SA due to design alternatives could, for instance, be managed by using simulator based design. Some of the differences due to roles in the design-use process could, for instance, be managed by the broader SM approach (cf. p. 45). Since SM includes both presentation and manoeuvring support, and is therefore a wider and more applicable term to interface design than SA, that is focused on the detection, understanding and prediction of surrounding elements. Some of the differences due to context could, for instance, be managed by use of on-line evaluation of SA. Some of the differences due to level of analysis could, for instance, be managed by measuring and analysing SA on multiple levels simultaneously.

If we have more than one assessment of SA, it often is interesting to compare the assessments. One assessment may vary from another due to reliability issues. However, the differences also often contain valuable information for the process of developing complex systems.

Simulator-based design is one of the means to manage differences in SA in the development of complex systems. Paper 6 describes the benefits of simulator-based HMI design in the development of complex systems, specifically vehicle systems, which has also been explained by Alm (2007) as well as by Alm, Alfredson, and Ohlsson (2007). Singer (2002) reported on benefits using part-task simulation early, when designing commercial aircraft cockpit systems, such as improvement of warning systems and warning message design, as well as for head-up display implementation. The differences due to design alternatives then can be used in further development of the system.

The study in Paper 1 is an example of how simulator based design can be used to link empirical knowledge in simulations regarding differences in SA into the development of complex systems, in this case a fighter aircraft, as described in Paper 6. In this case it was important that the colour schemes were specified in a standardized colour appearance system, so that the results from the simulations could be fed into the product design process. Also, it was important that the fidelity of the simulator was suitable, for simulation based design to be successfully used. In general, it is important that evaluations are designed so that the development process could benefit from results and lessons learned efficiently.

Here, iterative development is needed, since we lack detailed descriptions of a clearly defined end product, and lack knowledge about user behaviour, in advance. With an iterative development approach, we are able to find our way forward, as our understanding of the developing task is increasing. We can use gained knowledge in early development cycles to form an understanding that helps us direct the efforts in later development cycles. To do this in a structured manner we need to describe the process in terms of its components, including the human aspects. Using simulations or not, the development process could benefit from structural efforts on ensuring validity cost effectively. In light of this, referring to the terms in Figure 3, the *required awareness* is central to describe to achieve a valid situation, regarding SA, regardless of, for instance, the simulator fidelity, as argued in Paper 6.

Earlier in this dissertation SM has been described as focusing on the situation, rather than on a person or team experiencing the situation (cf. p. 46). By using the broader SM approach contributions to the situation could be regarded from any role in the design-use process. By analysing these contributions and potential alternatives, differences in SA could be detected and could be managed for the development of complex systems.

SA is a well suited concept to analyse what happens on one level, for instance with a person or a team, but less suited for analysis across a hierarchy of situations/meta-situations. For instance, it is not desirable for high level commanders to have the same SA as lower level operators performing a task, so it is hard to compare the quality and quantity of SA at various levels with each other.

In an analysis two complementary aspects could be fused to one approach, the aspect of having the focus on the situation, and what contributes to the situation, and the aspect of analysing SA on various levels simultaneously. In such an approach, differences at the level of an individual could be regarded together with differences at the level of a team, as well as at the level of roles in the design-use process as visualised in Figures 12, 13 and 14.

The differences at the level of an individual could be required awareness versus awareness (see Figure 3), and the relative awareness versus awareness of relative awareness (see Table 1) to assess the individual SA.

The differences at the level of a team could be assessment of the situation by the individual versus assessment of the situation by a team member (situation-oriented degree of agreement), assessed individual SA by the individual versus assessed individual SA by a team member (team-oriented degree of agreement), or team SA assessed by the individual versus team SA assessed by a team member (also, team-oriented degree of agreement), to assess team SA.

The differences at the level of roles in the design-use process could be the design-use process as it is perceived by one role versus the design-use process as it is perceived by another role (process-oriented degree of agreement), or the role as it is actually manifested in reality versus the role as it is perceived by an other role (role-oriented degree of agreement), to assess common understanding or SA between the roles in the design-use process. Extrapolating this approach, differences between agents on any level could be analysed, such as, at the level of SA for systems of systems or why not intergalactic SA, dependent on the question at hand.

Similar to an architect who is designing a building, an organisation could be designed as well (Simons, 2005). Also the design-use process can in itself be designed. The technological design (e.g. component, relations, etc.) and the institutional design (e.g. responsibilities, allocation of costs, etc.) can both be formed by the higher level process design (i.e. designing the design process) (Koppenjan & Groenewegen, 2005).

Alfredson, Oskarsson, Castor, and Svensson (2003) developed a meta instrument for evaluation of man-system interaction in systems engineering to support the process of defence material acquisition, for the support of FMV (Swedish Defence Material Administration) administrators. The meta instrument included a structured view of methods, including SA assessment methods, to assist the administrators in their dialogue with, for instance, industry developing complex military systems (Alfredson, Oskarsson, Castor, & Svensson, 2004). If the design process is properly designed to manage situations future users of products from the design process could indirectly benefit from this through increased SA.

Managing differences in SA in the development of complex systems are increasingly important since there is an increase in complexity of large socio-technical systems of today. Methods and techniques for designing and developing large socio-technical systems are increasingly important to produce efficient and usable solutions. In later years, attempts have been made to address this matter, through a variety of approaches, such as user-centred design, usability engineering, ecological interface design and many others.

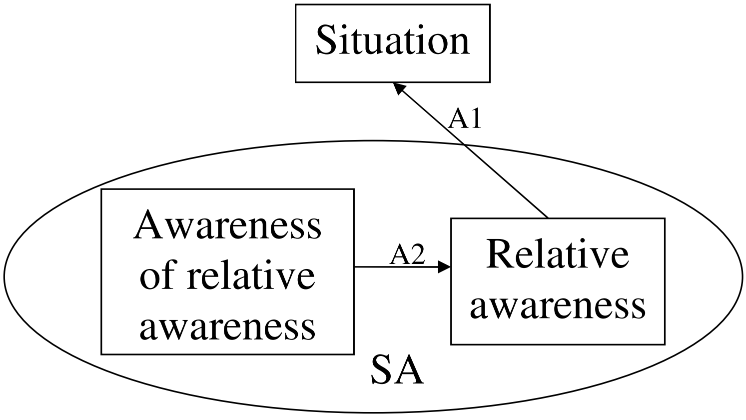

Figure 12. Differences at the level of an individual. High SA is when someone is fully aware of a high relative awareness. The arrows are representing the assessments potentially resulting in differences. At A1 the required awareness of the situation is compared to the awareness of the individual forming a relative awareness. At A2 the relative awareness is compared to the awareness of relative awareness of the individual forming the SA.

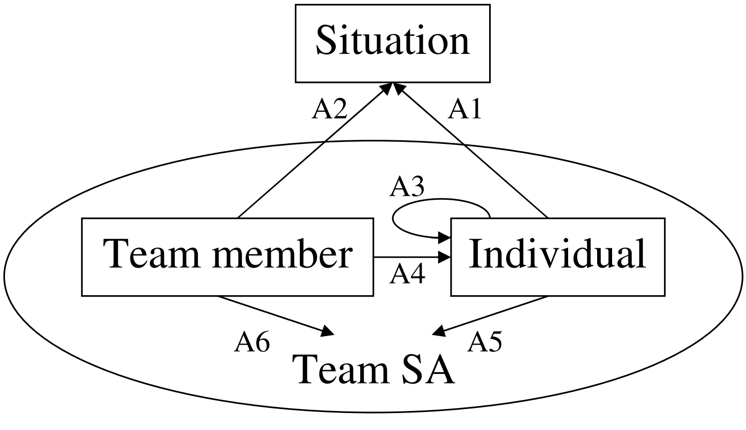

Figure 13. Differences at the level of a team. High team SA is when the team members are fully aware of each other's high individual SA. The arrows are representing the assessments potentially resulting in differences. At A1 and A2 the individual and another team member both assess the situation forming a situation-oriented degree of agreement as one measure of team SA. At A3 and A4 the individual and another team member both assess the individual SA of the individual forming one team-oriented degree of agreement as another measure of team SA. At A5 and A6 the individual and another team member both assess the team SA forming another team-oriented degree of agreement as yet another measure of team SA.

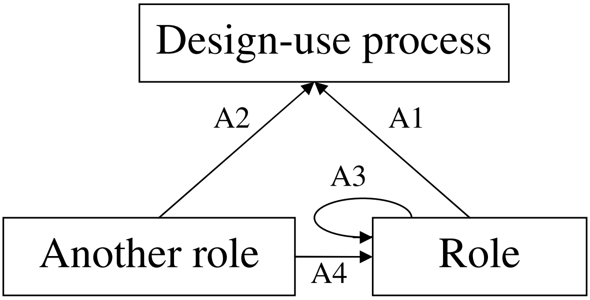

Figure 14. Differences at the level of roles in the design-use process. Even though it would be possible to do so, SA is not traditionally used to describe awareness between roles in the design-use process, but perhaps rather, common understanding. The arrows are representing the assessments potentially resulting in differences. At A1 and A2 the role and another role both perceive the design-use process forming process-oriented degree of agreement as a common understanding or SA. At A3 and A4 the role as it is actually manifested in reality is compared to the perception of that role by the other role, forming role-oriented degree of agreement as a common understanding or SA.

The role of the usability designer is important for successful systems development, and it is important that the actual design solutions are directly influenced by the usability designer (Boivie, Gulliksen, & Göransson, 2006; Gulliksen, Boivie, & Göransson, 2006). Usable systems can support user SA. Depending on how a situation is represented, different means for interpretations and solutions emerge. Novel techniques supplies new opportunities to knowledge representation.

Systems of systems are integrated into powerful networks and services available in the network are combined in real-time when needed. Powerful opportunities that never before were possible are created. Among them new opportunities to do powerful mistakes also emerge! To manage the new space of opportunities generated we need top-down conceptual design of the knowledge representation, creating preconditions for good system awareness (Alfredson, Derefeldt, & Jungert, 2004).

Severe consequences in distributed, safety-critical human activities often arise from problematic interaction between humans and systems. The ongoing development towards network-based defence is a typical example of a process dealing with complex socio-technical systems where misunderstandings and mishaps can lead to severe consequences. Decision makers here have to make decisions in situations where operational and system characteristics are highly dynamic and nonlinear, and the decisions affect real-time systems properties at several levels simultaneously such as individual operators, technical systems, and combinations of operators and systems of systems, which demand a multiple theory approach to system development (Worm, 2000; 2002). That is, the new context of complex systems of the future, demands new methods. In the search for new methods a practical approach is combining benefits from existing methods, developing them, and complementing them. After performing a review of SA measurement techniques, Salmon, Stanton, Walker, and Green (2006), recommended a multiple-measure approach for SA assessment in command, control, communication, computers and intelligence (C4i) environments.

One of the intended benefits of a flexible system is that it would be possible to deal with unforeseen events. To achieve this, a technical part of the system is expected, that can provide suitable solutions, for instance, services and new service configurations. Since flexibility is central the technical support has to be general (Wallenius, 2005).

Another part that is needed in the system is human decision makers with power of initiative. Both these ingredients are needed to fulfil the potential flexibility aimed for, and therefore both aspects ought to be regarded when developing service management tools for systems of systems (Alfredson, Lindoff, & Rencrantz, 2004). If something fails or work suboptimal, the reason for that could be found either in the technical part or in the human part or in the interaction between the two. Somewhat simplified the humans and the technical parts, are all part of the same system. Therefore, it is necessary from a development perspective to regard both of these parts of the total system simultaneously. The consequence of this is that the human contribution has to be regarded together with the technical. This is why it is necessary to include knowledge about human decision making into an iterative development process with a holistic approach to evaluations of the future system.

Of special interest is also the fact that the flexibility will be obtained at the cost of distributed control of validating to a real-time situation with a human decision maker and a technical implementation, rather that a predefined and controlled process. Giving the decision maker free hands, ties the hands of the predefined validation, at the same time. Since we do not know what will be performed in a fully flexible system, we neither know what will happen from a validation perspective.

However, this could be handled with regulations on the envelope of possible actions. It could also be handled with on-line efforts on, for instance, status surveillance and predictions of intended actions. Also, the methodology for performing validation for the predefined and controlled process could be adjusted. On-line evaluation of SA could be used for validation purposes. The system could thereby cope with differences due to a changing context. Short loops in the iterative evaluation cycle are needed to assess the situation fast. Extremely fast response can only be obtained if on-line assessment and evaluation is used.

One step towards on-line assessment of SA is to distribute questionnaires electronically instead of paper versions, since the electronic versions could get integrated into the system that the person is interacting with. Svensson, Rencrantz, Lindoff, Berggren, and Norlander (2006) demonstrated how a pocket PC could be used for questions, among other questions, also regarding SA, in a command and control environment. However, the use of electronic questionnaires should be further studied for validation purposes since, for instance, the paper version of NASA-TLX (Hart & Staveland, 1988) has been found to add less workload than a computerised version (Noyes & Bruneau, 2007).

Many systems become more and more complex, which increase the demands on the human handling the systems. Much of the automation efforts have been aimed at increasing performance of the technical systems. However, the focus has mainly been externally directed towards the performance of sub-systems and technological networks, and not so much on automation for the purpose of aiding the human handling the system. Multi-sensor systems on modern platforms effectively scanning the surrounding environment for data have been developed together with advanced computational systems for extracting information and distributing it over networks, connecting technical sub-systems.

Two major decision support programmes that have been reviewed by Svenmarck and Dekker (2003) are the US Pilot's Associate and the French Copilote Electronique, that both, at least partially, succeeded in developing a decision support system suited for military aircraft.

However, what current complex systems lack is an equivalent system for keeping track of what happens with the user, that is, information about the status of the user. With sensors directed towards the user, an adaptive system might be created that focuses on the human and the needs and current limitations of the user's handling abilities. Such a system may use implicit information from the user. Victor (2005) demonstrates how information about the user's visual behaviour could be automated into a driving environment to create in-vehicle attention support functions to deal with effects of inattention. Alfredson & Nählinder (2002) describes advantages and disadvantages of three eye-measuring techniques in the light of a future adaptive aircraft system.

Also, explicit information from the user, such as user needs expressed by the user, may be integrated into an adaptive system. The more influence the user's explicit needs have on the design, the closer we come to the situation were the user is designing.

When managing differences in SA in the development of complex systems it is central to regard that empirical results, however valid, comes from the past and design is about the future. It is important to get a good description about the future context, for instance, by means of task analysis, including cognitive task analysis, and various user inputs. Singer (2002) demonstrated the use of including final users early in the development process when studying methods for validating cockpit design. A recent example of when collecting users' opinions early in the design phase could enhance the situational awareness of a future complex system by regarding this in a development process is Barchéus and Mårtensson (2006) that reported on the views of air traffic controllers on future technology in air traffic management.

However, it is impossible to foresee all consequences of a suggested design of a future system. This calls for an iterative development process. Only if we can test ideas and implementations in an iterative process, will we be able to avoid the pitfalls of bad human-system interaction. We need to use state-of-the-art methods, simulations and scenarios explicitly suited for iterative development of human-system interaction.

It is not possible to develop a system guaranteeing to generate SA, since the SA is context dependent, as reported in Paper 4. When developing a system, however, it is important to try to estimate a future context in which the system are likely to be used. When designing a future system we will encounter the envisioned world problem, where we as developers have to envision a world that does not yet exist, and try to figure out what is important there (Dekker, 1996; Dekker & Woods, 1999; Woods & Dekker, 2001). Often developers tend to focus on the technological advances and do not put as much effort into understanding the new conditions that humans in the envisioned world will experience. When some technology is changing, people's understanding of the domain and their tasks change as well. In the evaluation of future system/product features or total performance, it is important to consider alternative design solutions and not only make comparisons with current systems/products, since work procedures, interaction dialogues, tasks, and organisation may be changed due to the implementation of the new technology.

From a design procedure point of view, a planned future system has to be compared with possible future alternatives. Comparing a non-existing future system with systems of today is not a fair comparison. What are the conclusions if the prototype of the future system seems to be better than the existing systems? It is a very modest goal to put a lot of money and effort into a new system, just to find that it is a slightly better than an existing system. Especially for early phases of an iterative evaluation process, it is more reasonable to compare various kinds of future concepts with each other, to find the best candidate concept or combination of concepts. Doing this, we will take part of the benefits with iterative evaluation, and gradually refine the technical solutions as well as gradually increase our understanding of what is central.

Another problem with designing future systems is that the prototypes are not as complete as an existing system. If they are, the cost of refusing the ideas implemented in the prototypes could be unjustified. Comparing a prototype with an existing system, or comparing several prototypes with each other therefore demands effort in thoroughly thinking through which hypotheses are possible and relevant to test.

A related problem is that users, that often have to be included in an evaluation to make it valid, are biased in their behaviour and preferences by the fact that they have experience of existing systems but not the future systems. The current users may not be valid users for the future system. Also, the estimated context for a future system is, in the development phase, never completely correct.

When developing a system it is important to measure and analyse SA on multiple levels of analysis simultaneously, since the outcome of the analyses of human behaviour sometimes is dependent on the level of analyses, as demonstrated in Paper 3, and in Paper 5.

One way of assisting users, developers as well as researchers, and others to analyse a situation on multiple data levels simultaneously is by the reconstruction-exploration approach, using behavioural protocols and interactive data presentation, to represent what happened and why it happened (Albinsson, 2004; Jenvald, 1999). The reconstruction-exploration approach could also be combined with observer methods and competence analysis (Alfredson, Albinsson, Borgvall, & Svensson, 2007). By catching essentials of the situation and giving the opportunity to reconstruct and explore the situation various aspects of the situation could be scrutinised and compared with other events, and a dialogue between persons in various roles could be facilitated.

This approach is focusing on situations that have happened, but could still be valid for design work, since it is important to learn from past experiences. There are not however, the same opportunities when it comes to future situations. The best we can do is to describe the future situation. Only if we know the characteristics of situations that the product is going to be used in, we know how to develop presentation and manoeuvring support for the user.

When we have realized that this is situation dependent, the next step is to enhance the user's SA through the technical design. It is not only important to regard *what* actual guidance on the design is provided as a result of a study of SA for the purpose of developing a complex system, but also to regard *how* this guidance is obtained. That is, it is important to regard what methods have been used and be aware of the strength and weaknesses of the applied methods to the development and evaluation process. Dekker and Nyce (2004) describes some of these strength and weaknesses by comparing three approaches (experimental, ethnomethodological, and surveys) with examples from research on comparable context. For instance, all three approaches "rely on rhetoric and argument as well as method and analysis, to justify findings in terms of their future applicability" (Dekker & Nyce, 2004, p. 1624). Being aware of the method used and using the methods pragmatically helps in managing differences in SA in the development of complex systems.

5 Conclusions

This dissertation contributes to research by exploring the concept of SA, and especially differences in SA and how to mange them in the development of complex systems. In particular, the appended papers contribute with empirical data on differences in SA and the dissertation, as a whole, contributes with a analysis of how differences in SA could be managed, including concepts and models as tools for that.

Suggestions have been made on how to manage differences in SA in the development of complex systems. Differences due to design alternatives, roles in the design-use process, context, and level of analysis, have been related to state-of-the-art theory and investigated through empirical evaluations reported in the appended papers. The dissertation has also presented a collection of examples of contributing factors to both good and poor SM. This is meaningful and useful for anyone having an interest in the development process of complex systems either as researchers, managers, or developers of such a process, or professionals active in the process. Specifically, the following aspects are worth stressing:

- SA is a multifaceted concept and is best assessed by a multi-methodological approach. By combining several methods, such as, psychophysiological measures, performance measures, and other objective measures, in combination with subjective ratings including peer-ratings and observer ratings indices of situational awareness could be obtained.
- When combining measurement techniques a straightforward strategy is suggested that allocates one or more techniques to every aspect of SA, and every level of SA that is of interest.
- SA assessments could provide both quantitative and qualitative information valuable for the process of development of complex systems.
- A practical approach focusing on *differences* is suggested, since differences contain a lot of important information that are of value in a development process.
- Suggestions have also been made on theoretical improvements on SA theory. It is suggested that *high SA is when someone is fully aware of a high relative awareness*, stressing the importance of the metacognitive ability of knowing what you know, or rather, being aware of your awareness. For instance, the difference between what you are required to be aware of and what you are aware of is suggested as a SA-indicator. Also, the difference between what you are aware of and what you think you are aware of are suggested as another SA-indicator. That is, SA could

be determined by first comparing the differences between what someone is aware of and what that person should be aware of (i.e. the required awareness), and then check on the differences between the assessed awareness and a person's notion of that awareness, for instance, by a subjective assessment of the own SA.

- Just as the differences between an externally assessed awareness for a person and a subjective self assessment could be used as a part of the assessment of the individual SA, differences within a team, or JCS, could be used for assessing the team SA in an equivalent procedure. It is suggested that *high team SA is when the team members are fully aware of each other's high individual SA*. Differences within the team such as variations of *degree of agreement* concerning, for instance, *the situation*, individual *SA*, or *team-SA*, could thereby all be used in an assessment of team SA. Extrapolating the focus on differences, the equivalent approach could be used also for large organisations (i.e. groups of teams), systems of systems (or JCS of JCS), roles in the design-use process and more.
- The term SM has been suggested as a broader term than SA, focusing on the situation and how to manage the situation, rather than a focus on an individual or a team. The focus on the situation and SM is particularly relevant from the view of developing complex systems, since *SM is about coping with the situation using all available means of awareness and action,* including efforts in design and development. SM thereby includes *the SA of an individual or team, every part of the perception-action cycle, the management of mental resources, and external means of managing the situation*.

The suggestions made in this dissertation are applicable to a wide range of complex systems. For all application areas implementing the suggestions in the development of complex systems require changes at three levels of the development process simultaneously, 1) at the level of comprehending and distributing the notion of the SA concept, 2) at the level of SA assessment, and 3) at the level of managing differences in SA in the development process. If all three levels are adjusted synchronised the projected impact on improved management of future situations through improved SA for a user of a complex system is best applied. Differences in SA could thereby be managed in development of complex systems.

6 Abstracts of Appended Papers

This chapter consists of an abstract for each of the appended papers:

- Paper 1, Improvement of tactical situation awareness with colour-coded horizontal-situation displays in combat aircraft
- Paper 2, Pilots' understanding of situational awareness
- Paper 3, Mode monitoring and call-outs: An eye-tracking study of 2-crew automated flight deck operations
- Paper 4, Effect of task complexity and decision maker separation on shared situational awareness in command and control in network centric warfare
- Paper 5, Individual differences in visual behaviour in simulated flight
- Paper 6, Business process reengineering in the automotive area by simulator-based design

6.1 Paper 1

When the multi-role combat aircraft Gripen monochrome displays were to be upgraded in colour, the present study, studied the effects of colour-coded displays on visual search and situation awareness (SA). The intended use of the studied horizontal-situation display is to provide the pilot with a good overview to promote SA. Some of its information content are the positions and velocities of the own and of sensor tracked aircraft. These aircraft are represented as symbols superimposed on an electronic map moving in the same direction as the velocity vector of the own aircraft.

The simulations were run in a real-time simulation of an air-to-air mission in the Saab Aerospace simulator for displays and controls (PMSIM). The cockpit with its simulated cockpit displays was placed in a 3-metre radius dome.

Gripen's monochrome colour scheme was compared to two chromatic (dichrome, polychrome) colour schemes. The monochrome scheme was a simulation of the present display of Gripen with green as the colour. The dichrome presentation had one more colour in that one of the enemy symbols (targets) was orange when activated as priority target, with the rest of the presentation as the monochrome. A multi-colour coded presentation was provided by the polychrome scheme, where the targets representing the priority threat and the manoeuvring aircraft were orange as with the dichrome scheme. Targets representing other enemy aircraft were yellow, the symbols for the pilot's own aircraft and for the steering command were green, the cursor was white, and the horizon (with velocity vector) was light blue. Background information was presented in darker damped colours with land and water in faint green and faint blue, respectively, and bases in black and cities in nearest to grey.

A 3x2x2 factorial within subjects design was used; the three colour schemes each with two different background conditions (simple and complex) and two different symbol configurations. The pilot had two tasks during the simulation: (1) to track a manoeuvring aircraft within specified limits by using the HUD, and (2) to detect the appearance of a priority target on the head down horizontal-situation display (HSD). Deviations in flight path angle and reaction times for target detection were recorded. After the test runs, the pilot answered questions and ranked the colour schemes in different respects. The pilot also rated them for SA using a subjective rating technique on cognitive compatibility, CC-SART, Cognitive Compatibility – Situational Awareness Rating Technique (Taylor, 1995a; 1995b). It is made up of three primary and ten subsidiary constructs. Only the primary constructs supporting cognitive compatibility were used, and they are defined as follows. (1) *Level of Processing* (LoP): The degree to which the situation involves, at the lower score level, natural, automatic, intuitive, and associated processing, or, at the higher score level, analytically considered, conceptual and abstract processing. (2) *Ease of Reasoning* (EoR): The degree to which the situation, at the lower score level, is confusing and contradictory, or, at the higher score level, is straightforward and understandable. (3) *Activation of Knowledge* (AoK): The degree to which the situation at the lower score level is strange and unusual, or at the higher score level, is recognisable and familiar.

Seven male fighter pilots with mean age 40 years participated. Four of these pilots were from the Saab Squadron and three were from the Swedish Defence Materiel Administration, Testing Directorate. They all had good experience with the simulated aircraft system.

The results indicate that colour is advantageous in comparison to the monochrome display: The ranks on SA and preference ratings were higher for the chromatic schemes, and with the complex background, the reaction times were significantly lower for the polychrome colour code. The dichrome and polychrome colour schemes were advantageous to the monochrome in that the pilots' ratings of them on the CC- SART and the preference ratings were higher than with the monochrome. The pilots found it easier to search for the priority threat and were more confident in their detection of it with these chromatic displays. Furthermore, in the complex background condition, the reaction times for target detection were shorter with the polychrome scheme, as compared to the monochrome. In summary, the results indicate that colour can improve tactical SA in combat aircraft.

6.2 Paper 2

A questionnaire about SA was distributed to a number of pilots. The idea was to better understand how pilots understand the concept of SA, and to investigate if there were differences between how the two groups of pilots interpreted and experienced SA. The main part of the questionnaire consisted of questions that were to be answered by rating on a seven point Likert-scale. Each of these questions delivers suggested alternatives as well as an opportunity for the respondent to add any other alternative. In addition the respondent was able to make any other comments regarding his/her interpretation of SA, so that nothing important was to be left out. The questionnaire consisted of questions about the respondent, as age, sex, and flying experience, and also questions about the respondents understanding of the term SA, for instance, abilities required for high SA. The pilots were British, military pilots and Swedish private pilots. The results are based on the response from 31 military pilots and 120 private pilots.

There were no major differences among the pilots on how to interpret the concept of SA. This implies that the concept of SA demonstrates strong face validity among the broad pilot community. This gives us further hope of revealing the common understanding and agreement on the essence of SA. The fact that there was not much difference among the pilots in their responses also strengthens the points that are pointed out in the further discussion below.

It is strikingly many abilities of various kinds required for high SA that are rated high. This implies that SA is not easily expressed in one, or few, factors, but requires a complex mix of abilities. The results imply that the respondents regard SA as more than “perception of the elements of the environment within a volume of time and space, the comprehension of their meaning and the projection of their status in the near future” as proposed by Endsley (1988, p. 792). It is clear that the respondents also relate metacognitive aspects, collaboration aspects and that the volume of time regards past events also, but perhaps even more coming events.

That some feelings, such as feeling in control, feeling self-confident, and feeling secure, are associated with SA is interesting while, for instance, a feeling of being stressed is not. This implies that SA is associated with feelings, by the respondents. Since feelings are subjective experiences this might also be seen as an incentive to regard subjectivity in the study of SA. While many researchers take the stand that SA is objective, this indicates that that point of regard might not be the best to reduce the gap between theoretically used concepts and practitioners notion of the concepts. Since strong face validity is a strong reason to use the concept of SA, the notions of the pilots are important.

It is interesting that the respondents state that they feel that they are experiencing high SA in very different situations, such as both during an emergency, and in familiar situations. Situational Awareness Rating Technique (SART), developed by Taylor (1990), is a rating technique for SA allowing operators to rate various factors. Later Taylor developed Cognitive Compatibility - Situational Awareness Rating Technique (CC-SART) (Taylor, 1995a; 1995b). One of the factors in CC-SART is activation of knowledge, which is the amount of familiarity with the situation. From the results of this study it seems tempting to doubt the relevance in trying to calculate SA from dimensions such as familiarity, when the feeling of experiencing high SA seem to be directly situation dependent. If SA is measured directly, instead of through the calculation of pre-defined dimensions, this risk will diminish. Further the rating was high for the situations during landing, and during takeoff. The wide spectra of situations that the respondents rated high is further underlined by that the statement *depending on the situation* received the highest rating.

6.3 Paper 3

Mode awareness has been suggested as a critical factor in safe operations of automated aircraft. This study investigated mode awareness by measuring Eye-Point-of-Gaze (EPOG) of both pilots during simulated commercial flights, while recording call-outs and tracking aircraft performance. The results of this study indicate that the compliance to manufacturer, or air carrier procedures regarding mode monitoring and call-outs was very low. However, this did not seem to have a negative effect on the flight path or safety during our observations.

The work described here aims to gain a better understanding of the creation and breakdown of crew mode awareness in a normal operating environment. Specifically, it empirically examines the links between the perception of mode display annunciations, verbal call-outs, pilot roles (flying or non-flying pilot) and pilot rank (captain, co-pilot). It also sheds light on how the procedures guiding these factors are handled by the crew, and the potential consequences these factors may have on flight progress and the aircraft flight path. As the interest here was in two pilots' monitoring and call-out behaviour during normal flight conditions, taskload was not varied systematically except in the case of a go-around at the end of the flight.

The participants were twelve professional pilots from four air carriers with minimal differences in their flight operational procedures, all of them prescribing visual verification and call outs when mode transitions occur. All pilots where male, with between 1850 to 9000 hours of flying time. The flights were performed in a high fidelity, motion-based flight simulator with Boeing

737NG (New Generation) displays, where each pilot had a Primary Flight Display located next to the Navigation Display. The EPOG of both pilots was measured by a head- and eye tracking system. The pilots, their respective visual fields, and the cockpit displays were recorded by a total of six video cameras. These also registered, and kept the time of all flight deck sounds including pilot verbalizations. The simulator itself tracked a large number of parameters for aircraft performance and flight path.

Each crew was to fly a normal out-and-return flight from Amsterdam (AMS) to London (LHR). Flying time between Amsterdam to London is typically under one hour, and normally takes place in busy controlled airspace. The route to be programmed into the flight management system, and to be followed, was a normal company route into the London Heathrow area, across its northeast cornerpost. At the end of the scenario an event was introduced to the pilots with the intention of increasing their workload. The higher workload event was a glide slope capture failure on ILS RWY 27L that forced a go-around. The crews' mental effort was assessed by a Rating Scale Mental Effort (RSME) – a Likert scale indicator that pilots used to subjectively rate their effort on particular sets of tasks. The ratings of mental effort were taken after the flight to ensure that the pilots were not distracted from their tasks during the simulation. The results could be synchronized with other data traces. EPOG data, crew communication and flight and system parameters were also synchronized. The experiment was set up as a factorial between-subject design, using proportion tests as the statistical method for analyzing the data.

Crew mental effort was rated as moderate during most of the flight, with the go-around at London Heathrow as only exception (Zon et al., 2003). During the go-around and missed approach, mental effort was consistently rated as higher. While call-out rates were too low for statistical analysis, there is some support for the effect of workload on call-out frequency. Of all crews and flights studied, only one co-pilot made a mode call-out during the missed approach procedure. All others remained silent with respect to mode status, not even using informal communication to guide each other's attention to current or future mode status. This would confirm that in higher workload situations, automation call-outs could be among the tasks to be put aside.

The data from the experiment reported here demonstrates that flight mode annunciators may not really be attention-getting. In this experiment, crews hardly followed the official procedures with respect to supervising and verbally coordinating mode transitions. There is likely a distance between official written guidance and the constraints and conditions governing actual practice. Fewer than 3 percent of all call-outs were formal expressions and occurred when they should occur (after visually checking the FMA). There appears to be a diversity of locally and personally tailored strategies that pilots develop to try to stay ahead of what the automation is doing. In the twelve flights of this study, close to twenty different such strategies were documented.

6.4 Paper 4

It is tempting to hope for benefits from the transformation to network centric warfare on decision maker shared SA. However, command and control in network centric warfare probably include both high task complexity and decision maker separation, which in turn may affect the shared SA. Two experiments were conducted to elucidate the effect of task complexity and decision maker separation on shared SA in command and control in network centric warfare. One hypothesis was that degree of agreement will decrease as a consequence of higher information complexity (Experiment 1). Another hypothesis was that degree of agreement will decrease as a consequence of decision makers' separation (Experiment 2). The experiments were performed in highly valid settings for command and control in network centric warfare.

Experiment 1 took place at the Swedish armed forces C4ISR development and experiment center, a top modern environment for experimentation of command and control in network centric warfare. Twelve officers participated, four army officers, four navy officers and four air force officers, all male. They were all matched on military experience – within different groups as well as between groups. The scenario was written down and read as follows: "The year is 2012 and a number of countries have become hostile and Sweden is in the process of being attacked". Their task was to prioritize how they believed they (given their position) should act, before and after a simulated chain of events. Three participants participated in the experiment at the same time, one from each service. The three services studied were used as a between participant variable in the design. Time (before versus after) was the first within-participant variable. Type of event was the second within-participant variable (events rated as high on complexity versus low on complexity, by SME:s). A high complexity event was operationalised as an event that contained uncertainty and risk for changing plans. For example, "one of the suspicious ships turns towards Gotland". A low complexity event was operationalised as events that contained information evidently important and with more typical possible responses. For example, "one missile is fired towards Gothenburg". Two different events with low complexity and two events with high complexity were orthogonally balanced between sessions. The dependent measure used in the analysis was degree of agreement between participants. Three indexing procedures were used, *average differences*, *average of squared differences*, and *average correlation*. A 3 x 2 x 2 ANOVA, Service (Army, Navy, Air Force) x Time (before (time 1) versus after (time 2)) x Complexity (Low Complexity, High Complexity), using average differences. Two main effects were obtained. First, there was a main effect of time ($F(1,15)=6.24$, $p<.05$, $MSe=2.72$) that indicate that degree of agreement was higher at before compared to after. Second, a main effect of complexity was significant ($F(1,15)=7.12$, $p<.05$, $MSe=5.48$) that indicate higher degree of agreement was obtained when the event was of low complexity compared to high complexity. The results concerning complexity of events suggest that operator's agreed more on low complexity tasks, especially at the beginning of the experiment.

Experiment 2 was performed during an air force exercise for the Swedish Air Force. The air force's top modern technical networks including data links between all aircraft and decision makers on the ground, and state-of-the-art decision support technology, constitute exceptional conditions for empirical research on topics of command and control in network centric warfare. All the participants were from the Air Force. Officials of six positions in the chain of command and control were observed, in three different locations: Fighter pilots, and squadron leader at the airbase, fighter controllers, and fighter allocators at the air command operation centre and, assistant air defence commander and air defence commander at the air force tactical command. All real time decision makers potentially affecting each mission were thereby included. The task for the participants was to perform a number of air operations during a training week. Seven specific operations were chosen for this experiment. After each of the 7 chosen operations all six officials answered a questionnaire. Their task was to rate (on a scale from 1-7) degree of SA from different individuals perspective, that is, all rated his own degree of SA as well as the officials they were interacting with during the specific operation. The experiment was set up as a factorial between-subject design since the officials were substituted between operations. The first factor referred to degree of agreement within and between units. The second factor referred to the three studied units. The dependent variable was degree of agreement between participants.
The 3 x 2 ANOVA, Unit (Air force tactical command, Air command operation centre, Air base) x Decision maker separation (within unit versus between units) revealed significant effects. There was a significant main effect for Decision maker separation $F(1,49)=8.24$, $p<.05$, $MSe=0.78$, that is, officials within a unit had a higher degree of agreement concerning SA compared to officials between units. Experiment 2 suggests that decision maker separation contributes to problems with assessing each other's SA, which obstructs shared SA. Even if the ratings as such were as high for officials within as between units (in an absolutes sense) degree of agreement varied significantly.

Experiment 1 demonstrated that task complexity contributes to disparity in SA between decision makers in a command and control setting, which makes shared SA more difficult to achieve. Experiment 2 demonstrated that decision maker separation makes it harder to establish shared SA. These results are contradictory to the assumption of better shared SA following from the transformation to network centric warfare.

6.5 Paper 5

Flying an aircraft is highly visually demanding. It is very important to map pilot visual behaviour, both for the purpose of evaluating the cockpit interface and to effectively integrate it with future adaptive interfaces and decision support systems. Pilots' visual behaviour was studied in two experiments. In the first experiment commercial aviation pilots were flying a commercial aviation scenario and eye point of gaze, and eye blinks were collected. In the second experiment military pilots were flying an air-to-air combat scenario and the visual behaviour was video recorded.

The first experiment used the Saab Commercial Aircraft Development Simulator consisting of a fixed base cockpit based on the SAAB-2000 twin-prop passenger aircraft. Only the Captain's side had indicators and instruments. The Eye-Piont-Of-Gaze (EPOG) system recorded: Fixation number, surface number, X-coordinate on surface, Y-coordinate on surface, X-coordinate eye position, Y-coordinate eye position, Z-coordinate eye position, pupil diameter, current time from beginning of the day, duration of fixation, eye blinks. Six male pilots participated. Three of the participants were recruited from the Swedish Air Force Commercial Pilot School (TFHS) and three participants were a subset of Saab's test pilots for the Saab 2000. Only one overhead scenario was used, since each participant was confronted with the same three scenarios, in the same order. Each participant was confronted with the same events and the same flight phases within each scenario. The idea behind only having one overhead scenario was that all participants were confronted with the same events, and thereby also a similar amount of workload. Three conditions were used, one normal flight, one flight with engine failure at 1000ft, and one flight with NDB approach and landing. Each of these flights were complete flights with a subset of events, such as, taxiing, take-off, climb, cruise, and so forth.

The second experiment used Saabs advanced simulator Presentation and Manoeuvring SIMulator (PMSIM). PMSIM is a fixed base, dome simulator, where the visual surroundings are displayed on a dome with a radius of three metres. The simulated aircraft was the combat aircraft JAS 39 Gripen. Seven male military pilots participated. Four were test pilots from Saab, and three were from the Swedish Defence Materiel Administration, Testing Directorate.

The scenario was an air-to-air scenario. The scenario consisted of two tasks; One with focus head-up (HU) and one with focus head-down (HD). The HU task was to follow a manoeuvring aircraft at a requested distance. The HD task was to detect a change in one of the symbols presented on the head-down display, displaying the tactical situation on the horizontal-situation moving map display. A video recording was made of the pilot's face recorded by a video camera placed in front of the participant. In this study the analyser played the videotape and conducted a qualitative study on the participant's eye blink behaviour. The analyser noted when the participant was looking head-down or head-up. He noted when the participants blinked and looked for personal characteristics among the participants blinking behaviour.

The areas of interest for a pilot are dependent on the current flight phase, resulting in differences between flight phases in the collected EPOG data. In the first experiment, the head-up time increased significantly during the taxi phase, the take off phases, the landing phases, and the approach phases. The rest of the time, the head-down time was clearly dominating. At a lower level of analysis, the pilots in this experiment reacted and performed with large individual differences. The results from the correlations in the first experiment demonstrated that there were large individual differences in how the data was correlated. For two of the participants the percentage of time spent on the display representing the forward outside view correlated negatively with percentage of time spent on the navigation display, indicating that one of these surfaces were frequently used in flight phases when the other was not.

In the second experiment the pilots exposed typical blinking behaviour in that they typically blinked during eye movements, during fixation changes between head-down and head-up, or vice versa, and that they blinked more after the runs. After the runs, when the visual demands decrease, the pilot could blink a lot. It is interesting however, that this experiment also demonstrates that it is mainly in the changes between head-up/head-down or head-down/head-up that that the blinks occur, rather than between fixations within each of the areas. It is also notable that there seems to be personal characteristics or personal regularities in when to blink. Of the six participants, two blinked almost only on the change from head-up to head-down, two almost only blinked on the change from head-down to head-up, and two used half-blinks at the change from head-up to head-down. Among the two that used half-blinks one also blinked almost every time at the change from head-down to head-up, while the other participant almost never blinked a fully completed blink, during the runs.

6.6 Paper 6

The automotive industry is facing economic and technical challenges. The economic situation calls for more efficient processes, not only production processes but also renewals in the development process, such as Simulator-Based Design (SBD). The technology changes have been much larger than the human adaptation. People today are in principle not better off than previous generations when it comes to car driving. Today the car has become a mechatronic system where computer and communication technology dominate the system design. Earlier, it was possible to understand by just looking at a blueprint. Today, most new functionality is expressed in software code. This apparently deteriorated situation could be compensated by the use of simulation, since a software-based function is easier to implement in a simulation at all levels compared to a mechanical construction. HMI related systems could be classified in the following four groups:

- Primary control systems (PCS, primarily steering, throttle control, brakes and related automation)
- Advanced driving assistance systems (ADAS, e.g., visual enhancement systems, obstacle warning, other alarm systems and related automation)
- In-vehicle information systems (IVIS, incl. communication functions and other information related functions that require driver-vehicle interaction not directly related to the driving control loop)
- Non-integrated systems (any system that the driver might bring in to the vehicle for example, palms, separate cell phones, GPS), also called nomad systems

In order to cope with the new circumstances a SBD approach will be of uttermost importance. A basic philosophy in SBD is to make as many shortcuts as possible in order to evaluate or just demonstrate a solution or to look at different concepts for the system design. Another part of the basic philosophy is to have a pragmatic view on reality. It is not necessary to copy a specific environment if the purpose is to design a vehicle system. The only ambition should be that the environment model supports the simulation purpose. However, validity is important. A methodological approach to overcome certain validity problems is to conduct comparative experimental studies.

Only if we know the characteristics of situations that the car is going to be used in, we know to develop presentation and manoeuvring support for the driver. When we have realized that this is situation dependant, the next step is to enhance situation management through the technical design. One of the major reasons for using the concept of situation management is pragmatic: Applying the concept to the design might help creating good design. Since situation management is used to evaluate human-machine interfaces, the concept should enhance the design process, and increase the benefits of the result of the improved design process. It is important to provide a context of valid situations all the way through the development process, in order for the design process to be effective, through early simulations to the final design. Understanding how a situation is managed in valid situations helps guiding the design process into a rational path. Through a simulator enhanced, iterative design-process the important features of the design could be detected and optimized.

We briefly present two examples of simulator use for design issues. In one project example the purpose was to study night vision systems using two different concepts. One was the approach with continuous presentation on the windshield and the other was a warning system based on the same kind of sensor presentation. We wanted to know if negative behavioural adaptation could be avoided or at least minimized by using a warning system concept for night vision instead of the traditional continuous IR-display implemented. The result of this research question was that average driving speed was 5 km/h higher for the original system. The most evident result in this experiment was that all subjects preferred the new system that is, the situation-dependent display. We do not pretend to say anything on how important such a system is for traffic safety in absolute terms. Such questions will fall outside the scope of simulator-based design.

Another project example was a study on visual search and situation awareness as the Swedish multi-role combat aircraft Gripen's monochrome colour scheme was compared to chromatic. A real-time simulation of an air-to-air mission with seven male fighter pilots was carried out in Saab's fixed-base simulator for display and control design (PMSIM). The pilot had to perform dual tasks: 1) to track a manoeuvring aircraft within specified limits by using the head-up display, and 2) to detect the appearance of a target on the head-down display. Reaction times were measured, and the pilots made rankings of situation awareness and performance under different phases of the mission. The results were that colour displays were more beneficial than monochrome displays. Since the colour schemes were specified in a standardized colour appearance system, and because of the adequate fidelity in the simulator, according to current requirements, the design of the applied colour schemes in this study was possible to use as a ground for further design efforts towards the use of colour in the display in today's version of the fighter aircraft. To let the design efforts made in an early study phase be transformed into the design of a final product is a cost effective way of using the benefits of simulators.

References

Adams, M. J., & Pew, R. W. (1990). Situational awareness in the commercial aircraft cockpit: A cognitive perspective. In *Proceedings of the 9th Digital Avionics Systems Conference* (pp. 519-524). Virginia Beach, VA: IEEE.

Adams, M. J., Tenney, Y. J., & Pew, R. W. (1995). Situation awareness and the cognitive management of complex systems. *Human Factors, 37*(1), 85-104.

Adams, S. (1998). Practical considerations for measuring situational awareness. In *Proceedings of the 3rd Annual Symposium and Exhibition on Situational Awareness in the Tactical Air Environment* (pp. 157-164). Patuxent River, MD: Naval Air Warfare Center Aircraft Division.

Albinsson, P-A. (2004). *Interacting with command and control systems: Tools for operators and designers.* Licentiate thesis (No. 1132). Linköping, Sweden: Linköping University.

Albinsson, P-A., & Alfredson, J. (2002). *Reflections on practical representation design* (FOI-R-0716-SE). Linköping, Sweden: Swedish Defence Research Agency.

Alfredson, J. (1997). Temporal situational awareness. In P. Svenmarck (Ed.), *Föredrag Presenterade på Kontaktkonferens Arrangerad av Nationella Rådet för Forskning, Teknikutveckling och Undervisning om Människan i Flygsystemet* (pp. 31-34). Linköping, Sweden: Swedish Centre for Human Factors in Aviation. In Swedish.

Alfredson, J. (2001a). *Aspects of situational awareness and its measures in an aircraft simulation context.* Licentiate thesis (No. 865). Linköping, Sweden: Linköping University.

Alfredson, J. (2001b). Situational Awareness. In L. Sundling (Ed.), *Proceedings of the National Conference on Human Factors in Aviation* (pp. 85-88). Linköping, Sweden: Swedish Centre for Human Factors in Aviation.

Alfredson, J., Albinsson, P-A., Borgvall, J., & Svensson, E. (2007). Evaluering av CAS-förmåga för NBG: Ett metodperspektiv. In A. W. Berggren (Ed.), *Människan i NBF: Med ökad internationell förmåga i fokus* (pp. 299-333). Stockholm: Swedish National Defense College. In Swedish.

Alfredson, J., Angelborg-Thanderz, M., Danielsson, M., Farkin, B., Farmer, E., Hadley, B., Ohlsson, K., Svensson, E., Westlund, J., & White, J. (1996). *Human awareness criteria* (Tech. Rep. No. VINTHEC-WP1-TR01). Amsterdam: National Aerospace Laboratory NLR.

Alfredson, J., Angelborg-Thanderz, M., van Avermaete, J., Bohnen, H. G. M., Farkin, B., Ohlsson, K., Svensson, E., & Zon, G. D. R. (1997). *Dynamic measures of pilot mental workLoad (PMWL), pilot performance (PP), and situational awareness (SA)* (Tech. Rep. No. VINTHEC-WP3-TR01). Amsterdam: National Aerospace Laboratory NLR.

Alfredson, J., Derefeldt, G., & Jungert, E. (2004). Kunskapsrepresentation för utveckling av gemensam lägesförståelse i nätverk. I. G. Derefeldt & H. Friman (Eds.), *Samhällsförsvar: Nya hot och ökat internationellt engagemang* (pp. 177-192). Stockholm: The Swedish Institute for International Affairs. In Swedish.

Alfredson, J., Lindoff, J., & Rencrantz, C. (2004). *Requirements analysis of the SitSyst-Builder. A methodology report to support the subsequent development of service management tools.* (FOI-R-1231-SE). Linköping, Sweden: Swedish Defence Research Agency. In Swedish.

Alfredson, J., & Nählinder, S. (2002). The pilot's eye as a source of information for a future adaptive aircraft system. In D. Caldenfors, J. Eklund, & L. Kiviloog (Eds.), *Proceedings of the 34th Annual Congress of the Nordic Ergonomics Society* (pp. 1.21-1.26). Linköping, Sweden: Linköping University.

Alfredson, J., Nählinder, S., & Castor, M. (2005). *Measuring eye movements in applied psychological research - five different techniques - five different approaches* (FOI-R-1406-SE). Linköping, Sweden: Swedish Defence Research Agency.

Alfredson, J., Oskarsson, P-A., Castor, M., & Svensson, J. (2003). Development of a meta instrument for evaluation of man-system interaction in systems engineering. In G. L. Rafnsdóttir, H. Gunnarsdóttir, & Þ. Sveinsdóttir (Eds.), *Proceedings of the 35th Annual Congress of the Nordic Ergonomics Society* (pp.77-79). Reykjavík, Island: Svansprent.

Alfredson, J., Oskarsson, P-A., Castor, M., & Svensson, J. (2004). *Metodvalsverktyg: Ett hjälpmedel vid planering av MSI-utvärdering* (FOI-R-1295-SE). Linköping, Sweden: Swedish Defence Research Agency.

Alfredson, J., & Romare, F. (2002). *Method and device for the presentation of objects in surrounding space* (PCT Pub No: WO0127652). United States Patent.

Alfredson, J., Rosenblad, T., & Skinnars, Ö. (2002). *Display device for aircraft and method for displaying detected threats* (PCT Pub No: WO02090890). United States Patent.

Alm, T. (2001). How to put the real world into a 3D aircraft display. In *Proceedings of the 2nd International Conference on Human Interfaces in Control Rooms, Cockpits and Command Centres* (pp. 223-227). London: Institution of Electrical Engineers.

Alm, T. (2007). *Simulator-based design: Methodology and vehicle display applications.* Doctoral dissertation (No. 1078). Linköping, Sweden: Linköping University.

Alm, T., Alfredson, J., & Ohlsson, K. (submitted). Simulator-based HMI design for in vehicle systems. *International Journal of Vehicle Systems Modelling and Testing.*

Andersson, J. (1996). *Two is one too many: Dyadic memory collaboration effects on encoding and retrieval of episodes*. Doctoral dissertation (No. 47). Linköping, Sweden: Linköping University.

Andersson, J. (2001). Net effect of memory collaboration: How is collaboration affected by factors such as friendship, gender and age. *Scandinavian Journal of Psychology, 42*(4), 367-375.

Angelborg-Thanderz, M. (1990). *Military flight training at a reasonable price and risk.* Doctoral dissertation (No. 311). Stockholm: School of Economics. In Swedish.

Angelborg-Thanderz, M. (1997). Civil och militär MMI-forskning i komplicerade flygsystem: Bonuseffekter eller interferens? En sammanfattning. In P. Svenmarck (Ed.), *Föredrag Presenterade på Kontaktkonferens Arrangerad av Nationella Rådet för Forskning, Teknikutveckling och Undervisning om Människan i Flygsystemet* (pp. 51-56). Linköping, Sweden: Swedish Centre for Human Factors in Aviation. In Swedish.

Annett, J. (2002). Subjective rating scales: Science or art? *Ergonomics, 45*(14), 966-987.

Artman, H. (1999). *Fördelade kunskapsprocesser i ledningscentraler vid nödsituationer: Koordination och situationsmedvetenhet.* Doctoral dissertation (No. 186). Linköping, Sweden: Linköping University. In Swedish.

Artman, H. (2000). Team situation assessment and information distribution. *Ergonomics, 43*(8), 1111-1128.

Artman, H., & Garbis, C. (1998). Situation awareness as distributed cognition. In T. R. G. Green, L. Bannon, C. P. Warren, & J. Buckley (Eds.), *Proceedings of the 9th European Conference on Cognitive Ergonomics* (pp. 151-156). Paris: European Association of Cognitive Ergonomics.

Banbury, S., Dudfield, H., Hoermann, H-J., & Soll, H. (2007). FASA: Development and validation of a novel measure to assess the effectiveness of commercial airline pilot situation awareness training. *The International Journal of Aviation Psychology, 17*(2), 131-152.

Barchéus, F., & Mårtensson, L. (2006). Air traffic management and future technology: The views of the controllers. *Human Factors and Aerospace Safety, 6*(1), 1-16.

Bell, H. H., & Lyon, D. R. (2000). Using observer ratings to assess situation awareness. In M. R. Endsley & D. J. Garland (Eds.), *Situation awareness analysis and measurement* (pp. 129-146). Mahwah, NJ: Erlbaum.

Bell, H. H., & Waag, W. L. (1995). Using observer ratings to assess situational awareness in tactical air environments. In D. J. Garland & M. R. Endsley (Eds.), *Proceedings of an International Conference on Experimental Analysis and Measurement of Situation Awareness* (pp. 93-99). Daytona Beach, FL: Embry-Riddle Aeronautical University Press.

Bennett, K. B., Cress, J. D., Hettinger, L. J., Stautberg, D., & Haas, M. W. (2001). A theoretical analysis and preliminary investigation of dynamically adaptive interfaces. *The International Journal of Aviation Psychology, 11*(2), 169-195.

Berggren, P. (1999). *Situational awareness, mental workload, and pilot performance.* Master thesis. Linköping, Sweden: Linköping University.

Berggren, P., Alfredson, J., Andersson, J., & Modéer, B. (2007). Comparing measurements of shared situational awareness in a microworld with a development environment. *Paper presented at the IFAC-HMS 2007*, 4-6 September, Seoul, Korea.

Björklund, C., Alfredson, J., & Dekker, S. (2003). Shared mode awareness of the FMA in commercial aviation: An eye-point-of-gaze, and communication data analysis in a high fidelity simulator environment. In *Proceedings of the 22nd European Annual Conference on Human Decision Making and Control* (pp. 119-126). Linköping, Sweden: Linköping University.

Boag, C. C., Neale, M., & Neal, N. (1999). Measuring situation awareness: A comparison of four measurement tools. In R. S. Jensen, B. Cox, J. D. Callister, & R. Lavis (Eds.), *Proceedings of the 10th International Symposium on Aviation Psychology* (pp. 1240-1246). Columbus, OH: The Ohio State University Press.

Boehm-Davis, D. A., Gray, W. D., & Schoelles, M. J. (2000). The eye blink as a physiological indicator of cognitive workload. In *Proceedings of the 14th Triennial Congress of the International Ergonomics Association and the 44th Annual Meeting of the Human Factors and Ergonomics Society* (pp. 6.116-6.119). Santa Monica, CA: Human Factors and Ergonomics Society.

Boivie, I., Gulliksen, J., & Göransson, B. (2006). The lonesome cowboy: A study of the usability designer role in systems development. *Interacting with Computers, 18*(4), 601-634.

Bolstad, C. A., Endsley, M. R., & Howell, C. D. (2002). General aviation pilot training for situation awareness: An evaluation. In *Proceedings of the Human Factors and Ergonomics Society 46th Annual Meeting* (pp. 21-25). Santa Monica, CA: Human Factors and Ergonomics Society.

Borst, C., Suijkerbuijk, H. C. H., Mulder, M., & Van Paassen, M. M. (2006). Ecological interface design for terrain awareness. *The International Journal of Aviation Psychology, 16*(4), 375-400.

Boy, G. (2000). A vision for cognitive engineering in the early twenty first century. In *Proceedings of the 14th Triennial Congress of the International Ergonomics Association and the 44th Annual Meeting of the Human Factors and Ergonomics Society* (pp. 1.576-1.579). Santa Monica, CA: Human Factors and Ergonomics Society.

Brehmer, B. (1991). Distributed decision making: Some notes on the literature. In J. Rasmussen, B. Brehmer, & J. Leplat (Eds.), *Distributed decision making: cognitive models for cooperative work* (pp. 3-14). New York: Wiley.

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: Elsevier.

Brickman, B. J., Hettinger, L. J., Roe, M. M., Stautberg, D., Vidulich, M. A., Haas, M. W., & Shaw, R. L. (1995). An assessment of situation awareness in an air combat simulation: The global implicit measure approach. In D. J. Garland & M. R. Endsley (Eds.), *Proceedings of an International Conference on Experimental Analysis and Measurement of Situation Awareness* (pp. 339-344). Daytona Beach, FL: Embry-Riddle Aeronautical University Press.

Brown, A. (1987). Metacognition, executive control, self regulation, and other more mysterious mechanisms. In F. E. Weinert & R. H. Kluwe (Eds.), *Metacognition, motivation and understanding* (pp. 65-116). Hillsdale, NJ: Erlbaum.

Burdick, M. D., & Shively, R. J. (2000). Evaluation of a computational model of situational awareness. In *Proceedings of the 14th Triennial Congress of the International Ergonomics Association and the 44th Annual Meeting of the Human Factors and Ergonomics Society* (pp. 1.109-1.112). Santa Monica, CA: Human Factors and Ergonomics Society.

Byrne, E. A. (1995). Role of volitional effort in the application of psychophysiological measures to situation awareness. In D. J. Garland & M. R. Endsley (Eds.), *Proceedings of an International Conference on Experimental Analysis and Measurement of Situation Awareness* (pp. 147-153). Daytona Beach, FL: Embry-Riddle Aeronautical University Press.

Candell, O. (2004). Development of user information products for complex technical systems. Licentiate thesis (No. 2004:25). Luleå, Sweden: Luleå University of Technology.

Carlander, O., & Eriksson, L. (2006). Uni- and bimodal threat cueing with vibrotactile and 3D audio technologies in a combat vehicle. In *Proceedings of the Human Factors and Ergonomics Society 50th Annual Meeting* (pp. 1552-1556). Santa Monica, CA: Human Factors and Ergonomics Society.

Castor, M., Hanson, E., Svensson, E., Nählinder, S., Le Blaye, P., Macleod, I., Wright, N., Alfredson, J., Ågren, L., Berggren, P., Juppet, V., Hilburn, B., & Ohlsson, K. (2003). *GARTEUR Handbook of mental workload measurement* (GARTEUR TP 145). The Group for Aeronautical Research and Technology in Europe.

Chambres, P., Izaute, M, & Marescaux P-J. (Eds.). (2002). *Metacognition: Process, function and use*. Boston: Kluwer Academic Publishers.

Cooke, N. J. (1997). Broadening the measurement of situation awareness through cognitive engineering methods. In *Proceedings of the Human Factors and Ergonomics Society 41st Annual Meeting* (pp. 215-219). Santa Monica, CA: Human Factors and Ergonomics Society.

Cross, N. (1995). Discovering design ability. In R. Buchanan & V. Margolin (Eds.), *Discovering Design* (105-120). Chicago: The University of Chicago Press.

Cummings, M. L., & Tsonis, C. G. (2006). Partitioning complexity in air traffic management tasks. *The International Journal of Aviation Psychology, 16*(3), 277-295.

Dahlström, N. (2007). *Exploring levers for improvement of basic civil aviation training.* Doctoral dissertation. Lund, Sweden: Lund University.

Darses, F., & Wolff, M. (2006). How do designers represent to themselves the users' needs? *Applied Ergonomics, 37*(6), 757-764.

Dekker, S. W. A. (1996). Cognitive complexity in management by exception: Deriving early human factors requirements for an envisioned air traffic management world. In D. Harris (Ed.), *Engineering psychology and cognitive ergonomics, Volume I: Transportation systems* (pp. 201-210). Aldershot, England: Ashgate.

Dekker, S. W. A. (2000). Crew situation awareness in high-tech settings: Tactics for research into an ill-defined phenomenon. *Transport Human Factors, 2*(1), 49-62.

Dekker, S. W. A. (2003). Illusions of explanation: A critical essay on error classification. *The International Journal of Aviation Psychology, 13*(2), 95-106.

Dekker, S. W. A. (2005). *Ten questions about human error: A new view of human factors and systems safety*. Mahwah, NJ: Erlbaum.

Dekker, S. W. A. (2007). Doctors are more dangerous than gun owners: A rejoinder to error counting. *Human Factors, 49*(2), 177-184.

Dekker, S. W. A., & Lützhöft, M. (2004). Correspondence, cognition and sensemaking: A radical empiricist view of situation awareness. In S. Banbury & S. Tremblay (Eds.), *A cognitive approach to situation awareness: Theory and application* (pp. 22-41). Aldershot, England: Ashgate.

Dekker, S. W. A., & Nyce, J. M. (2004). How can ergonomics influence design? Moving from research findings to future systems. *Ergonomics, 47*(15), 1624-1639.

Dekker, S. W. A., & Woods, D. D. (1999). Extracting data from the future: Assessment and certification of envisioned systems. In S. Dekker & E. Hollnagel (Eds.), *Coping with computers in the cockpit* (pp. 131-143). Aldershot, England: Ashgate.

Derefeldt, G., Eriksson, L., Andersson, P., Berggrund, U., Rozen, U.,Westlund, J., Skinnars, Ö., Alfredson, J., Santesson, R., & Holmberg, J. (1998a). *Using color in horizontal-situation displays (HDS) improves tactical situation awareness in combat aircraft* (FOA-R-98-00767-706-SE). Linköping, Sweden: National Defence Research Establishment.

Derefeldt, G., Westlund, J., Skinnars, Ö., Alfredson, J., Andersson, P., Holmberg, J. Eriksson, L., Berggrund, U., & Santesson R. (1998b). Color for moving map cockpit displays. In B. E. Rogowitz & T. N. Pappas (Eds.), *Proceedings of IS&T/SPIE's 10th Annual Symposium: Vol. 3299. Human Vision and Electronic Imaging VII* (pp. 378-389). Bellingham, WA: SPIE Press.

Dick, A. O. (1980). *Instrument scanning and controlling: Using eye movement data to understand pilot behaviour and strategies.* (Contractor Report 3306). Langley Research Centre, VA: National Aeronautics and Space Administration.

Dreyfus, H. L. (1979). *What computers can't do: The limits of artificial intelligence.* New York: Haper & Colophon Books.

Durso, F. T., Bleckley, M. K., & Dattel, A. R. (2006). Does situation awareness add to the validity of cognitive tests? *Human Factors, 48*(4), 721-733.

Durso, F. T., & Dattel, A. R. (2004). SPAM: The real-time assessment of SA. In S. Banbury & S. Tremblay (Eds.), *A cognitive approach to situation awareness: Theory and application* (pp. 137-154). Aldershot, England: Ashgate.

Durso, F. T., Truitt, T. R., Hackworth, C. A., Crutchfield, J. M., Nikolic, D., Moertl, P. M., Ohrt, D., & Manning, C. A. (1995). Expertise and chess: A pilot study comparing situation awareness methodologies. In D. J. Garland & M. R. Endsley (Eds.), *Proceedings of an International Conference on Experimental Analysis and Measurement of Situation Awareness* (pp. 295-303). Daytona Beach, FL: Embry-Riddle Aeronautical University Press.

Dutarte, D. (2001). *A human perspective on ship and aircraft control*. Licentiate thesis. Stockholm: Royal Institute of Technology.

Edgar, G. K., & Edgar, H. E. (2007). Using signal detection theory to measure situation awareness: The technique, the tool (QUASA™), the test, the way forward. In M. Cook, J. Noyes, & Y. Masakowski (Eds.), *Decision making in complex environments* (pp. 373-385). Aldershot, England: Ashgate.

Edgar, G. K., Edgar, H. E., & Curry, M. B. (2003). Using signal detection theory to measure situation awareness in command and control. In *Proceedings of the Human Factors and Ergonomics Society 47th Annual Meeting* (pp. 2019-2023). Santa Monica, CA: Human Factors and Ergonomics Society.

Edgar, G. K., Smith, A. J., Stone, H. E., Beetham, D. L., & Pritchard, C. (2000). QUASA: Quantifying and analysing situational awareness. In *Proceedings of the IMCD People in Digitized Command and Control Symposium*, RMCS Shrivenham, UK. 12-14 Dec, 2000. CD-ROM.

Ek, Å. (2006). *Safety culture in sea and aviation transport.* Doctoral dissertation. Lund, Sweden: Lund University.

Endsley, M. R. (1988). Situation awareness global assessment technique (SAGAT). In *Proceedings of the IEEE National Aerospace and Electronics Conference* (pp. 789-795). New York: IEEE.

Endsley, M. R. (1990). Predictive utility of an objective measure of situation awareness. In *Proceedings of the Human Factors Society 34th Annual Meeting* (pp. 41-45). Santa Monica, CA: Human Factors Society.

Endsley, M. R. (1993). Situation awareness and workload: Flip sides of the same coin. In R. S. Jensen & D. Neumeister (Eds.), *Proceedings of the 7th International Symposium on Aviation Psychology* (pp. 906-911). Columbus: Ohio State University.

Endsley, M. R. (1995a). Measurement of situation awareness in dynamic systems. *Human Factors, 37*(1), 65-84.

Endsley, M. R. (1995b). Toward a theory of situation awareness in dynamic systems. *Human Factors, 37*(1), 32-64.

Endsley, M. R. (2000). Direct measurement of situation awareness: Validity and use of SAGAT. In M. R. Endsley & D. J. Garland (Eds.), *Situation awareness analysis and measurement* (pp. 147-173). Mahwah, NJ: Erlbaum.

Endsley, M. R. (2004). Situation awareness: Progress and directions. In S. Banbury & S. Tremblay (Eds.), *A cognitive approach to situation awareness: Theory and application* (pp. 317-341). Aldershot, England: Ashgate.

Endsley, M. R., Bolstad, C. A., Jones, D. G., & Riley, J. M. (2003). Situation awareness oriented design: From user's cognitive requirements to creating effective supporting technologies. In *Proceedings of the Human Factors and Ergonomics Society 47th Annual Meeting* (pp. 268-272). Santa Monica, CA: Human Factors and Ergonomics Society.

Endsley, M. R., Bolté, B., & Jones, D. G. (2003). *Designing for situation awareness: An approach to user-centered design*. London: Taylor & Francis.

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* (pp. 46-67). Santa Monica, CA: Human Factors and Ergonomics Society.

Endsley, M. R., Selcon, S. J., Hardiman, T. D., & Croft, D. G. (1998). A comparative analysis of SAGAT and SART for the evaluation of situation awareness. In *Proceedings of the Human Factors and Ergonomics Society 42nd Annual Meeting* (pp. 82-86). Santa Monica, CA: Human Factors and Ergonomics Society.

Eriksson, L., Tribukait, A., von Hofsten, C., & Eiken, O. (2006). On reducing the somatogravic illusion with HMD conveyed visual flow. In *Proceedings of the Human Factors and Ergonomics Society 50th Annual Meeting* (pp. 1599-1603). Santa Monica, CA: Human Factors and Ergonomics Society.

Eriksson, L., van Erp, J., Carlander, O., Levin, B., van Veen, H., & Veltman, H. (2006). Vibrotactile and visual threat cueing with high G threat intercept in dynamic flight simulation. In *Proceedings of the Human Factors and Ergonomics Society 50th Annual Meeting* (pp. 1547-1551). Santa Monica, CA: Human Factors and Ergonomics Society.

Fitts, P. M. (1951). *Human engineering for an effective air navigation and traffic control system*. Ohio State University Research Foundation Report, Columbus, OH.

Fitts, P. M., Jones, R. E., & Milton, J. L. (1950). Eye movements of aircraft pilots during instrument-landing approaches. *Aeronautical Engineering Review, 9*(2), 24-29.

Flach, J. M. (1995a). Maintaining situation awareness when stalking cognition in the wild. In D. J. Garland & M. R. Endsley (Eds.), *Proceedings of an International Conference on Experimental Analysis and Measurement of Situation Awareness* (pp. 25-34). Daytona Beach, FL: Embry-Riddle Aeronautical University Press.

Flach, J. M. (1995b). Situation awareness: Proceed with caution. *Human Factors, 37*(1), 149-157.

Flavell, J. H. (1976). Metacognitive aspects of problem solving. In L. B. Resnick (Ed.) *The Nature of Intelligence* (pp. 231-235). Hillsdale, NJ: Erlbaum.

Fracker, M. L. (1988). A theory of situation assessment: Implications for measuring situation awareness. In *Proceedings of the Human Factors Society 32nd Annual Meeting* (pp. 102-106). Santa Monica, CA: Human Factors Society.

Fracker, M. L., & Vidulich, M. A. (1991). Measurement of situation awareness: A brief review. In *Proceedings of the 11th Congress of the International Ergonomics Association* (pp. 795-797). London: Taylor & Francis.

Fredriksson, M. (2004). *Online engineering: On the nature of open computational systems.* Doctoral dissertation (No. 2004:05). Karlskrona, Sweden: Blekinge Institute of Technology.

French, H. T., Clarke, E., Pomeroy, D., Seymour, M., & Clark, R. (2007). Psycho-physiological measures of situation awareness. In M. Cook, J. Noyes, & Y. Masakowski (Eds.), *Decision making in complex environments* (pp. 291-298). Aldershot, England: Ashgate.

Funk, K. H., Lyall, E. A., & Niemczyk, M. C. (1997). Flightdeck automation problems: Perceptions and reality. In M. Mouloua & J. M. Koonce (Eds.), *Human-automation interaction: Research and practice* (pp. 29-34). Mahwah, NJ: Erlbaum.

Fuster, J. M. (1989). *The prefrontal cortex: Anatomy, physiology, and neuropsychology of the frontal lobe* (2nd ed.). New York: Raven Press.

Fuster, J. M. (1995). *Memory in the cerebral cortex: An empirical approach to neural networks in the human and nonhuman primate.* Cambridge, MA: MIT Press.

Fällman, D. (2003). *In romance with the materials of mobile interaction: A phenomenological approach to the design of mobile information technology*. Doctoral dissertation. Umeå, Sweden: University of Umeå.

Garbis, C., & Artman, H. (2004). Team situation awareness as communicative practices. In S. Banbury & S. Tremblay (Eds.), *A cognitive approach to situation awareness: Theory and application* (pp. 275-296). Aldershot, England: Ashgate.

Geddie, J. C., Boer, L. C., Edwards, R. J., Enderwick, T. P., Graff, N., Pfendler, C., Ruisseau, J.-Y., & Van Loon, P. A. (2001). *NATO guidelines on human engineering testing and evaluation* (RTO-TR-021). Neuilly-Sur-Seine, France: NATO Research and Technology Organization.

Goldberg, J. H. (2000). Eye movement-based interface evaluation: What can and cannot be assessed? In *Proceedings of the 14th Triennial Congress of the International Ergonomics Association and the 44th Annual Meeting of the Human Factors and Ergonomics Society* (pp. 6.625-6.628). Santa Monica, CA: Human Factors and Ergonomics Society.

Gorman, J. C., Cooke, N. J., Pederson, H. K., Connor, O. O., & DeJoode, J. A. (2005). Coordinated awareness of situation by teams (CAST): Measuring team situation awareness of a communication glitch. In *Proceedings of the Human Factors and Ergonomics Society 49th Annual Meeting* (pp. 274-277). Santa Monica, CA: Human Factors and Ergonomics Society.

Gorman, J. C., Cooke, N. J., & Winner, J. L. (2006). Measuring team situation awareness in decentralized command and control environments. *Ergonomics, 49*(12-13), 1312-1325.

Goteman, Ö. (2006). *Airborne cognitive systems in search of an appropriate context: Studies of the introduction of new technology in aviation using a cognitive systems engineering approach.* Doctoral dissertation. Lund, Sweden: Lund University.

Gould, J. D. (1976). Looking at pictures. In R. A. Monty & J. W. Senders (Eds.), *Eye movements and psychological processes* (pp. 323-345). Hillsdale, NJ: Erlbaum.

Graf, W., & Krueger, H. (1989). Ergonomic evaluation of user-interfaces by means of eye movement data. In M. J. Smith & G. Salvendy (Eds.), *Work with computers: Organizational, management, stress and health aspects* (pp. 659-665). Amsterdam: Elsevier.

Granlund, R. (2003). Monitoring experiences from command and control research with the C3Fire microworld. *Cognition, Technology & Work, 5*(3), 183-190.

Gugerty, L., & Tirre, W. (1995). Comparing explicit and implicit measures of situation awareness. In D. J. Garland & M. R. Endsley (Eds.), *Proceedings of an International Conference on Experimental Analysis and Measurement of Situation Awareness* (pp. 259-265). Daytona Beach, FL: Embry-Riddle Aeronautical University Press.

Gulliksen, J., Boivie, I., & Göransson, B. (2006). Usability professionals: Current practices and future development. *Interacting with Computers, 18*(4), 568-600.

Harris, R. L., & Christhilf, D. M. (1980). What do pilots see in displays? In *Proceedings of the Human Factors Society 24th Annual Meeting* (p. 22.6). Santa Monica, CA: Human Factors Society.

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In P. A. Hancock & N. Meshkati (Eds.), *Human mental workload* (pp. 5-39). New York: Elsevier.

Hartman, B. O., & Secrist, G. E. (1991). Situational awareness is more than exceptional vision. *Aviation, Space, and Environmental Medicine, 62*, 1084-1089.

Harwood, K., Barnett, B., & Wickens, C. D. (1988). Situational awareness: A conceptual and methodological framework. In F. E. Mc Intire (Ed.) *Proceedings of the 11th Symposium of Psychology in the Department of Defense* (pp. 316-320). Colorado Springs, CO: US Air Force Academy.

Hauland, G. (2002). *Measuring team situation awareness in training of en route air traffic control: Process oriented measures for experimental studies*. Doctoral dissertation (Risø-R-1343). Risø, Denmark: Risø National Laboratory.

Hauss, Y., & Eyferth, K. (2003). Securing future ATM-concepts' safety by measuring situation awareness in ATC. *Aerospace Science and Technology, 7*(6), 417-427.

Helander, M. G. (2000). Theories and models of electronic commerce. In *Proceedings of the 14th Triennial Congress of the International Ergonomics Association and the 44th Annual Meeting of the Human Factors and Ergonomics Society* (pp. 2.770-2.773). Santa Monica, CA: Human Factors and Ergonomics Society.

Helander, M. (2006). *A guide to human factors and ergonomics* (2nd ed.). Boca Raton, FL: Taylor & Francis.

Helander, M. G., & Khalid, H. M. (2000). Modeling the customer in electronic commerce. *Applied Ergonomics, 31*(6), 609-619.

Hoc, J. M. (2000). From human-machine interaction to human-machine cooperation. *Ergonomics, 43*(7), 833-843.

Hogg, D. N., Follesø, K., Strand-Volden, F., & Torralba, B. (1995). Development of a situation awareness measure to evaluate advanced alarm systems in nuclear power plant control rooms. *Ergonomics, 38*(11), 2394-2413.

Hollnagel, E. (1993). *Human reliability analysis, context and control*. London: Academic Press.

Hollnagel, E. (1999). From function allocation to function congruence. In S. Dekker & E. Hollnagel (Eds.), *Coping with computers in the cockpit* (pp. 29-53). Aldershot, England: Ashgate.

Hollnagel, E. (2007). Flight decks and free flight: Where are the system boundaries? *Applied Ergonomics, 38*(4), 409-416.

Hollnagel, E., & Woods, D. D. (1983). Cognitive systems engineering: New wine in new bottles. *International Journal of Man-Machine Studies, 18*(6), 583-600.

Hollnagel, E., & Woods, D. D. (2005). *Joint cognitive systems: Foundations of cognitive systems engineering*. Boca Raton, FL: Taylor & Francis.

Horswill, M. S., & McKenna, F. P. (2004). Drivers' hazard perception ability: Situation awareness on the road. In S. Banbury & S. Tremblay (Eds.), *A cognitive approach to situation awareness: Theory and application* (pp. 155-175). Aldershot, England: Ashgate.

Howard, M. (2002). *Usefulness in representation design.* Doctoral dissertation (No. 753). Linköping, Sweden: Linköping University.

Hughes, E. R., Hassoun, J. A., & Ward, F. (1990). *Advanced tactical fighter (ATF) simulation support: Final report.* (CSEF-TR-ATF-90-91). Wright-Patterson Air Force Base, OH: Air Force Systems Command, Aeronautical Systems Division.

Hult, L., Irestig, M., & Lundberg, J. (2006). Design perspectives. *Human-Computer Interaction, 21*(1), 5-48.

Itoh, Y., Hayashi, Y., Tsukui, I., & Saito, S. (1990). The ergonomic evaluation of eye movements and mental workload in aircraft pilots. *Ergonomics, 33*(6), 719-733.

Jeannot, E., Kelly, C., & Thompson, D. (2003). *The development of situation awareness measures in ATM systems* (HRS/HSP-005-REP-01). Bretigny-sur-Orge, France: Eurocontrol Experimental Centre.

Jenvald, J. (1999). *Methods and tools in computer-supported taskforce training*. Doctoral dissertation (No. 598). Linköping, Sweden: Linköping University.

Johansson, B. (2005). *Joint control in dynamic situations.* Doctoral dissertation (No. 972). Linköping, Sweden: Linköping University.

Jones, D. G. (2000). Subjective measures of situation awareness. In M. R. Endsley & D. J. Garland (Eds.), *Situation awareness analysis and measurement* (pp. 113-128). Mahwah, NJ: Erlbaum.

Jones, D. G., & Endsley, M. R. (1996). Sources of situation awareness errors in aviation. *Aviation, Space, and Environmental Medicine, 67*(6), 507-512.

Jones, D. G., & Endsley, M. R. (2001). Overcoming representational errors in complex environments. *Human Factors, 42*(3), 367-378.

Jones, D. G., & Endsley, M. R. (2004). Use of real-time probes for measuring situation awareness. *The International Journal of Aviation Psychology, 14*(4), 343-367.

Jones, D. G., Endsley, M. R., Bolstad, M, & Estes, G. (2004). The designer's situation awareness toolkit: Support for user-centered design. In *Proceedings of the Human Factors Society 48th Annual Meeting* (pp. 653-657). Santa Monica, CA: Human Factors Society.

Jones, D. G., & Kaber, D. B. (2005). Situation awareness measurement and the situation awareness global assessment technique. In N. Stanton, A. Hedge, K. Brookhuis, E. Salas, & H. Hendrick (Eds.), *Handbook of human factors and ergonomics methods* (pp. 42.1-42.7). Boca Raton, FL: CRC Press.

Judge, C. L. (1992). Situation awareness: Modelling, measurement, and impacts introduction. In *Proceedings of the Human Factors Society 36th Annual Meeting* (pp. 40-42). Santa Monica, CA: Human Factors Society.

Kennedy, R. S., Braun, C. C., & Massey, C. J. (1995). *Performance, workload, and eye movement activity* (SAE Paper 951570). San Diego, CA: SAE.

Khalid, H. M. (1999). Uncovering customer needs for web-based DIY product design. In J. Axelsson, B. Bergman, & J. Eklund (Eds.), *Proceedings of the International Conference on TQM and Human Factors* (343-348). Linköping, Sweden: Linköping University.

Klein, G., Moon, B., & Hoffman, R. R. (2006). Making sense of sensemaking 1: Alternative perspectives. *IEEE Intelligent Systems, 21*(5), 70-73.

Klein, G., Orasanu, J., Calderwood, R., & Zsambok, C. (Eds.) (1993). *Decision making in action: Models and methods*. Norwood, NJ: Ablex.

Klinger, D. W., & Gomes, M. E. (1993). A cognitive systems engineering application for interface design. In *Proceedings of the Human Factors and Ergonomics Society 37th Annual Meeting* (pp. 1.16-1.20). Santa Monica, CA: Human Factors and Ergonomics Society.

Knez, W. L., & Ham, D. J. (2006). A comparison of the effects of fatigue on subjective and objective assessment of situation awareness in cycling. *Journal of Sports Science & Medicine, 5*(1), 89-96.

Koester, T. (2007). Psycho-physiological measurements of mental activity, stress reactions and situation awareness in the maritime full mission simulator. In M. Cook, J. Noyes, & Y. Masakowski (Eds.), *Decision making in complex environments* (pp. 311-320). Aldershot, England: Ashgate.

Koppenjan, J., & Groenewegen, J. (2005). Istitutional design for complex technological systems. *International Journal of Technology, Policy and Management, 5*(3), 40-257.

Lankford, C. P., Shannon, P. F., Beling, P. A., McLaughlin, P. J., Israelski, E. W., Ellis, S. H., & Hutchinson, T. E. (1997). Graphical user interface design using eye gaze tracking and pupil response with ERICA. In *Proceedings of the Human Factors and Ergonomics Society 41st Annual Meeting* (p. 1371). Santa Monica, CA: Human Factors and Ergonomics Society.

Lichacz, F. M. J., Cain, B., & Patel, S. (2003). Calibration of confidence in situation awareness queries. In *Proceedings of the Human Factors and Ergonomics Society 47th Annual Meeting* (pp. 222-226). Santa Monica, CA: Human Factors and Ergonomics Society.

Lichacz, F. M. J., & Farrell, P. S. E. (2005). The calibration of situation awareness and confidence within a multinational operational net assessment. *Military Psychology, 17*(4), 247-268.

Linde, L., Strömberg, D., Pettersson, G., Andersson, R., & Alfredson, J. (2001). Evaluation of a temporal display for presentation of threats in air combat. In R. Onken (Ed.), *Proceedings of CSAPC'01 8th Conference on Cognitive Science Approaches to Process Control* (pp. 137-146). Neubiberg, Germany: University of the German Armed Forces.

Lindh, J-O. (2006). *On observation of and interaction in open distributed systems*. Doctoral dissertation (No. 2006:06). Karlskrona, Sweden: Blekinge Institute of Technology.

Lorenz, B., Biella, M., Teegen, U., Stelling, D., Wenzel, J., Jakobi, J., Ludwig, T., & Korn, B. (2006). Performance, situation awareness, and visual scanning of pilots receiving onboard taxi navigation support during simulated airport surface operation. *Human Factors and Aerospace Safety, 6*(2), 135-154.

Löwgren, J. (1993). *Human computer interaction: What every system developer should know.* Lund, Sweden: Studentlitteratur.

Ma, R., & Kaber, D. B. (2007). Situation awareness and driving performance in a simulated navigation task. *Ergonomics, 50*(8), 1351-1364.

Magnusson, S. (2002). Similarities and differences in psychophysiological reactions between simulated and real air-to-ground missions. *The International Journal of Aviation Psychology, 12*(1), 49-61.

MANPRINT in acquisition: A handbook. (2000). Prepared by office of the deputy chief of staff for personnel. Washington, DC: Personnel Technologies Directorate 300 Army Pentagon.

Marshall, S. P. (1996). *Monitoring situational awareness in tactical information displays with eye-tracking instrumentation* (N00014-95-1-1091). San Diego, CA: San Diego State University.

Matthews, M. D., & Beal, S. A. (2002). *Assessing situation awareness in field training exercises* (Report No. 1795). Arlington, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Matthews, M. D., Pleban, R. J., Endsley, M. R., & Strater, L. D. (2000). Measures of infantry situation awareness for a virtual MOUT environment. In *Proceedings of the Human Performance, Situation Awareness and Automation: User Centered Design for the New Millennium Conference* (pp. 262-267). Savannah, GA: SA Technologies.

Matthews, M. D., Strater, L., D., & Endsley, M. R. (2004). Situation awareness requirements for infantry platoon leaders. *Military Psychology, 16*(3), 149-161.

May, J. G., Kennedy, R. S., Williams, M. C., Dunlap, W. P., & Brannan, J. R. (1990). Eye movements indices of mental workload. *Acta Psychologica, 75,* 75-89.

McGowan, A. M., & Banbury, S. P. (2004). Evaluating interruption-based techniques using embedded measures of driver anticipation. In S. Banbury & S. Tremblay (Eds.), *A cognitive approach to situation awareness: Theory and application* (pp. 176-192). Aldershot, England: Ashgate.

McGuinness, B. (1995). Situational awareness measurement in cockpit evaluation trials. In *Situation Awareness: Limitations and Enhancement in the Aviation Environment* (AGARD-CP-575, pp. 7.1-7.8). Neuilly-Sur-Seine, France: NATO Research and Technology Organization.

McGuinness, B. (2004). Quantitative analysis of situational awareness (QUASA): Applying signal detection theory to true/false probes and self-ratings. In *9th International Command and Control Research and Technology Symposium* (pp. 159-178). Washington, DC: US Department of Defence Command and Control Research Program.

McGuinness, B. (2007). Signal detection theory and the assessment of situation awareness. In M. Cook, J. Noyes, & Y. Masakowski (Eds.), *Decision making in complex environments* (pp. 299-309). Aldershot, England: Ashgate.

McGuinness, B., & Foy, L. (2000). A subjective measure of SA: The crew awareness rating scale (CARS). In *Proceedings of the 1st Human Performance, Situation Awareness and Automation Conference*. Savannah, Georgia: SA Technologies.

McKinnon, F. B. (1986). *Final report of the intraflight command, control, and communications symposium.* Las Vegas, NV: Nellis Air Force Base.

Means, B., Salas, E., Crandall, B., & Jacobs, T. O. (1993). Training decision makers for the real world. In G. A. Klein, J. Orasanu, R. Calderwood, & C. E. Zsambok (Eds.), *Decision making in action: Models and methods* (pp. 306-326). Norwood, NJ: Ablex.

Metalis, S. A. (1993). Assessment of pilot situational awareness: Measurement via simulation. In *Proceedings of the Human Factors and Ergonomics Society 37th Annual Meeting* (pp. 113-117). Santa Monica, CA: Human Factors and Ergonomics Society.

Milham, L. M., Barnett, J. S., & Oser, R. L. (2000). Application of an event-based situation awareness methodology: Measuring situation awareness in an operational context. In *Proceedings of the 14th Triennial Congress of the International Ergonomics Association and the 44th Annual Meeting of the Human Factors and Ergonomics Society* (pp. 2.423-2.426). Santa Monica, CA: Human Factors and Ergonomics Society.

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review, 63,* 81-97.

Mocharnuk, J. B. (1978). Visual target acquisition and ocular scanning performance. *Human Factors, 20*(5), 611-631.

Mourant, R. R., & Rockwell, T. H. (1972). Strategies of visual search by novice and experienced drivers. *Human Factors, 14*(4), 325-335.

Mulgund, S., Rinkus, G., Illgen, C., Zacharias, G., & Friskie, J. (1997). OLIPSA: On-line intelligent processor for situation assessment. In K. Garner (Ed.), *Second Annual Symposium and Exhibition on Situational Awareness in the Tactical Air Environment*. Patuxent River, MD: Naval Air Warfare Center Aircraft Division.

Mumaw, R. J., Roth, E. M., Vicente, K. J., & Burns, C. M. (2000). There is more to monitoring a nuclear power plant than meets the eye. *Human Factors, 42*(1), 36-55.

Muniz, E., Stout, R., Bowers, C., & Salas, E. (1998). A methodology for measuring team situational awareness: Situational linked indicators adapted to novel tasks (SALIANT). *RTO Human Factors and Medicine Symposium on Collaborative Crew Performance in Complex Operational Systems* (8 p.). Neuilly-Sur-Seine, France: NATO Research and Technology Organization.

Nagel, D. C. (1988). Human error in aviation operations. In E. L. Weiner & D. C. Nagel (Eds.), *Human Factors in Aviation* (263-303). San Diego, CA: Academic Press.

Neisser, U. (1976). *Cognition and reality principles and implications of cognitive psychology*. San Francisco: Freeman.

Noyes, J. M., & Bruneau, D. P. J. (2007). A self-analysis of the NASA-TLX workload measure. *Ergonomics, 50*(4), 514-519.

O'Brien, K. S., & O'Hare, D. O. (2007). Situational awareness ability and cognitive skills training in a complex real-world task. *Ergonomics, 50*(7), 1064-1091.

Olmos, O., Liang, C-C., & Wickens, C. D. (1995). Construct validity of situation awareness measurements related to display design. In D. J. Garland & M. R. Endsley (Eds.), *Proceedings of an International Conference on Experimental Analysis and Measurement of Situation Awareness* (pp. 219-225). Daytona Beach, FL: Embry-Riddle Aeronautical University Press.

Papantonopoulos, S. (2004). How system designers think: A study of design thinking in human factors engineering. *Ergonomics, 47*(14), 1528-1548.

Patrick, J., & James, N. (2004). A task-oriented perspective of situation awareness. In S. Banbury & S. Tremblay (Eds.), *A cognitive approach to situation awareness: Theory and application* (pp. 61-81). Aldershot, England: Ashgate.

Patrick, J., James, N., & Ahmed, A. (2007). Awareness of control room teams. *Le Travail Humain, 70*(1), 67-94.

Patrick, J., James, N., Ahmed, A., & Halliday, P. (2006). Observational assessment of situation awareness, team differences and training implications. *Ergonomics 49*(4), 393-417.

Perkins, D. N. (1989). Reasoning as it is and could be: An empirical perspective. In D. M. Topping, D. C. Crowell, & V. N. Kobayashi (Eds.), *The 3rd International Conference on Thinking* (pp. 175-194). Hillsdale, NJ: Erlbaum.

Pew, R. W. (1995). The state of situation awareness measurement: Circa 1995. In D. J. Garland & M. R. Endsley (Eds.), *Proceedings of an International Conference on Experimental Analysis and Measurement of Situation Awareness* (pp. 7-15). Daytona Beach, FL: Embry-Riddle Aeronautical University Press.

Phillips, C. (1986). *Final report of the intraflight command, control and communications symposium*. Las Vegas, NV: Nellis Air Force Base.

Previc, F. H., & Ercoline, W. R. (2004). Spatial disorientation in aviation: Historical background, concepts, and terminology. In F. H. Previc & W. R. Ercoline (Eds.), *Spatial disorientation in aviation* (pp. 1-36). Reston, VA: American institute of astronautics and aeronautics.

Prince, C., Ellis, E., Brannick, M. T., & Salas, E. (2007). Measurement of team situation awareness in low experience level aviators. *The International Journal of Aviation Psychology, 17*(1), 41-57.

Pritchett, A. R., & Hansman, R. J. (2000). Use of testable responses for performance-based measurement of situation awareness. In M. R. Endsley & D. J. Garland (Eds.), *Situation awareness analysis and measurement* (pp. 189-209). Mahwah, NJ: Erlbaum.

Proctor, M., & Lipinski, M. J. (2000). Technical performance measures and distributed-simulation training systems. *Acquisition Review Quarterly, 7*(1), 19-32.

Proctor, M. D., Panko, M., & Donovan, S. J. (2004). Considerations for training team situation awareness and task performance through PC-gamer simulated multiship helicopter operations. *The International Journal of Aviation Psychology, 14*(2), 191-205.

Proctor, R. W., & Dutta, A. (1995). *Skill acquisition and human performance.* London: SAGE.

Rasmussen, J. (1983). Skills, rules, and knowledge: Signals, signs, and symbols, and other distinction in human performance models. *IEEE Transactions on Systems, Man, & Cybernetics, 13*(3), 257-266.

Rasmussen, J. (1986). *Information processing and human-machine interaction: An approach to cognitive engineering.* New York: North-Holland.

Reason, J. (1988). Cognitive aids in process environments: Prostheses or tools? In E. Hollnagel, G. Mancini, & D. D. Woods (Eds.), *Cognitive engineering in complex dynamic worlds* (pp. 7-14). London: Academic Press.

Reason, J. (1990). *Human error.* Cambridge, England: Cambridge University Press.

Reason, J. (1997). *Managing the risks of organizational accidents*. Aldershot, England: Ashgate.

Regal, D. M., Rogers, W. H., & Boucek, G. P. (1988). Situational awareness in the commercial flight deck: Definition, measurement, and enhancement. In *Proceedings of the Aerospace Technology Conference and Exposition* (pp. 65-69). Warrendale, PA: Society of Automotive Engineers.

Riley, J. M., Endsley, M. R., Bolstad, C. A., & Cuevas, H. M. (2006). Collaborative planning and situation awareness in army command and control. *Ergonomics, 49*(12-13), 1139-1153.

Riley, J. M., Kaber, D. B., & Draper, J. V. (2004). Situation awareness and attention allocation measures for quantifying telepresence experiences in teleoperation. *Human Factors and Ergonomics in Manufacturing, 14*(1), 51-67.

Salmon, P., Stanton, N., Walker, G., & Green, D. (2006). Situation awareness measurement: A review of applicability for C4i environments. *Applied Ergonomics, 37*, 225-238.

Sarter, N. B., Mumaw, R. J., & Wickens, C. D. (2007). Pilots' monitoring strategies and performance on automated flight decks: An empirical study combining behavioral and eye-tracking data. *Human Factors, 49*(3), 347-357.

Sarter, N. B., & Woods, D. D. (1991). Situation Awareness: A Critical but ill-defined phenomenon. *The International Journal of Aviation Psychology, 1*(1), 45-57.

Sarter, N. B., & Woods, D. D. (1995). How in the world did we get into that mode? Mode and awareness in supervisory control. *Human Factors, 37*(1), 5-19.

Saus, E-R., Johnsen, B. H., Eid, J., Riisem, P. K., Andersen, R, & Thayer, J. F. (2006). The effect of brief situational awareness training in a police shooting simulator: An experimental study. *Military Psychology, 18*(Suppl.), S3-S21.

Scallen, S. F., & Hancock, P. A. (2001). Implementing adaptive function allocation. *The International Journal of Aviation Psychology, 11*(2), 197-221.

Schieber, F., Harms, M. L., Berkhont, J., & Spangler, D. S. (1997). Precision and accuracy of video-based measurements of driver's gaze location. In *Proceedings of the Human Factors and Ergonomics Society 41st Annual Meeting* (pp. 929-933). Santa Monica, CA: Human Factors and Ergonomics Society.

Selcon, S. J., & Taylor, R. M. (1990). Evaluation of the situational awareness rating technique (SART) as a tool for aircrew system design. In *Situational Awareness in Aerospace Operations* (AGARD-CP-478, pp. 5.1-5.8). Neuilly-Sur-Seine, France: NATO Research and Technology Organization.

Sennersten, C., Alfredson, J., Castor, M., Hedström, J., Lindahl, B., Lindley, C. A., Nählinder, S., & Svensson, E. (2007). *Verification of an experimental platform integrating a Tobii eyetraking system with the HiFi game engine* (FOI-R-2227-SE). Linköping, Sweden: Swedish Defence Research Agency.

Shively, R. J., Brickner, M., & Silbiger, J. (1997). A computational model of situational awareness instantiated in MIDAS. In R. S. Jensen & L. A. Rakovan (Eds.), *Proceedings of the Ninth International Symposium on Aviation Psychology* (pp. 1454-1459). Columbus, OH: University of Ohio.

Shively, R. J., & Goodman, A. D. (1994). Effects of perceptual augmentation of visual displays: Dissociation of performance and situational awareness. In *Proceedings of the Human Factors and Ergonomics Society 38th Annual Meeting* (pp. 1271-1274). Santa Monica, CA: Human Factors and Ergonomics Society.

Shook, R. W. C., & Garland, D. J. (2000). Situation awareness problems in general aviation. In *Proceedings of the 14th Triennial Congress of the International Ergonomics Association and the 44th Annual Meeting of the Human Factors and Ergonomics Society* (pp. 1.185-1.188). Santa Monica, CA: Human Factors and Ergonomics Society.

Siem, F. M. (1992). A test battery to predict individual differences in situational awareness. In *Proceedings of the Human Factors Society 36th Annual Meeting* (p. 41). Santa Monica, CA: Human Factors Society.

Simons, R. (2005). *Levers of organization design: How managers use accountability systems for greater performance and commitment.* Boston: Harvard Business School Press.

Singer, G. (2002). *Methods for validating cockpit design: The best tool for the task.* Doctoral dissertation. Stockholm: Royal Institute of Technology.

Smallman, H. S., & St.John, M. (2005). Naïve realism: Misplaced faith in realistic displays. *Ergonomics in Design, 13*(3), 6-13.

Smith, K., & Hancock, P. A. (1995). Situation awareness is adaptive, externally directed consciousness. *Human Factors, 37*(1), 137-148.

Smolensky, M. W. (1993). Toward the physiological measurement of situational awareness: The case for eye movement measures. In *Proceedings of the Human Factors and Ergonomics Society 37th Annual Meeting* (p. 41). Santa Monica, CA: Human Factors and Ergonomics Society.

Sneddon, A., Mearns, K., & Flin, R. (2006). Situation awareness and safety in offshore drill crews. *Cognition, Technology & Work, 8*(4), 255-267.

Snow, M. P., & Reising, J. M. (2000). Comparison of two situation awareness metrics: SAGAT and SA-SWORD. In *Proceedings of the 14th Triennial Congress of the International Ergonomics Association and the 44th Annual Meeting of the Human Factors and Ergonomics Society* (pp. 3.49-3.52). Santa Monica, CA: Human Factors and Ergonomics Society.

Sohn, Y. W., & Doane, S. M. (2004). Memory processes of flight situation awareness: Interactive roles of working memory capacity, long-term working memory, and expertise. *Human Factors, 46*(3), 461-475.

Sonnenwald, D. H., & Pierce, L. G. (2000). Information behavior in dynamic group work contexts: Interwoven situational awareness, dense social networks and contested collaboration in command and control. *Information Processing and Management, 36*(3), 461-479.

Stanton, N. A., Stewart, R., Harris, D., Houghton. R. J., Baber, C., McMaster, R., Salmon, P., Hoyle, G., Walker, G., Young, M. S., Linsell, M., Dymott, R., & Green, D. (2006). Distributed situation awareness in dynamic systems: Theoretical development and application of an ergonomics methodology. *Ergonomics, 49*(12-13), 1288-1311.

Stark, L., & Ellis, S. R. (1981). Scanpaths revisited: Cognitive hypothesis direct active looking. In J. W. Senders, D. F. Fisher, & R. A. Monty (Eds.), *Eye movements: Cognition, and visual perception* (pp. 192-226). Hillsdale, NJ: Erlbaum.

Stern, J. A. (1988). Blink rate: A possible measure of fatigue. *Human Factors, 36*(2), 285-297.

Stock, O. (2003). Password swordfish: Verbal humor in the interface. *Humor 16*(3), 281-295.

Strömberg, D., & Alfredson, J. (1997). Temporal displays for fighter pilots. In S. Bagnara, E. Hollnagel, M. Mariani & L. Norros (Eds.), *Proceedings of the 6th European Conferense on Cognitive Science Approaches to Process Control* (pp. 209-212). Rome: CNR.

Svenmarck, P., & Dekker, S. (2003). Decision support in fighter aircraft: From expert systems to cognitive modelling. *Behaviour & Information Technology, 22*(3), 175-184.

Svensson, E., Angelborg-Thanderz, M., Olsson, S., & Sjöberg, L. (1992). *Risk för informationsöverflöde? Mental arbetsbelastning och prestation vid militär flygning* (Report C 50097-5.1). Linköping, Sweden: Swedish Defence Research Agency.

Svensson, E., Angelborg-Thanderz, M., & Sjöberg, L. (1993). Mission challenge, mental workload and performance in military aviation. *Aviation, Space, and Environmental Medicine, 64*(11), 985-991.

Svensson, E., Angelborg-Thanderz, M., Sjöberg, L., & Gillberg, M. (1988). Military flight experience and sympatho-adrenal activity. *Aviation, Space, and Environmental Medicine, 59*(5), 411-416.

Svensson, E., Angelborg-Thanderz, M., Sjöberg, L., & Olsson, S. (1997). Information complexity: Mental workload and performance in combat aircraft. *Ergonomics, 40*(3), 362-380.

Svensson, E., Angelborg-Thanderz, M., & Wilson, G. F. (1999). *Models of pilot performance for systems and mission evaluation: Psychological and psychophysiological aspects* (Report No. AFRL-HE-WP-TR-1999-0215). Wright-Patterson Air Force Base, OH: United States Air Force Research Laboratory.

Svensson, E., Rencrantz, C., Lindoff, J., Berggren, P., & Norlander, A. (2006). Dynamic measures for performance assessment in complex environments. In *Proceedings of the Human Factors and Ergonomics Society 50th Annual Meeting* (pp. 2585-2589). Santa Monica, CA: Human Factors and Ergonomics Society.

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.

Svensson, J., & Andersson, J. (2006). Speech acts, communication problems, and fighter pilot team performance. *Ergonomics, 49*(12-13), 1226-1237.

Taylor, R. M. (1990). Situational awareness rating technique (SART): The development of a tool for aircrew systems design. In *Situational Awareness in Aerospace Operations* (AGARD-CP-478, pp. 3.1-3.17). Neuilly-Sur-Seine, France: NATO Research and Technology Organization.

Taylor, R. M. (1995a). Experimental measures: Performance-based self ratings of situational awareness. In D. J. Garland & M. R. Endsley (Eds.), *Proceedings of an International Conference on Experimental Analysis and Measurement of Situation Awareness* (pp. 83-91). Daytona Beach, FL: Embry-Riddle Aeronautical University Press.

Taylor, R. M. (1995b). CC-SART: The development of an experimental measure of cognitive compatibility in systems design. *Report to TTCP UTP-7 Human Factors in Aircraft Environments, Annual Meeting*, DCIEM, Toronto, Canada, 12-16 June.

Taylor, R. M., & Selcon, S. J. (1990). Understanding situational awareness. *Contemporary Ergonomics* (pp. 105-111). London: HMSO.

Taylor, R. M., & Selcon, S. J. (1991). Subjective measurement of situational awareness. In *Proceedings of the 11th Congress of the International Ergonomics Association* (pp. 789-791). London: Taylor & Francis.

Tenney, Y. J., Adams, M. J., Pew, R. W., Huggins, A. W. F., & Rogers, W. H. (1992). *A principled approach to the measurement of situation awareness in commercial aviation* (Report No. 4451). Langley, VA: NASA.

Thomas, E. L. (1963). The eye movements of pilot during aircraft landing. *Aerospace Medicine, 34,* 424-426.

Tolk, J. D., & Keether, G. A. (1982). *Advanced medium-range air-to-air missile (AMRAAM) operational evaluation (OUE) final report (U).* Kirtland Air Force Base, NM: Air Force Test and Evaluation Center.

Van DerWeert, R. (1986). *Final report of the intraflight command, control and communications symposium*. Las Vegas, NV: Nellis Air Force Base.

Veltman, J. A., & Gaillard, A. W. K. (1998). Physiological workload reactions to increasing levels of task difficulty. *Ergonomics, 41*(5), 656-669.

Victor, T. (2005). *Keeping eye and mind on the road*. Doctoral dissertation. Uppsala, Sweden: Uppsala University.

Vidulich, M. A. (1989). The use of judgment matrices in subjective workload assessment: The subjective workload dominance (SWORD) technique. In *Proceedings of the Human Factors Society 33rd Annual Meeting* (pp. 1406-1410). Santa Monica, CA: Human Factors Society.

Vidulich, M. A. (1995). The role of scope as a feature of situation awareness metrics. In D. J. Garland & M. R. Endsley (Eds.), *Proceedings of an International Conference on Experimental Analysis and Measurement of Situation Awareness* (pp. 69-74). Daytona Beach, FL: Embry-Riddle Aeronautical University Press.

Vidulich, M. A. (2000). The relationship between mental workload and situation awareness. In *Proceedings of the 14th Triennial Congress of the International Ergonomics Association and the 44th Annual Meeting of the Human Factors and Ergonomics Society* (pp. 3.460-3.463). Santa Monica, CA: Human Factors and Ergonomics Society.

Vidulich, M. A., & Hughes, E. R. (1991). Testing a subjective metric of situation awareness. In *Proceedings of the Human Factors Society 35th Annual Meeting* (pp. 1307-1311). Santa Monica, CA: Human Factors Society.

Vidulich, M. A., Stratton, M., Crabtree, M., & Wilson G. (1994). Performance-based and physiological measures of situational awareness. *Aviation, Space, and Environmental Medicine, 65*(5), A7-A12.

Vidulich, M. A., Ward, G. F., & Schueren, J., (1991). Using the subjective workload dominance (SWORD) technique for projective workload assessment. *Human Factors, 33*(6), 677-692.

Waag, W. L., & Houck, M. R. (1994). Tools for assessing situational awareness in an operational fighter environment. *Aviation, Space, and Environmental Medicine, 65(*5), A13-A19.

Wachtel, P. L. (1967). Conceptions of broad and narrow attention. *Psychological Bulletin, 68*(6), 417-429.

Walker, G., Salmon, P., Green, D., Stanton, N., & Baber, C. (2004). Toward a theory of agent-based systemic situational awareness. In D. A. Vincenzi, M. Mouloua, & P. A. Hancock (Eds.), *Human performance, situation awareness and automation: Current research and trends* (Vol. 1, pp. 83-87). Mahwah, NJ: Erlbaum.

Walker, G. H., Stanton, N. A., & Young, M. S. (2006). The ironies of vehicle feedback in car design. *Ergonomics, 49*(2), 161-179.

Walker, G. H., Stanton, N. A., & Young, M. S. (2007). What's happened to car design? An exploratory study into the effect of 15 years of progress on driver situation awareness. *International Journal of Vehicle Design, 45*(1/2), 266-282.

Wallenius, K. (2005). *Generic support for decision-making in effects-based management of operations.* Doctoral dissertation. Stockholm: Royal Institute of Technology.

Walls, G. L. (1962). The evolutionary history of eye movements. *Vision Research, 2,* 69-80.

Waltz, E., & Llinas, J. (1990). *Multisensor data fusion*. Norwood, MA: Artech House.

Weick, K. E. (1995). *Sensemaking in organizations*. Thousand Oaks, CA: Sage.

Weick, K. E. (2001). *Making sense of the organization*. Oxford, UK: Blackwell.

Whitaker, L. A., & Klein, G. A. (1988). Situation awareness in the virtual world: Situation assessment report. In F. E. Mc Intire (Ed.) *Proceedings of the 11th Symposium of Psychology in the Department of Defense* (pp. 321-325). Colorado Springs, CO: US Air Force Academy.

White, L. J., & Cohen, E. I. (1980). A domain strategy for computer program testing. *IEEE Transactions on Software Engineering, 6*(3). 247-257.

Wickens, C. D. (1992). *Engineering psychology and human performance* (2nd ed.). New York: Harper-Collins Publishers.

Wickens, C. D. (2002). Situation awareness and workload in aviation. *Current Directions in Psychological Science, 11*(4), 128-133.

Willén, B. (1995). *Hur kan ergonomi integreras i konstruktionsprocessen?* (Report No. LiTH-IKP-R-890). Linköping, Sweden: Linköping University. In Swedish.

Wilson, G. F. (1995). Psychophysiological assessment of SA? In D. J. Garland & M. R. Endsley (Eds.), *Proceedings of an International Conference on Experimental Analysis and Measurement of Situation Awareness* (pp. 141-145). Daytona Beach, FL: Embry-Riddle Aeronautical University Press.

Wilson, G. F. (2000). Strategies for psychophysiological assessment of situation awareness. In M. R. Endsley & D. J. Garland (Eds.), *Situation awareness analysis and measurement* (pp. 175-188). Mahwah, NJ: Erlbaum.

Wilson, G. F., Badeau, A., & Gawron, V. J. (1993). Evaluation of two in-flight sensors using heart rate and eye blink measures. In R. S. Jensen & D. Neumeister (Eds.), *Proceedings of the 7th International Symposium on Aviation Psychology* (pp. 886-890). Columbus: Ohio State University.

Wilson, G. F., & Fisher, M. D. (1991). The use of cardiac and eye blink measures to determine flight segment in F4 crews. *Aviation, Space, and Environmental Medicine, 62*(10), 959-962.

Wilson, G. F., Fullenkamp, P., & Davis, I. (1994). Evoked potential, cardiac, blink, and respiration measures of pilot workload in air-to-ground missions. *Aviation, Space, and Environmental Medicine, 65*(2), 100-105.

Wilson, G. F., Lambert, J. D., & Russell, C. A. (2000). Performance enhancement with real-time physiologically controlled adaptive aiding. In *Proceedings of the 14th Triennial Congress of the International Ergonomics Association and the 44th Annual Meeting of the Human Factors and Ergonomics Society* (pp. 3.61-3.64). Santa Monica, CA: Human Factors and Ergonomics Society.

Wilson, G. F., Purvis, B., Skelly, J., Fullenkamp, P., & Davis, I. (1987). Physiological data used to measure pilot workload in actual flight and simulator conditions. In *Proceedings of the Human Factors Society 31st Annual Meeting* (pp. 779-783). Santa Monica, CA: Human Factors Society.

Woltjer, R. (2005). *On how constraints shape action.* Licentiate thesis (No. 1221). Linköping, Sweden: Linköping University.

Wood, J. G., & Hassett, J. (1983). Eyeblinking during problem solving: The effect of problem difficulty and internally vs. externally directed attention. *Psychophysiology, 20*(1), 18-20.

Woods, D. D. (1995). Toward a theoretical base for representation design in the computer medium: Ecological perception and aiding human cognition. In J. M. Flach, P. A. Hancock, J. Caird, & K. Vicente (Eds.), *Global perspectives on the ecology of human-machine systems* (pp. 157-188). Hillsdale, NJ: Erlbaum.

Woods, D. D., & Dekker, S. W. A. (2001). Anticipating the effects of technology change: A new era of dynamics for ergonomics. *Theoretical Issues in Ergonomics Science, 1*(3), 272-282.

Woods, D. D., & Hollnagel, E. (2006). *Joint cognitive systems: Patterns in cognitive systems engineering*. Boca Raton, FL: Taylor & Francis.

Worm, A. (2000). *On control and interaction in complex distributed systems and environments.* Doctoral dissertation (No. 664). Linköping, Sweden: Linköping University.

Worm, A. (2002). An integrated framework for information-centred human-machine systems analysis. *International Journal of Emergency Management, 1*(2), 125-143.

Xiao, Y., Hunter, W. A., Mackenzie, C. F., Jefferies, N. J., & Horst, R. L. (1996). Task complexity in emergency medical care and its implications for team coordination. *Human Factors, 38*(4), 636-645.

Zacharias, G. L., Miao, A. X., Illgen, C. X., Yara, J. M., & Siouris, G. M. (1996). SAMPLE: Situation awareness model for pilot in-the-loop evaluation. In *Proceedings of the First Annual Symposium on Situational Awareness in the Tactical Air Environment.* Patuxent River, MD: Naval Air Warfare Center Aircraft Division.

Zachary, W. W., & Ryder, J. M. (1997). Decision support systems: Integrating decision aiding and decision training. In M. Helander, T. Landauer, & P. Prabhu (Eds.), *Handbook of human-computer interaction* (2nd ed., pp. 1235-1258). Amsterdam: North Holland.

Zon, G. D. R., Alfredson, J., Andersen, H. H. K., Berggren, P., Catchpole, K. R., Dylak, J. et al. (2003). *Full-scale simulator experiment* (Tech. Rep. No. VINTHEC II-WP7-TR01). Amsterdam: National Aerospace Laboratory NLR.