**Situation Awareness and Working Memory: an Integration of an Applied Concept with a
Fundamental Cognitive Process**

By

Kamilla Rún Jóhannsdóttir, B. A.

**A thesis submitted for examination
to the Faculty of Graduate Studies and Research
in partial fulfillment of
the requirements for the degree of**

Doctor of Philosophy

Institute of Cognitive Science

Carleton University

Ottawa, Ontario

July 12, 2004

© Kamilla Rún Jóhannsdóttir, 2004

Library and
Archives Canada

Bibliothèque et
Archives Canada

Published Heritage
Branch

Direction du
Patrimoine de l'édition

395 Wellington Street
Ottawa ON K1A 0N4
Canada

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence

ISBN: 0-612-97828-1

Our file Notre référence

ISBN: 0-612-97828-1

NOTICE:

The author has granted a non-exclusive license allowing Library and Archives Canada to reproduce, publish, archive, preserve, conserve, communicate to the public by telecommunication or on the Internet, loan, distribute and sell theses worldwide, for commercial or non-commercial purposes, in microform, paper, electronic and/or any other formats.

The author retains copyright ownership and moral rights in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

AVIS:

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque et Archives Canada de reproduire, publier, archiver, sauvegarder, conserver, transmettre au public par télécommunication ou par l'Internet, prêter, distribuer et vendre des thèses partout dans le monde, à des fins commerciales ou autres, sur support microforme, papier, électronique et/ou autres formats.

L'auteur conserve la propriété du droit d'auteur et des droits moraux qui protègent cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

In compliance with the Canadian Privacy Act some supporting forms may have been removed from this thesis.

Conformément à la loi canadienne sur la protection de la vie privée, quelques formulaires secondaires ont été enlevés de cette thèse.

While these forms may be included in the document page count, their removal does not represent any loss of content from the thesis.

Bien que ces formulaires aient inclus dans la pagination, il n'y aura aucun contenu manquant.

Canada

Abstract

This dissertation explored the relation between the applied concept of situation awareness (SA) and the basic cognitive mechanism of working memory (WM). SA is an important concept in applied research. It captures how operators perceive, comprehend and anticipate events in dynamic environments. There is substantial evidence both in basic and applied research to suggest that WM might be the cognitive mechanism necessary to support such dynamic comprehension. However, the potential relation between SA and WM is not well articulated. This lack of articulated relation between the two concepts is largely due to how researchers have conceptualized SA. Rather than thinking about SA in terms of processing of information by the underlying cognitive mechanisms, SA is commonly thought of as conscious comprehension that is separate from the underlying processes. The role of WM is therefore of little interest to the study of SA. Alternatively, the existing theoretical frameworks are too vague to fully explain the potential link between SA and WM.

It was hypothesized in this dissertation that SA depends on maintaining and processing information in the different WM systems. An existing WM model was used to establish a conceptual link between the two concepts. The role of WM in supporting SA was then tested empirically. A series of four experiments was conducted using a dual-task paradigm in which tracking tasks were combined with different WM load tasks. In experiment 1 and 2 participants tracked a moving object against a dynamic background. In experiment 3 participants had to periodically predict or search for the tracked object in addition to tracking it. In experiment 4 participants drove in a driving simulator. Their task was to maintain a central lane position (lateral tracking) and maintain SA for information

around the vehicle. Simultaneously to the primary task, participants performed one of four load tasks. Each load task was designed to tap into the different WM systems. In all four experiments, the results provided support for the hypothesis that WM plays an active role in maintaining SA in a dynamic environment. A unique contribution of this research was to show that maintaining SA in a dynamic environment involves selectively the verbal and the spatial WM systems and central executive control. For example, in experiment 4 maintaining SA of surrounding cars required the spatial system for cars in front of the driver and the verbal system for cars in back of the driver. The implication of the results for SA theory and research is explored.

To
My parents
Jóhann Þór Halldórsson og
Auður Eiríksdóttir

Með kærri þökk fyrir ást ykkar og umhyggju
og ómetanlegan stuðning í gegnum árin

Acknowledgements

I have many people to thank for their support and inspiration during the process of writing this thesis. First of all, I would like to thank my supervisor Chris Herdman who has taught me a great deal about the importance of a well presented material and good writing. He has been generous in his support, providing me with great opportunities to present my research at international conferences and to participate in exciting research projects, thus giving me an immeasurable learning experience.

I am especially grateful to Jo-Anne LeFevre for always being ready to put in her time and energy to give helpful comments, providing me with tools to enhance my writing.

I am thankful to Andrew Brook for creating and maintaining a wonderful program and for his interest and enthusiasm for students' work.

I want to thank Lew Stelmach who sadly passed away this year for inspiring me with his enthusiasm for his work and for life in general. Lew taught me a great deal about the process of publishing a paper and was always ready to help in any way he could.

I want to thank my external examiner Malcolm Cook for his lively presence and intriguing questions.

I want to thank my other committee members for their contribution and their helpful feedback.

I am thankful to Colleen Fulton and Lianne Dubreuil for providing a warm and loving environment in the department and for always being there to help.

I am grateful to Jerzy Jarmasz for his friendship and support throughout my studies. Ideas were born and developed in our many conversations about everything between how to do science and how to live life.

I am grateful to Diana DeStefano who has no idea how much she helped me in the last stages of the writing with her positive attitude and strong spirit.

I want to thank the people in my lab. Lisa for being a great friend and for her participation in shaping my thinking and running the experiments. Carol for her friendship and for programming three out of four experiments. Kishore and Dan for programming the driving simulator. In particular, I want to thank Kishore for the countless hours he put into running the last experiment. All of you made it worthwhile to come in to work everyday.

I want to thank Sanjay Chandrasekharan for his editing and all the fun discussions we had about various topics within cognitive science. I also want to thank Jennifer Schellinck for her feedback and for moral support during the last stages of writing this document.

I thank all my friends for their inspiration and support. In particular I am grateful to my good friend Nina Stefánsdóttir for wonderful friendship and support during my studies.

My family who has always been there for me through good times and bad, showing me endless support and believing in me. I am forever grateful for them and their support.

Last but certainly not least, I'm most grateful to my loving partner Gerardo Reynaga for bringing joy and happiness into my life and for reminding me to live when things seemed difficult.

Abstract	ii
Acknowledgements	v
Chapter 1. Introduction	1
Chapter 2. Situation Awareness: A Critical Review.....	5
2.1 The concept.....	5
2.1.1 Product or process?.....	6
2.2 Endsley's Information-Processing Framework.....	9
2.2.1 Three levels of SA	10
2.2.2 The underlying mechanism.....	11
2.2.3 The situation model	13
2.3 Measuring SA according to the product view.....	15
2.3.1 Verbal Protocol.....	18
2.3.2 Subjective assessment.....	19
2.3.3 Knowledge queries	23
2.3.4 Situation manipulation.....	28
2.3.5 Performance assessment	28
2.3.6 Summary.....	30
2.4 Problems with the product view	31
2.5 An alternative approach	32
2.5.1 The perception-action cycle.....	32
2.5.1.1 The underlying mechanisms	34
2.5.2 SA as a skill	36
2.6 Alternative methodology	38
2.6.1 The WOMBAT test	39
2.7 Problems with the alternative view	41
2.8 Empirical findings	42
2.9 WM: A cognitive model for SA	47
2.10 Summary.....	48
Chapter 3: The meaning of "awareness of the situation"	50
3.1 What is SA awareness of?	50
3.2 Explicit vs. implicit knowledge	51
3.3 SA and Expertise	54
3.4 Blindsight, inattention blindness and other evidence from neuroscience	56
3.4.1 Two separate pathways.....	61

3.5 Some philosophical interpretations.....	63
3.6 Back to SA.....	66
Chapter 4. Working memory: outlining the role of the underlying mechanism.....	71
4.1 Working Memory and SA.....	71
4.2 Structure and processing: The multi-component model	72
4.2.1 The phonological loop or the verbal subsystem	74
4.2.2 The visuospatial subsystem	76
4.2.3 Executive control.....	78
4.2.4 The episodic buffer: Connection to long-term memory.....	80
4.3 The role of WM in complex cognition	83
4.3.1 Playing chess	84
4.3.2 Inner speech and action control	85
4.3.3 Text comprehension.....	87
4.3.4 Navigating	88
4.3.5 Dynamic environments	89
4.4 Potential role of WM in maintaining SA	95
Chapter 5. Empirical research.....	97
5.1 Overview	97
5.2 Experiment 1.....	98
5.2.1 Method.....	99
5.2.2 Results	102
5.2.3 Discussion.....	105
5.3 Experiment 2.....	105
5.3.1 Method.....	107
5.3.2 Results	113
5.3.3 Discussion.....	118
5.4 Experiment 3.....	118
5.4.1 Method.....	121
5.4.2 Results	123
5.4.3 Discussion.....	131
5.5 Experiment 4.....	133
5.5.1 Method.....	136
5.5.2 Results	143
5.5.3 Discussion.....	151
5.6 General Discussion	153
5.6.1 SA and the subsystems	155

5.6.2 SA and the central executive	159
5.6.3 Implications for WM research and dynamic environments	163
Chapter 6. Conclusions: Implications for SA theory and research.....	166
6.1 Summary of the dissertation work	166
6.2 Implications for SA theory	168
6.3 The CSE framework	174
6.3.1 The DMM.....	175
6.3.2 The measured constructs.....	179
6.4 Conclusion	181
6.5 Final words	183
References.....	187
Appendix A.....	213
Appendix B	214

List of Figures

Figure 1. Framework of SA (Endsley, 2000, p.6).....	9
Figure 2. The perception action cycle adopted from Neisser 1967 (adapted from Adams et al., 1995). The schema is further divided into implicit (LTS) and explicit (WM) focus	33
Figure 3. The Multi-Component Model of WM (Baddeley, 2002)	73
Figure 4. RMSE for baseline (tracking only) and for tracking with verbal and visuospatial load tasks. The numbers presented in the bars represent the mean RMSE for each condition.	103
Figure 5. Response time for verbal and spatial load tasks for single (load task only) and dual (load task with tracking) conditions. For the verbal task, participants responded to changes in a number. For the spatial task, participants responded to changes in a vertical bar.	105
Figure 6. RMSE for baseline (tracking only) and for tracking with different load tasks.....	115
Figure 7. Response time for verbal and visual discrimination tasks for single (discrimination task only) and dual (discrimination and tracking) conditions.	116
Figure 8. Time between keystrokes for spatial and random tapping tasks for single (tapping only) and dual (tapping and tracking) conditions.	117
Figure 9. RMSE for baseline (tracking only) and tracking with the different load tasks.....	124
Figure 10. RTs for moving the controlled rectangle to the corner where the tracked rectangle reappeared. Prediction: The tracked rectangle reappeared in the lower right corner. Search: The tracked rectangle reappeared in any of the four corners.	126
Figure 11. Response times for verbal and visual discrimination tasks for different primary task conditions. Single conditions: Discrimination task without the tracking. Tracking: Discrimination task performed with tracking. Prediction: Discrimination task perform.....	127
Figure 12. Response times for spatial and random tapping tasks for different primary task conditions. Single conditions: Tapping without the tracking. Tracking: Tapping performed with tracking. Prediction: Tapping performed with the prediction task. Search: Tapping performed with the search task.	128
Figure 13. Percentage of total key presses that deviate from balanced count (equal number of key presses for each key) for spatial and random tapping and different primary task conditions.	130
Figure 14. RMSE for baseline (driving only) and for driving with different load tasks.....	144
Figure 15. Response time for verbal discrimination task for baseline (discrimination task only) and for different load conditions (3 cars vs. 5 cars).....	145
Figure 16. Percentage of total key presses that deviate from balanced count (equal number of key presses for each key) for spatial and random tapping for single (tapping only) and dual (tapping and driving) conditions.	146

Figure 17. Deviation in estimating positions of cars behind the participants' car measured in car lengths for baseline (driving only) and for driving with different load tasks..... 149

Figure 18. Deviation in estimating positions of cars in front of the participants' car measured in car lengths for baseline (driving only) and for driving with different load tasks..... 151

Figure 19. CSE Framework 168

Table 1: <i>ANOVA for prediction and search</i>	125
Table 2: <i>ANOVA for SA of cars in front and back</i>	147
Table 3: <i>ANOVA for SA of cars in back</i>	148
Table 4: <i>ANOVA for SA of cars in front</i>	150

Chapter 1. Introduction

This dissertation explores the relation between situation awareness (SA) and working memory (WM). The former is a concept studied in application contexts like aviation, while the latter is a fundamental cognitive mechanism.

SA is an important concept in aviation research, originating as a description of what makes up the *ace factor*. In aviation, the term “ace” is used in reference to exceptional fighter pilots who were noticeably better at their tasks than average pilots (Spick, 1988). In an attempt to enhance the performance and success of average pilots, researchers and trainers focused on finding out what made up the ace factor, with the intent of providing average pilots with specific, ace factor information during training. Spick (1988) claimed that aces were not better trained, did not necessarily have more experience or better-equipped aircraft, nor could they be distinguished from the average pilot in terms of any personal or physical traits. Rather, aces simply seemed to have better awareness of what was going on. As one pilot noted, in order to perform well in the air, a pilot must have “...a three-dimensional sense of awareness and feel time, distance and relative motion as if they were part of his soul; only if you have a feeling for what is going on around you, you can take action and make correct decisions” (Spick, 1988, p.165). The notion that exceptional awareness leads to exceptional performance has been generalized from the aviation field to performance involving any complex human-machine interaction. Today, it is widely accepted that good SA is required for successful and safe performance in tasks such as driving a vehicle and air traffic control (ATC)

(Endsley, 2000a). In contrast, low SA has been linked to operator error (Jones & Endsley, 1996).

As the interest in operator SA has developed, researchers have articulated SA as a formal psychological construct that can be operationalized and measured (Uhlarik & Comerford, 2002). As a formal construct, SA refers to how operators perceive relevant situational elements, integrate those elements into a coherent mental picture and anticipate future events. Therefore, understanding SA provides an important insight into how a new system, implemented in the cockpit or the automobile, can be best designed to provide optimum support to the human user.

Working Memory (WM) is a well-studied topic in the fundamental research literature. Historically, memory has been broadly classified into short-term storage (STS) and long-term storage (LTS) systems (Atkinson & Shiffrin, 1968), or as James (1890) proposed, primary versus secondary memory. STS was viewed as a limited capacity system for temporarily holding information before it is transferred to LTS. LTS refers to more permanent, although not necessarily invariant, memory. WM includes the traditional STS concept of a temporary storage system. In addition, however, WM is the place where information is integrated, manipulated and possibly recorded. Researchers have shown that ability to activate and maintain sub-goals and intermediate solutions in WM is the key to success in many cognitive tasks. On this view, WM is believed to support the generation and maintenance of representations of complex environments. WM is what allows some chess players to store sub-goals and to plan and anticipate future moves (Robbins et al., 1996). This mechanism is also known to underlie the ability

to track spatial and temporal developments in a story (Friedman & Miyake, 2000).

Temporary maintenance of information from each sentence in WM, allows the reader to connect one sentence to another and to make the necessary inferences.

Given the role of WM in many complex cognitive tasks, it is likely that this mechanism also supports the integration and anticipation of information in dynamic environments. In fact, there is evidence both in basic and applied research suggesting that WM may play a key role in maintaining SA. However, the relationship between SA and WM has not been well articulated. Although WM is a well-researched mechanism, it has not been explored in the context of complex cognitive processing in dynamic environments. In applied research, WM is frequently referred to as an important mechanism supporting SA. But there is no conceptual link between WM and SA, specifying how WM contributes to SA. A major problem seems to be how SA has been conceptualized: the most common approach is to view SA as something that an operator can consciously introspect upon and thus researchers have often used conscious reports as a primary measure of SA (Endsley, 1995a, 1995b). On this view, the cognitive mechanisms that are fundamental to generating and maintaining SA are of little interest. An alternative, but less common approach is to consider SA as a product of fundamental processes underlying human cognition. On this view, an operator's ability to consciously report what they know about a situation may be useful, but must be considered relative to measures concerning how the operator was able to achieve a high-level of performance. The alternative approach recognizes WM as critical to the maintenance of SA. However, similar to the first approach, the role of WM has not been properly specified in this

tradition either. But a conceptual link between SA and WM would provide the needed foundation for further theoretical and empirical progress in the field.

The goal of this dissertation is to provide a conceptual and empirical link between SA and WM. The present work reports a series of experiments, testing the premise that WM subsystems play specific and measurable roles in operator SA. The roles of these WM subsystems depend on the type of task that is being performed and thus the types of cognitive processes that are being activated in WM (Garden et al., 2002; Logie et al., 1989). In chapter 2, conceptual and methodological approaches to SA are discussed and related to the distinction between measuring SA as what an operator knows versus measuring SA as how the operator knows. Chapter 3 addresses the issue of whether the concept of SA must necessarily refer to conscious awareness. In chapter 4, a current model of WM is described and linked to SA. In chapter 5, hypotheses generated from the WM model are examined and linked to SA by way of a series of empirical studies. In the last chapter, the implications of the present research are discussed. Linking the two concepts provides an independent and novel contribution to both basic research on WM and applied theory and research on SA. On the one hand, SA provides WM research with a concept that allows researchers to examine the role of WM in complex cognitive processes that take place in dynamic environments. On the other hand, WM provides a cognitive model for further guiding theory and research on SA.

Chapter 2. Situation Awareness: A Critical Review

2.1 The concept

It has proven difficult to define SA for use in empirical research (Baxter & Bass, 1998; Sarter & Wood, 1991). As a description, SA refers to a wide range of behavioural and cognitive factors. Therefore, definitions of SA often tend to be very general, referring to almost every cognitive process that is relevant to the operator's performance in a complex task environment (Durso & Gronlund, 1999). On the other hand, some researchers have taken the opposite approach by referring to SA in terms of specific task-related knowledge. On this view, SA refers to different knowledge for different situations: What the combat pilot, the commercial pilot, or the car driver requires for good SA is not necessarily the same.

Some researchers have also attempted to capture cognitive activity common to SA across different situations (e.g., Endsley, 2000a). The definition of SA proposed by Endsley (1988) is the most frequently used in the field. Endsley (1988) defined SA as, "...the perception of the elements in the environment within volume of time and space, the comprehension of their meaning and the projection of their status in the future" (p. 97). A similar definition was proposed by Wickens (1995) who described SA as, "...the continuous extraction of environmental information about a system or environment, the integration of this information with previous knowledge to form a coherent picture, and

the use of that picture in directing further perception, anticipating and responding to future events” (p.1). Generally, SA refers to knowing or comprehending in complex, dynamic environments. The reference to dynamic environments distinguishes SA from more stable knowledge structures stored in LTS, such as those used to represent facts or memories of events (Wickens, 2002). SA is the online comprehension of multiple sources of information in an environment, as they change constantly both within space and time. Therefore, an important part of SA is the process of constantly extracting required information from the environment, integrating this information, and interpreting this information within what the operator already knows and has experienced. Good SA also involves the ability to anticipate what may happen next (Regal, William, & Boucek, 1988). This predictive aspect of SA allows the operator to respond quickly and accurately to critical changes in a system and/or a task situation. For example, a fighter pilot who has a good understanding of how the various parameters in a situation relate is more likely to recognize and anticipate potential threats that are developing and thereby be able to respond to these threats effectively.

2.1.1 Product or process?

A distinction has developed in the literature between SA as a product versus SA as an underlying process. According to the *product* view, SA is a product of what the operator knows at a specific point in time. For example, measuring a pilot’s reported knowledge of aircraft location reflects SA as a product (Endsley & Bolstad, 1994). Therefore, researchers who have adopted a product view focus on identifying the type of

knowledge needed for optimum SA rather than on how that knowledge is obtained. Four dimensions of knowledge have been identified as necessary for good SA, namely, *where, what, who and when* (Harwood, Barnett & Wickens, 1988). These dimensions map onto knowing where one is located in space (spatial awareness), what the relevant external elements are (e.g., hazard awareness), who is responsible for a task (e.g., aircrew awareness), and how the situation is evolving over time (temporal awareness) (Aretz, 1991; Endsley, 1995a; 1997; Harwood et al., 1988; Olmos, Liang & Wikens, 1995; Wickens, 1995; 2000; Williams, 2002). Recently, researchers have extended a product view of SA to measure other type of “awareness”. For example, work awareness has been examined to understand and measure how people comprehend their workspace and the responsibility and activity of other workers (Gutwin & Greenberg, 2002). In a product view of SA, awareness effectively refers to the type of knowledge required to carry out a particular task.

In contrast to the product view of SA, some researchers have adopted a *situation assessment* view (Fracker, 1988; Fredrico, 1995; Williams, 2002). The situation assessment view focuses on the processes underlying SA, rather than on the knowledge-based products of SA. An example would be measurement of SA in terms of lateral and vertical tracking errors in a flight path task (Williams, 2002). Researchers tend to use either an SA or a situation assessment approach, not both. Indeed, some researchers (e.g., Fracker, 1988; Fredrico, 1995) have argued that SA can only be explained in terms of a theory of situation assessment.

The obvious problem with distinguishing between SA and situation assessment is that it is not always clear (especially in dynamic environments) where information-

gathering ends and understanding begins (Clark, 1997; Kirsh & Maglio, 1994). The operator's knowledge, for example, is largely determined by constant updating of transient information, selectively attending to what is relevant, ignoring distracters and maintaining task-relevant knowledge (in WM) until that task can be executed.

Understanding the nature of the dynamic knowledge cannot be dissociated from the question of how that knowledge is acquired and maintained. Hence, distinguishing *what* the operator needs to know from the issue of *how* the operator knows may generate an incomplete view of what is involved in the operator's SA. The focus on what an operator knows has often led researchers to ignore the functions of the underlying cognitive mechanisms (Sarter & Wood, 1991; Taylor & Selcon, 1994).

There is an alternative, but rarely recognized approach, whereby emphasis is placed on SA in terms of how the operator acquires and maintains knowledge by the underlying mechanism (Adams, Tenney & Pew, 1995; Dominguez, 1994; Tenney et al., 1992; Salas et al., 1995). On this view, SA is seen as both the product and the processing of information. Understanding SA in terms of how knowledge is acquired involves specifying both the requisite knowledge and the functions of the underlying mechanisms. Ultimately the combination of *what and how* should provide researchers with the ability to predict SA for a given operator across different situations.

Although, the emphasis in the literature is on the product view or measuring *what* an operator knows, there is an underlying tension both in theory and methodology regarding whether researchers should approach SA as *what* the operator knows or *how* the operator acquires and maintains knowledge in a dynamic environment. The literature will be reviewed in the following sections according to this underlying distinction.

2.2 Endsley's Information-Processing Framework

In Endsley's (1995a) approach, SA is added to an existing information-processing model (see Figure 1). The core of Endsley's model (see double arrow) is a simple feedback loop where the state of the environment feeds into the operator's mental picture or SA. SA then feeds into a decision-making mechanism which leads to operator performance. The loop is closed when the operator's performance feeds back into SA by changing the state of the environment.

Figure 1. Framework of SA (Endsley, 2000, p.6).

In Endsley's approach, the operator forms an understanding of the environment by integrating incoming information and existing knowledge into a coherent mental picture of the task situation. It is this mental picture that the operator uses to make decisions and then to perform (Endsley, 2000). Therefore, although the representation of

the situation is complete, the operator can make an error in performance or in judgment based on the information that is available. A good example is an inexperienced operator who understands the situation but who selects a poor strategy for executing a decision. Hence, in Endsley's approach, SA, performance and decision-making form separate stages in the information processing framework. As such, the operator's knowledge is separated from the underlying processes and cognitive mechanisms. In accord with a product view of SA, it is this knowledge that the operator can introspect upon and report.

2.2.1 Three levels of SA

In Endsley's model, SA has three levels. Each level provides information for the next level. Level 1 of SA is the perception of the individual elements in the task environment. For instance, in the cockpit or the automobile, level 1 SA includes the perception of the various indicators such as speed, RPM, or altitude. Level 1 also includes the perception of objects in the environment such as trees or other vehicles. Therefore, if an operator has perceived relevant elements in a situation, then he or she presumably has good level 1 SA. Level 1 SA, however, is necessary but not sufficient for level 2 and level 3 SA.

Level 2 SA refers to how a situation is understood by an operator. That is, how the various elements that are perceived at level 1 are integrated and comprehended. At level 2 SA, the various perceived elements are integrated relative to the operator's goals and the requirements of the task. At level 2, the operator forms a holistic picture of the environment that combines the individual elements with knowledge in long-term store that are relevant to the task.

Level 3 SA refers to the ability to predict the future status of the situational elements. This is the highest level of SA and is based on the input of both level 1 and level 2. Therefore, at level 3 SA, operators understand a situation in a way that allows them to predict possible future events and scenarios.

The following is an example of how the three levels of SA function in Endsley's model. A driver who approaches a narrow bridge and perceives both an oncoming car and the bridge has level 1 SA. However, the driver must also comprehend other factors such as relative distance and speed of both cars and must also take into consideration external factors such as the weather (e.g., if it is raining this means the roads are wet and thus more braking time may be required). This type of understanding and awareness reflects level 2 SA, since it requires an integrated mental picture of the relevant information. Finally, given the current information and the driver's prior knowledge, a driver with good level 3 SA will be able to predict which car will reach the bridge first (see Uhlarik & Comerford, 2002, p. 2 for a similar example). Therefore, an operator who can report the relevant situational elements, but who does not understand how they relate to each other, or what they mean for future development, has incomplete understanding of the situation and as such, incomplete SA. In order to have good SA, all three levels must be attained.

2.2.2 The underlying mechanism

In Endsley's model, SA involves attention, WM, and knowledge in LTS. Attention allows the operator to selectively process information in a complex environment or to divide limited attentional resources between multiple elements in a

task situation. Limited attentional resources may cause an overall lack of SA as the operator does not perceive the relevant information or attends to one aspect of the system while neglecting others. In accord with this view, research has shown that a large percentage of accidents that are classified as a loss of SA are due to attentional failures (Durso & Gronlund, 1998; Hartel, Smith & Prince, 1991; Jones & Endsley, 1996).

Once information is perceived, it is stored in WM, which functions as a medium for activating knowledge in LTS. Over time an operator will gain understanding of the system he or she is operating as well as of the task environment. This knowledge is stored as a mental model in LTS and is the most important factor in determining SA in Endsley's model.

Stored mental models get activated when the operator is faced with familiar situations and are modified and adapted according to incoming information. This process of activating stored mental models allows the operator to form expectations which in turn are beneficial for distinguishing relevant from irrelevant information, thus enhancing information processing. In turn, incoming information may cause the mental model to be modified or may activate a new model. Mental models can also be activated by the operator's task-related goals. With experience, established mental models in LTS will guide the operator's attention to relevant information, resulting in minimum attention load or demands on WM. Therefore, an expert operator uses a mental model to lower resource demands and to possibly bypass the limitations of extensive processing in WM. In contrast, the novice operator has less-developed mental models and is therefore more dependent on WM resources. With expertise, some aspects of performance may become automatic, thereby enhancing overall SA by freeing up attentional resources.

Paradoxically, however, the more automatic the operator's behaviour is, the less he is able to access what he knows about the environment. But reported knowledge is frequently used as a measure of SA.

In summary, Endsley (1995a, 2000a) emphasizes mental models and LTS as the main underlying cognitive mechanism in determining SA. In her framework, WM functions mainly as a medium for activating stored mental models.

2.2.3 The situation model

Endsley emphasizes that SA is not the activated mental model but rather a new representation of the current situation that is formed in WM. Endsley refers to this representation as a *situation model*. The concept of a situation model was proposed by Van Dijk and Kintsch (1983) to explain how people comprehend a written text. Situation models have been used to explain how people maintain and constantly update a representation of a situation as things evolve and change within space and time. The construct is thought to accurately reflect spatial, temporal and causal structure of the situation described in the text, as well as representing object properties and relations (Van Dijk & Kintsch, 1983). Furthermore, situation models combine incoming online information with stored information about similar situations or about the objects and persons described (Zwaan & Radvansky, 1998). The concept is well suited as the underlying structure for SA. Researchers generally agree that a mental model refers to a more abstract knowledge stored in LTS whereas a situation model is the specific instantiation of that knowledge (Durso & Gronlund, 1999; Shebilske et al., 2000). In research on text comprehension, WM is recognized as the critical mechanism whereby

the situation model is formed and maintained (Friedman & Miyake, 2000). Researchers have suggested that the situation model is based on representing information in various subsystems of WM (Friedman & Miyake, 2000; Shah & Miyake, 1996).

The reference to a situation model in relation to SA is interesting. In both cases researchers are trying to explain how information is monitored, integrated and updated in a given situation. There is a key difference, however, between Endsley's notion of a situation model and how this term has been used in research on text comprehension. For Endsley, the situation model is SA. Therefore, the concept of a situation model in her framework refers to what the operator consciously comprehends. Furthermore, Endsley does not explain how the situation model is formed and maintained in WM. On Endsley's view, the situation model is matched to a stored mental model in WM that already has a script for task performance associated with it. Thus, by recognizing a pattern in the environment the operator gains access to stored mental models about similar situations that further determine his or her performance. The activated mental model will also determine the operator's higher levels of SA; understanding and prediction. Endsley argues that the matching of the situation model to the mental model is a key mechanism for maintaining SA in a complex, dynamic environment. This process ultimately allows the operator to bypass WM limits. Therefore Endsley does not specify the potential operations of WM in maintaining the situation model.

Conversely, in the literature on text comprehension, the link between comprehension, the situation model and WM is well laid out. Durso and Gronlund (1999) criticized Endsley's assumption that the concept of a situation model can be equated with SA. Rather, they suggested that the situation model is the underlying

representation maintained in WM that supports SA (see also Mogford, 1997). Viewing the two concepts as referring to the same thing is confusing, as it assumes that the situation model as a representation of the task situation includes only information the operator is aware of.

In sum, Endsley's framework has been very influential in the field and it is commonly used in SA research. Her emphasis on SA as three stages of conscious comprehension or knowledge has had huge impact on research focus and methods developed in the field. However, Endsley focuses more on what the operator knows rather than how the operator gets there. She emphasizes stored knowledge structures as the main source of operator's SA. But the actual role of the underlying cognitive mechanism in maintaining SA is not well specified.

2.3 Measuring SA according to the product view

Research on SA has been shaped almost entirely by the product view. Hence, researchers have relied on measures that tap directly into the operator's state of knowledge about a situation. Measures of SA can be divided into five broad categories; verbal protocol, subjective assessment, awareness queries, situation manipulations and performance assessment (Herdman & LeFevre, 2003; Pew, 2000; see also Endsley, 1995a; Fracker, 1991 and Vidulich, 2000 for reviews of SA measures). These measures typically depend on operators' self-reports of information that is accessible to them or on having operators subjectively evaluate (e.g., using rating scale) their own behaviour and understanding of the situation. It is difficult to assess how reliable or valid these subjective measures are. However, both Fracker (1991) and Uhlarik and Comerford

(2002) reviewed the existing measures in light of their actual (tested) or potential reliability and validity.

A measure is reliable if it produces consistent results on different occasions. The validity of a measure is harder to judge. There are four types of validity; content, face, predictive (criterion) and concept validity. Content validity refers to how well the measure represents the particular task domain or knowledge that it is supposed to measure. Therefore, high content validity in one task situation may mean low content validity for another task situation. Face validity refers to whether the measure seems appropriate for its purpose. For example, is the measure capturing SA rather than workload, or how much cognitive resources are required for performing a particular task? Most SA measures are designed according to the assumption that SA is the operator's explicit knowledge that can be reported. As such, most of the measures have face validity. Also, with the focus on the type of knowledge required in different task situations, the measures are likely to have some content validity. Criterion or predictive validity refers to how well the measure correlates with some externally defined criterion. That is, a measure indicating high SA should correlate with superior performance or general success in other relevant aspects of the task. A measure with high predictive validity should thus predict performance in real-world settings (Uhlarik & Comerford, 2002). The problem with assessing criterion validity for SA measures is that SA is assumed to be only indirectly related to performance (Endsley, 1995a). As such, good SA may or may not lead to good performance. This conundrum leaves researchers in a position where they must measure how a particular design affects operators' SA but they are unable to determine whether the design therefore causes optimum performance.

Hence, measuring SA may provide an incomplete view of how new systems affect the human operator.

Finally, construct validity refers to whether the measure is actually measuring the psychological attribute of interest. With a lack of clear-cut definition of what SA is it is difficult to determine the construct validity of the existing measures. A particularly important issue in determining construct validity is the distinction between reflective and momentary awareness (Fracker, 1991). Momentary awareness is what the operator is aware of at any given point in time. Reflective awareness refers to making inferences based on general knowledge and experience about what happened at an earlier point in time. The operator may answer a particular question based not on what was his or her awareness at that moment but rather on inferring what was likely to have been given what the operator knows about that particular task. As Fracker points out, the operator may be aware in the sense that he or she can respond to events but may not recall them when asked. The fact that the operator does not recall an event does not mean the operator was not aware of the event when it took place or when (if) the information was needed. Confounding reflective and momentary knowledge is particularly a problem for more abstract questions regarding protocols, goals or more general knowledge. The operator may be likely to answer based on what he or she knows about the protocols for that situation rather than report accurately on what he or she was aware of at that particular moment in time. The issue here really has to do with whether we take SA to be what the operator explicitly knows and can access.

Other criteria that should be considered when examining measures of SA include sensitivity, selectivity, diagnosticity, and obtrusiveness (Uhlarik & Comerford, 2002).

Sensitivity refers to how well a measure distinguishes between different experimental manipulations and conditions. Selectivity refers to how well the measure captures only changes in the construct of interest. For example, does the measure capture changes in SA but not in workload? Diagnosticity refers to not only how well the measure captures changes in SA across conditions but also identifies the cause of the change in SA. Most SA measures are not diagnostic as they are designed to capture what the operator knows or does not know, but not why the operator knows. Finally, it is important that the SA measure does not interfere with the primary task. Many SA measures are intrusive as they are based on asking the operator questions about their SA during performance on a task.

2.3.1 Verbal Protocol

This technique relies on information recorded by the operator immediately after completing a task or during a video replay of the task scenario. The operator is asked to think aloud as he or she follows the activities in the task scenario. In essence, the operator verbalizes what is happening, what activities are taking place and why. How informative this technique is depends on the task. This method has been used extensively in basic cognitive research to understand peoples' strategies in abstract problem solving (Ericsson & Simon, 1984). However, verbal protocols are rarely used in SA research. In fact, it is not clear that having the operator verbalizing about what is going on in the situation captures accurately the underlying representation of knowledge in a dynamic, fast changing environment. Sullivan and Blackman (1991) found no difference in the nature of verbal protocols for expert and novice pilots. Also, people are often poor at accurately reporting their own mental processes (Nisbett & Wilson, 1977).

2.3.2 Subjective assessment

Subjective assessment is frequently used as a method to evaluate the operator's SA (Endsley, 1995b; Jones, 2000, Taylor & Selcon, 1991). Subjective assessments are obtained by having operators generate ratings of their SA or by gathering ratings from experienced observers, instructors, or supervisors regarding the operator's SA (observers' assessments). In general, subjective assessments are easy and inexpensive to use (Endsley, 1995b). However, there are some obvious limitations to asking operators to assess their own SA. Self-assessment methods assume that the operators can judge accurately what they experience, feel and think and are able to distinguish facts from experiences (Jones, 2000). It has been shown, however, that self-ratings can depend on the outcome of performance (Sarter & Woods, 1991). For example, if performance is good regardless of SA, the operators may tend to rate their SA as good: Venturion, Hamilton and Dvorchak (1989) found that there was a high correlation between pilots' rated SA and mission success. Also operators may fail to realize that they are unaware of information when this information could be critical to their success. For example, Fracker and Vidulich (1991) tested how different displays supported pilots' spatial orientation. They found that the display that was rated as providing the best SA by the pilots was also the one that resulted in the highest numbers of inverted recovery errors (where pilots believed they were upright when they were in fact inverted). Thus subjective measures should be used in combination with other measures.

Self-assessment can be useful in providing information on how comfortable and confident an operator is in a situation. If used in combination with other measures,

subjective measures can be informative. At times, a misalignment between self-rated confidence in their SA and their actual performance can reveal operator's false confidence with a system.

Various subjective rating scales have been created for subjective assessments of SA, including the Situation Awareness Rating Technique (SART) and Situation Awareness-Subjective Workload Dominance (SA-SWORD) (see Jones, 2000 for a further discussion of SA rating scales). These scales fall into two distinct categories: direct versus comparative ratings. Direct ratings require the observers to rate their SA according to a Likert (or similar) scale (e.g., ranging from 1 to 7). This type of rating method is absolute in the sense that operators rate different conditions independently using abstract scales (Vidulich, 1989). The direct rating method assumes a one-to-one relation between the rating and some absolute value of SA. The comparative rating scale offers a different method where the operator compares SA in one mission to SA during another mission. Comparative SA evaluations are therefore based on difference of ratio estimates rather than correlating SA to a fixed point on a scale (Uhlarik & Comerford, 2002).

SART is one of the best-known direct self-rating scales for evaluating SA. SART is based on the assumption that SA is a multidimensional concept and hence it includes multiple scales (Taylor, 1989; Taylor, 1990; Taylor & Selcon, 1991). The scales are organized according to three main dimensions; attentional demand, attentional supply, and understanding. Each of these three dimensions includes other concepts. Attentional demand is further divided into instability of situation, complexity of situation, and variability of situation. Supply of attentional resources refers to arousal, concentration of

attention and division of attention. Understanding refers to information quantity, information quality and familiarity (Taylor, 1990). Hence, a SART rating can be based on as few as three or as many as ten dimensions. Selcon and Taylor (1989) found that the three and ten dimension scales essentially capture the same information.

The SART approach was developed by asking experienced aircrew which elements they considered essential for good SA. Hence, the SART scale has some ecological validity. Furthermore, the scale is considered useful as it refers to various cognitive and perceptual processes such as attention and WM resources. As such, SART may provide some insight into different components of SA. However, as with any other subjective rating scale, the operators' evaluation of their own SA using SART may or may not be accurate and it does not tap into the cognitive processes underlying SA. SART has been found to have moderate reliability (Waag & Houck, 1994). However, SART has been criticized for confounding SA with workload (Uhlarik & Comerford, 2002). The focus on issues such as supply and demand on resources has made researchers in the field question whether SART in fact measures SA. For example, it has been argued that the correlation between SART ratings and performance does not reflect SA but rather general processing and workload issues (Jones, 2000). Taylor (1990) claimed that maintaining SA adds to the overall workload and has to do with processing demands and available resources (see also Adams et al., 1995; Tenney et al., 1992). On the other hand, Selcon and Taylor (1989) found the SART scale to be sensitive to issues such as pilot's experience whereas the NASA TLX workload scale (Hart & Staveland, 1988) was not. The findings by Selcon and Taylor suggest that the two scales were measuring something different.

Fracker (1991) argued that SART does not have good content validity as it was based on specific experiences of pilots and as such may not generalize to other situations. The SART scale however, is sensitive to experimental manipulations. For example, Endsley (1998) found the SART method to be more sensitive to capturing differences in different conditions than performance measures and Taylor and Selcon (1989) found the scale to be more sensitive than a single overall score of SA. Although SART does not necessarily predict operator performance, Jones (2000) claims that SART does have some level of diagnosticity. The individual dimensions of the scale provide some information regarding possible causes of SA. That is, the scale may show that in a particular condition the operator experiences overload for attentional resources or lack of understanding etc. The necessity or sufficiency of the subscales for SART, however, have not been thoroughly evaluated.

SA-SWORD (Situation Awareness-Subjective Workload Dominance technique) is a comparative rating technique based on a workload measure that originally allowed subjects to rate one design concept in relation to another along a continuum, to evaluate the degree to which one concept entails less workload than another (Vidulich, 1989; Vidulich & Hughes, 1991). This metric was modified to allow for the rating of the extent to which a different system design concept reflects good SA (Endsley, 1995b). With SA-SWORD, operators are exposed to different designs and then rate their SA with each condition compared to their SA in the other conditions. It is unclear how effective SA-SWORD is as the scale has not been thoroughly tested for evaluating SA (Jones, 2000). It is also likely that the SA-SWORD scale provides a limited picture of SA. No validity or reliability has been established for the scale, although there are some indications that the

scale may have some construct validity but the evidence is weak (Fracker & Davis, 1990).

The problem with SA-SWORD is that it is limited to a within-subject design as the rating is based on having the operator compare SA across different conditions. Vidulich and Huges (1991) found that the SA-SWORD scale was sensitive to variations in a cockpit display that was supposed to enhance SA by providing more information. However, the scale was not sensitive to different segments of the flight mission. Further analysis showed that the problem was due to the fact that different pilots rated their SA differently. Half of the pilots rated their SA according to how much information they had to attend to while the other half rated their SA according to the amount of information they thought they missed. It is therefore important to make sure the operators are rating their SA according to a common definition of the concept. Endsley (1995b) also argued that the sensitivity of the SA-SWORD scale may reveal the preference of the operators as they are basing their rating on comparison between different systems, rather than tapping into SA.

2.3.3 Knowledge queries

Knowledge queries involve asking the operators questions regarding their perception of situational elements and overall understanding of the situation. Knowledge queries are considered to be the most accurate measure of SA as this method taps directly into the operator's knowledge of the specific situation (Endsley, 1995b; Pew, 2000; Vidulich, 2000). The query method is thought to be objective as the operator's answers are compared against what is actually going on in the situation (Endsley, 1995b; 2000b).

Fracker (1991) noted, however, that knowledge queries are still subjective measures as they rely on having the operator self-report on information that is momentarily in memory. Knowledge queries, often referred to as memory probes, can be divided into post-test, online queries and the freezing technique.

With post-test queries, operators are asked questions regarding their perception and comprehension of the situation after the completion of each experimental trial. The benefit of the post-test technique is that it does not interrupt the operator during task performance. One problem, however, is that sometimes the operator may not remember correctly what happened or may tend to over-generalize based on his or her experience in the last part of the trial. A particular problem with the post-test method is the confounding of reflective and momentary knowledge (Fracker, 1991). Longer periods between what is experienced and reporting about that experience may result in people relying more on general knowledge about the situation than what actually took place. Also, much of the information the operator is aware of at any given point in time may not be registered in memory and thus may be forgotten. The post-test method is therefore unlikely to provide a thorough picture of the operator's SA throughout performance on a task. There are no studies reporting reliability of the posttest method in SA research. The method is also considered to have low construct validity as it does not adequately distinguish between the operator's momentary and reflective awareness. Furthermore, the test does not have much diagnosticity or predictive value.

Another possibility is to ask the operator questions regarding their perception and comprehension of the environment during their task performance. The benefit of using the on-line query is that the operator reports accurate knowledge for any given point in

time and does not have to refer back in time to recall the past status of the system and/or the environment. However, the on-line query method is limited in that it adds to the operator's workload and as such may alter the operator's performance and perception of SA at that time (Endsley, 1995b). Furthermore, on-line queries may cue operators to attend to certain factors which they might not have been aware of otherwise, giving a false picture of their SA. Similar to the post-test, on-line queries may not offer much in terms of diagnosticity and this method has not been assessed for predictive validity.

In the freeze technique the operators' task is randomly frozen at various points in time. For example, a display screen might be blanked and an operator asked various questions regarding their current SA. This technique is better known as the Situation Awareness Global Assessment Technique (SAGAT) and was proposed by Endsley (1995b). The operators' answers are then compared to what was actually occurring in the task environment at the time the task was frozen. SAGAT is based on Endsley's definition of SA as perception, comprehension, and projection into the future. Hence, the questions asked during the freezing of the task refer to all three levels of SA. For example, in the case of using SAGAT to evaluate the SA of pilots, questions may include status of airspeed (level 1 SA), predicting future status of elements (level 3 SA), and how the various situational elements are integrated (level 2 SA) (see also Endsley, 2000b).

SAGAT is often considered to be the best SA measure that is available (Endsley, 1995; Pew, 2000; Vidulich, 2000). This is because it taps directly into SA at different points in time and therefore is not influenced by issues such as operator memory loss that are inherent in the post-test method. The fact that the task is frozen means that the operator is not being subjected to changes in workload while performing their task.

Finally, SAGAT is considered to be partially objective and as such it does not have the same problems as the subjective assessment. Accordingly, the SAGAT method has proven to be a useful measure for evaluating what the operator knows and how the operator comprehends the various elements in the task situation at different points in time. However, the SAGAT method has several limitations. For example, stopping a task or simulation may disrupt ongoing task performance and memory and create an unrealistic task environment (Pew, 2000). Indeed, freezing the simulation and then asking operators various questions about certain aspects of the situation may disturb the phenomenon that is being studied (Sarter & Wood, 1995, p.17). The data provided by SAGAT may therefore be an artifact of the method itself as it creates expectations for the operators: the operators know that they will be questioned and thus may alter their behaviour. Furthermore, because the simulation is stopped at random, a measure of SA at specific and critical points in a task may not be obtained (Pew, 2000). Also, as Selcon, Taylor and Koritsas (1991) noted, the freezing technique is based on the assumption that the operator's knowledge is stored and accessed in memory. However, some knowledge may be used without awareness and as such is not captured by the SAGAT.

Researchers generally agree that the knowledge query method, in particular SAGAT, have good construct validity (Endsley, 1995b; Wickens 1992). The measure is designed to capture what SA actually is –the operator's explicit knowledge about the task environment. In general, knowledge queries are designed to capture the three levels, of SA: perception, comprehension and prediction. As such, techniques like the SAGAT have good content validity and work has been devoted to adapting the SAGAT to

different task environments (Endsley, 1995b; 2000b). It is not clear, however, whether the SAGAT method confounds reflective and momentary awareness, which would question the construct validity of the measure. That is, even though the task/simulation is frozen during task performance there is no guarantee that the operator is not answering at least some of the SAGAT questions according to inferences based on prior knowledge and past experience. Confounding reflective and momentary awareness would be more of a problem for certain types of questions. For example, Gugerty (1997) used SAGAT to measure SA in a car simulation. The simulation was frozen at random and the participants had to mark the locations of the surrounding cars using a grid map of the simulated driving environment. This type of assessment is not likely to tap into inferred knowledge. However, asking the driver to interpret the situation (i.e., what is going on), is more likely to tap into general knowledge based on experience and prior learning.

In general, the knowledge query technique is limited insofar as it tells us about the knowledge the operator is aware of at any given point in time but not much about how that knowledge will be used or whether it will be used. For example, the method does not tell us much about how knowledge maps onto performance. This is a crucial limitation. Also, the knowledge query method is not very predictive of the operator's performance. Although SAGAT scores have been shown to correlate with performance (Endsley, 1990, 2000b), this correlation may depend on how SA is defined and measured. For example, Endsley (1990) found that awareness of aircraft location as measured by SAGAT correlated with number of hits in a simulated air-to-air battle. However, Vidulich et al. (1994) found that SAGAT failed to distinguish between the effectiveness of different cockpit display conditions. Similarly, Durso et al. (1995) found that memory probes were

not very sensitive in distinguishing different conditions for chess players except for questions regarding future moves. The method has limited diagnosticity as it does not allow insight into the cause of the gained or lack of SA reported.

2.3.4 Situation manipulation

Situation manipulation is based on the assumption that certain performance is largely or solely driven by SA, and therefore indicates the status of the operator's SA (Pew, 2000). In this method, certain information is artificially changed or removed from a display during a simulation of a task environment. The time it takes the operator to react to the missing information is then measured. The limitation of this method is that manipulating the display and measuring responses is based on the assumption that the operator will react to the missing information. However, there is no guarantee that the operator will react to the sudden disappearance of information from the display. The operator could assume a malfunction in the system or decide not to react for other valid reasons, or the information that has disappeared may not be considered critical by the operator and hence it may be ignored. Situation manipulation however, could be useful to study performance when the environment (the system) is manipulated systematically (see Adams et al., 1995). Systematic manipulation might provide researchers with an important insight of when and how SA brakes down.

2.3.5 Performance assessment

Performance assessment is often considered an implicit measure of SA as opposed to explicit measures that tap directly into the state of SA. The claim is that SA can only

be inferred indirectly from performance measures. Therefore, researchers have emphasized that it is important to create an environment where the observed performance must indicate the operator's SA (Pew, 2000). In general though researchers caution the use of performance measures, although, they are the most frequently used method in the field (Vidulich, 2000).

Global task measures assess the overall performance during a mission or on a task (Endsley, 1995b). This method is rarely used, as it is hard to interpret. For example, an important difference in some aspects of performance for different experimental conditions may not be revealed when a global task measure is used (Endsley, 1995b). The embedded task measure method therefore is more common, where performance on an individual task is assumed to reflect SA for that particular task. Williams (2002) for example, measured how much pilots deviated from a defined path as an indication of spatial SA. This use of performance measures raises the point of whether path maintenance indirectly assesses some explicit knowledge that could be more directly revealed through knowledge queries. It seems more likely that path maintenance is a direct measure of maintaining SA of the path, and has little to do with what the operator can report. But again the issue here is how SA is conceptualized. Gugerty and Tirre (1995) and Gugerty (1997) found a high correlation between explicit and implicit measures of SA in a driving simulation. That does not mean that SA should be thought of as being exclusively what the operator can explicitly report. Maintaining awareness may involve a broad representation of information and knowledge that sometimes is better tapped into by using performance measures and sometimes by using knowledge queries, or both.

Fracker (1991) argued that performance measures have more construct validity than knowledge queries, as they do not confound reflective and momentary knowledge. Performance measures are also better than many of the explicit measures in that they are objective and unobtrusive. Furthermore, performance measures have more predictive validity than the knowledge queries (Pritchett, Hansman & Johnson, 1995). Knowledge queries measure what the operator knows but not how the operator uses that knowledge. The operator for example, may not act on his or her knowledge about the situation due to lack of confidence. Performance measures however, provide information on how reliable operator's SA is by measuring how the operator uses his or her knowledge.

2.3.6 Summary

There is a tension in the literature between whether SA should be measured as *what* an operator knows versus *how* the operator knows. The approach that one adopts will be reflected in the methodology that is used. Clearly, measuring what the operator knows faces some serious limitations. These measures are generally not diagnostic nor do they have much predictive power. A great concern here is that measures such as knowledge queries confound reflective and momentary awareness. Furthermore, what the operator knows (or is aware of) may not always be what the operator can articulate. The operator's own assessment of what is going on is also prone to errors. Operators may not remember what took place or they may overestimate their SA. Thus, it is unlikely that measuring what the operator knows and can report provides a complete assessment of the human-machine interaction in question.

Vidulich (2000) reviewed the various studies in the SA literature and found that performance measures and subjective questionnaires were sensitive to system manipulations meant to affect SA. This is true regardless of whether performance measures are defined as measuring SA or not. In contrast, memory probes showed less sensitivity to system manipulation. However, wide scope questions were more sensitive than narrow scope questions (Vidulich, 2000). In general, the existing measures do not provide information on how knowledge in a dynamic environment is maintained and acquired. The focus on SA as the operator's conscious comprehension is a very narrow approach to understanding knowledge in a dynamic environment. It is not clear that what the operator knows is necessarily what the operator can report. Furthermore, measuring what the operator can report provides little information on how knowledge is acquired and maintained. Thus, the view that SA is what the operator knows and can report may not best reflect on the operator's cognitive processing and comprehension in dynamic environments.

2.4 Problems with the product view

On the product view to SA, researchers are essentially trying to understand what the operator knows under different conditions and how that knowledge relates to performance. But there is no specification of how the underlying cognitive mechanisms such as WM might function to acquire and maintain the requisite knowledge. Endsley seems to view WM in terms of a temporary storage rather than in terms of online processing and maintenance of information to carry out a task. This interpretation is in stark contrast to the accepted view of WM in basic cognitive research (Miyake & Shah,

1999). Endsley's framework therefore leaves many questions unanswered or even unasked, such as, how information from the environment and WM is integrated. What is the relationship between SA or the situation model, WM and the three levels of perception, comprehension and projection? It seems that attempts to answer those questions are essential for further theoretical progress in the field.

Without an understanding of the functions of the underlying cognitive mechanisms it is hard to arrive at a model that would allow researchers to predict operator SA. Ultimately, such a model would identify the operations needed to support SA and would portray how the different stages (perception, comprehension and projection) are processed and how they interact to support SA (Adams & Pew, 1992; Garland, Stein, Wise & Blanchard, 1993; Hopkin, 1980; Sarter & Wood, 1995; Shebilske, Goettl & Garland, 2000).

2.5 An alternative approach

An alternative approach to the product view of SA involves attempting to model the functions of the underlying mechanisms and processes that support SA in a dynamic, multi-tasked environment. On this view, SA is considered as both a product and as a process.

2.5.1 The perception-action cycle

Adams et al. (1995) incorporate the notion of WM into the perception/action cycle proposed by Neisser (1967). On this view, SA is an ongoing cycle of perceived information and upgrading of the operator's mental picture (see Figure 2). Activated

schemas in WM cause exploration in the environment, directing attention to the right situational elements at the right levels of abstraction. The exploration may in turn modify the activated schemas causing further exploration. Available information may also cause new schemas to be activated. Therefore, SA in this model refers to both the activated schema (the product) and the information processing activity (the processes). It is the continuous update and maintenance of a representation of the dynamic situation where stored schemas, incoming information and environmental structures closely interact. Adams et al. argue that distinguishing between the product and the underlying processes is rarely tenable in psychological research. Rather, the two are interconnected where the underlying processes determine SA and in turn the operator's SA determines the processes of assessing the situation.

Figure 2. The perception action cycle adopted from Neisser 1967 (adapted from Adams et al., 1995). The schema is further divided into implicit (LTS) and explicit (WM) focus

2.5.1.1 The underlying mechanisms

Adams et al. (1995) focused entirely on the functions of WM in maintaining SA. They expanded Neisser's model by adding the elements of explicit and implicit focus based on the work by Stanford and Garrod (1981) that explored comprehending events through story reading. The explicit focus refers to the online processing and maintenance of information in WM. The implicit focus refers to knowledge in LTS that is accessible but not currently activated. The inactive part of LTS is further divided into episodic and semantic memory. Episodic memory refers to the complete record of knowledge structure gathered over the course of operating a system. Semantic memory refers to the person's lifetime of knowledge in general. This information is particularly costly to bring into active memory.

For each task, WM (the explicit focus) contains pointers to a much larger mental model stored in LTS (the implicit focus). The pointers refer to quickly retrievable information about similar situations or systems and about causal and functional workings of the task environment. This is similar to the idea of long-term WM (LT-WM) proposed by Ericsson and Kintsch (1995). LT-WM is not limited by a fixed capacity. Rather, it contains pointers to task relevant knowledge and skills stored in LTS. On this view, operators' SA at any point in time will include knowledge currently active and maintained in WM but it will also include relevant knowledge in LTS that has to do with the operator's experience and skills.

In a complex environment, the operator must prioritize tasks, know what task to perform and when, maintain task-related information in WM until the task can be executed, selectively attend to information and so on. Understanding how the operator

does this is important to understanding SA. Adams et al. argue that WM carries information relevant to each task at hand. WM also carries pointers to a larger task scheduling structure in LTS, allowing the operator to know what tasks to perform in what order. Furthermore, WM may also represent pointers to the relevant task-related information allowing the operator to go back to performing a task if interrupted. Otherwise interruption of a task might cause the operator to forget where in the previous task he or she was. Each task has sub steps, therefore there is much task-related information that needs to be kept in active memory. It may therefore be hard to go back to a task that has been interrupted unless the operator has pointers to some sort of a task scheduling in the implicit focus. This can help the operator to reactivate information related to a previous task. Incoming tasks might be ignored or performed depending on its priority as represented in the more general task schema in LTS. The activated schema also determines what information is processed and attended to. That is, information that is related to the task is more easily assimilated with the current SA of the operator than unrelated information.

In addition to carrying pointers to activated schema in LTS, WM also controls and selects action, adapts hypothesis and evaluates consequences. Hence, WM essentially is the mechanism where information is integrated, manipulated and maintained to support SA. Therefore, maintaining SA, even for just a single demanding task, may require considerable WM resources. The idea of the explicit focus being responsible for managing resources and selecting and controlling action, sounds very similar to the multiple component model of WM proposed by Baddeley and Hitch (1974; Baddeley, 1986), particularly the central executive. In their model, the central executive is involved

in coordinating and integrating information, focusing and switching attention and activating representations within LTS. The central executive also integrates information from the two slave systems; the phonological loop and the visuospatial sketchpad, which serve as a temporary storage for task relevant information.

In sum, Adams et al. (1995) proposed a more dynamic approach than Endsley where SA is both the product and processing of information. Their focus is on understanding how SA is maintained and acquired by outlining the functions of WM.

2.5.2 SA as a skill

Hartman and Secrist (1991) took a slightly different approach from Adams et al. (1995) in explaining how pilots maintain SA while in air-to-air combat. They reviewed the literature in order to identify which cognitive factors are important to maintain SA while in combat. They found that good SA of a superior pilot is largely based on: “a) accurate acquisition of performance-critical information, even when the cues are of low intensity and short duration or are embedded in an excessively rich stimulus environment; b) rapid, highly efficient processing; c) exceptional application of trend prediction and anticipation, to include anticipation of adversarial moves; and d) selection of a response optimized on an internal risk/benefit metric” (p. 1085). That is, maintaining SA in a fast and dynamic environment such as air-to-air combat is largely based on skilled management and integration of information and performance. Hence, understanding what the pilot knows at a different point in time may not prove to be all that informative.

With experience an operator will establish important links between certain stimuli in the environment and effective performance. The skilled operator will also learn to use

memory rapidly and efficiently in dynamic, fast changing situations. Rapid and effortless use of memory allows the operator to quickly analyze pieces of information and integrate them into the larger picture. Hartman and Secrist (1991) point out that skilled operators (e.g., expert pilots) can be affected by information that is below their level of awareness. As research has shown, near threshold information may impact performance as well as perception and even comprehension of the situation. Hence, SA for the ace fighter pilot is largely determined by skilled and automatic processing and integration of information. Accordingly, it should be possible to train pilots to acquire superior SA. Secrist and Hartman (1993) decided to test this claim by training participants (10 college students who were members of the Air Force Reserve Officers' Training Corps AFROTC) on early target detection. In their study they included three successively more difficult protocols, where the subjects detected whether the target was present or not (detection), whether it was the target or one of the five non-targets (recognition) and finally which of four possible targets it was (identification). They trained the subjects in responding to targets using long (67 ms) or short (33 ms) viewing time. The subjects quickly exceeded 95% accuracy when viewing the targets for 67 ms for all three protocols. Interestingly when viewing the target for 33 ms, subjects could be trained in recognizing the target as quickly and accurately as detecting them. However, identification was slower and harder to train at 33 ms viewing time than both detection and recognition.

The fact that participants could be trained to recognize the target as quickly and accurately as merely detecting it is important. Kass, Herschler and Companion (1991) also managed to train civilians in recognizing different targets which further improved their SA. In general, Hartman and Secrist (1991) raise the point that SA is more than just

what the operator knows. This is particularly true in a dynamic environment where changes are quick and constant. SA may be better explained in terms of how knowledge about the environment is acquired and used.

2.6 Alternative methodology

Adams et al. (1995) argue that SA is best measured by using online assessment of performance. As such, they are not proposing new measures but rather emphasizing certain types of measures over others. They point out that performance measures can be used along with knowledge queries but knowledge queries should not be used on their own. Furthermore, it is critical to systematically manipulate the environment in order to answer certain questions regarding the maintenance of SA. For instance, the researcher may add subtle information to the display and measure how quickly and accurately the operator responds. Another manipulation is to blank a display or perhaps alter the course and measure how well the operator finds his or her way back. Introducing an interrupting task also allows the researcher to see how easy it is for the operator to divide his or her attention between different tasks. The operator can furthermore be put in a situation where information is only available upon request.

By manipulating the environment, researchers are in a better position to make inferences regarding how information is gathered, how easy it is to share attention between two tasks and about what information is needed/used and when. In the product approach, measures are generally not thought of in terms of systematically altering the environment over and above the system that is being tested. This lack of systematic measures is largely because researchers are not trying to understand how SA is acquired

but what the operator knows. In contrast Adams et al. are trying to guide measurements in a way that provides information on the relevant cognitive processes. Obviously further theoretical guidance is needed where the environment is systematically altered in order to test some hypothesis regarding how SA is maintained.

2.6.1 The WOMBAT test

The WOMBAT Situational Awareness and Stress Tolerance Test was proposed by Roscoe (1993; Roscoe & Corl, 1987; Roscoe, Corl, & LaRoche, 1997). Although the test is not associated with any specific approach, it fits well with the alternative view. The test was designed to understand how participants can selectively attend to, share attention, and prioritize tasks in a dynamic, multitasked environment. In WOMBAT participants are placed in front of a console containing two joysticks and 13-button keypad with the number 0-9, up and down arrow key, and a button labeled “bonus” allowing the participants to engage in bonus tasks. The primary task is a pursuit-tracking task. With their left hand participants control two vertical lines at each side of a hexagon drawn on the computer screen. Their task is to make sure the lines are kept at both sides of the hexagon as it contracts and expands. With their right hand participants make sure that the center of a cross is inside a moving circle at all times. The tracking task gives points for how well the participants are tracking, error also means lost points. The two joysticks operate in one of two control modes. In velocity mode the targets are stationary when the joystick is centered and their speed is proportional to the movement of the joystick. In acceleration mode the targets move at a constant rate when the joystick is centered and the force of control input determines the change in rate of movement for the

target. If tracking is within error limits participants can engage in auto-tracking performance. However, the auto tracking function is prone to failure and therefore the participants must monitor the tracking objects at all times and decide when to disengage the auto-tracking function and gain control. The penalty depends on the severity of the error in tracking. Participants must evaluate when it is worth using the auto-tracking mode, when the errors are acceptable, and whether they need to focus on one of the bonus tasks or not.

Simultaneous to the tracking task, participants are instructed to engage in one of three sub tasks or bonus tasks as frequently as possible. The bonus tasks give extra points but the participants can disengage from performing them without losing points. The three bonus tasks are spatial orienting, quadrant locating, and digit cancellation. For the spatial orientation task, participants are presented with two 3D figures (from Shepard & Metzler, 1971). The task is to figure out as quickly as possible if the two figures are identical, mirror images or otherwise different by manipulating both or one of them using the joysticks. Participants are also asked to indicate how confident they are in the answer they gave. A correct answer where the participant was uncertain gives fewer points than correct answer where the participant was certain in his or her choice. The digit cancellation task has participants monitor digits that are placed at several second intervals. After three digits have been presented the participant must press the button on the keypad corresponding to the digit displayed two back in the sequence. Finally, in the quadrant location task, the numbers from 1 to 32 are placed in each of 4 quadrants. The participants' task is to find the numbers in ascending order and cancel them by pushing the number corresponding to the number on the quadrant (see Roscoe, et al., 1997).

The stark contrast between the WOMBAT test and for example, SAGAT is interesting. SAGAT is based on the idea that SA must be measured by capturing the operators' conscious comprehension. In the WOMBAT test however, the emphasis is on how SA is maintained in a multitask environment. The test has been found to predict exceptional pilot performance (O'Hare, 1997). Furthermore, the WOMBAT test does not correlate highly with performance on individual tasks, suggesting that it is capturing something over and above single task performance. The drawback of the test is that it is not clear how scores on WOMBAT are related to SA. On the other hand, the test offers a useful method for testing the role of the underlying mechanism of WM in supporting SA. For example, the bonus tasks could be systematically manipulated to test the relationship between SA and the different components of WM. However, in order to systematically test the relationship between WM and SA a more complete theory specifying this relationship is needed.

2.7 Problems with the alternative view

Adams et al. (1995) specify the functions of WM and relate it to the maintenance of SA in a multitask environment. Their approach therefore is more prominent in forming a complete theoretical model that explains SA and allows researchers to generate testable hypotheses. According to Adams et al. WM plays a central role in supporting SA as it handles online processing of information. However, the framework by Adams et al. is still very general. The relationship between SA and WM is not specified fully enough to support empirical research. Although they clearly identify aspects of known WM models in basic research, there is no connection to empirical evidence that might provide further

insight into the role of WM in maintaining SA. For example, in order to use the WOMBAT test to study the link between SA and WM, a stronger conceptual foundation is needed.

2.8 Empirical findings

There are empirical findings, both in basic and applied research, that directly or indirectly support the role of WM in maintaining SA. Understanding the existing evidence is important for further linking the two concepts. This section focuses on the relevant evidence in applied research. Further evidence from basic cognitive research is reviewed in chapter 4.

Majority of applied research on SA is guided by the product view, where the impact of specific factors (such as a particular display) on SA is examined. For example, Andre, Wickens, Moorman and Boschelli (1991) varied the perspective (3D or 2D), frame of reference (inside-out or outside-in attitude indicator) and visual momentum (the display was either in colour or monochrome). They tested how the different features of the display affected the pilots' SA. This study is quite characteristic for SA research in general. Although, it is important to measure how SA is affected by different system design, there are some obvious shortcomings to this type of research. First, SA is both what is being measured and what explains the problem. We know that SA was lost because the observer didn't perform in a certain way and at the same time we explain the decrement in performance by referring to the operator's SA being lost (Flach, 1995, p. 151). Second, this type of research is not systematic in that there is no necessary or sufficient connection between the different studies or between research and theory.

Research within the product view also looks at the correlation between measured SA and various cognitive factors. Research here provides some insight into the cognitive factors that contribute to the operator's SA. However, even though there is a correlation between SA and cognitive factors such as WM, it is still unclear how WM contributes to SA. Thus, research based on the product view provides knowledge on the correlation between a particular design and the measured SA but no explanation on how SA is acquired and maintained.

Another, more relevant approach for the present work, is to systematically manipulate the environment in a way that allows the researchers to make inferences regarding how knowledge is acquired and maintained in a complex, dynamic environment. Many experiments examine cognitive processing in complex operational environments without specifically measuring SA. However, these experiments may still provide information on the potential role of the underlying cognitive mechanisms in maintaining SA.

A recent study, using the fMRI technique to localize mental functions in a simulated flight environment strongly suggests that there is a complex relationship between WM and SA (Perez et al., 2000). In Perez et al. expert and novice pilots followed a defined flight path that was projected on the screen in front of them. The flight path was determined by a sequence of red rectangles that were placed one after the other at different rates to simulate variation in speed. In the easier condition the rectangles simulated a constant speed of 100 knots. In the more difficult condition the rectangles simulated a constant speed of 200 knots. Both expert and novice pilots showed dominant activation in the right frontal cortex while performing the task. The right frontal cortex

contains the premotor and the prefrontal areas that are both thought to be involved in WM processes (Goldman-Rakic, 1992). In particular the prefrontal area has been associated with spatial WM (Jonides, et al., 1993) and planning behaviour. Perez et al. found that there was more prefrontal activation for the experts compared to novices. Both frontal and prefrontal areas are associated with executive control function in WM, and are responsible for coordinating activity, selectively attending to information and switching between tasks (Baddeley, 1996; Baddeley, Chincotta, & Adlam, 2001). Furthermore, the expert pilots seemed to process information along a well-defined path from high-level processing areas in visual cortex (V3 and V5) directly to frontal and prefrontal areas. Whereas novice pilots showed less selective and defined processing. Their performance seemed to be based on widespread activity in both the motor region and the prefrontal lobe. Perez, et al. concluded that performance on the track-following task is based on maintaining an active representation of motion and direction of flight in WM, allowing the pilots to anticipate and plan future movements. This process is less developed or trained for the novice pilot than the expert pilot.

Perez et al's. (2000) results are particularly interesting for SA. First, these results suggest a complex role for WM in maintaining SA. Although SA was not measured, the track-following task clearly involves SA of the projected path: the participant must maintain a complex representation of the direction and movement of the rectangles that allow him or her to anticipate future moves. This particular task also relied mainly on representing and comprehending the incoming information (the moving path) rather than involving knowledge stored in LTS. The results therefore suggest a complex role of WM in maintaining SA, even for the expert operator. In fact, the experts showed more

activation in areas associated selectively with WM where as the novice pilots showed more activation in areas associated with mental load. In the SA literature, WM is often seen as short-term storage for information that puts more constrictions on the novice operator who cannot activate sophisticated mental models in LTS (Endsley, 1995a). In contrast, the expert operator can bypass the limitations of WM by relying on mental models activated from LTS. However, Perez et al.'s. results indicate that expert performance may largely be based on well-developed WM functions. Similar results have been found in basic research on WM and chess (Robbins et al., 1996). The results by Perez et al. also suggest that examining knowledge may not allow researchers to make the needed inference regarding the underlying representation in WM. Asking the pilots about their knowledge of the path at different points in time is unlikely to reveal the complex WM function shown by the fMRI.

The potential role of spatial WM in the track-following task is interesting. It is possible that verbal and spatial WM systems play a role in maintaining SA. Caretta, Perry, and Ree (1996) found that spatial WM along with verbal WM, spatial reasoning, and divided attention predicted SA performance measured as the ability to form and follow tactical game plans, to communicate, to operate the system and to interpret information (e.g. threat prioritization). In studies by Gugerty and Tirre (1995; 2000), maintaining awareness of location and avoiding hazards was highly correlated with WM measures. Aretz (1991) also found evidence suggesting that both the task of controlling an aircraft and navigating utilizes spatial resources, possibly in spatial WM. Furthermore, when the navigation task became too demanding, pilots seemed to switch from relying

primarily on spatial resources to relying on both spatial and verbal resources for carrying out the flight task.

Perez et al's. (2000) results also suggest that maintaining SA in a dynamic environment is highly dependant on executive control because the dominant activity was measured in the frontal lobe, area associated with the central executive system in WM (Baddeley, 1996). The central executive system is, for example, responsible for task switching, attentional control, and time sharing (Baddeley, 1996; Baddeley, et al. 2001). Tham and Kramer (1994) found that flight instructors were better than student pilots in task switching and in ignoring distracting information, whereas both groups were similar in their ability to scan the environment. Similarly, Gopher (1982) found that performance on a dichotic listening task distinguished between successful and unsuccessful flight cadets. In particular switching errors were highly correlated with successful performance in flight training.

Hopkin (1993) argued that SA is largely determined by higher-order attentional management ability and that this ability is somewhat independent from the individual abilities. This claim was supported by O'Hare (1997) who found that scores on WOMBAT, where participants are engaged in multiple tasks, did not correlate with scores on individual tasks. Performance on the WOMBAT test was also highly predictive of elite pilot performance. Similarly, Damos (1993) found that studies that used multiple tasks had significantly higher predictive validity for pilot performance than studies that used single task measures. Furthermore, Endsley and Bolstad (1994) used an extensive selection of tests to assess individual differences in SA. The tests were divided into five categories; spatial abilities, attention, memory, perception, and cognitive abilities.

Interestingly, the measure of tracking difficulty in a dual task context correlated highly with SA scores. Participants performed a two-dimensional tracking task along with a digit cancellation task and tracking difficulty automatically increased or decreased depending on each participant's performance. No other measure had a significant correlation with SA scores. Similar results were found by Bolstad (1991). But in addition to dual task tracking, Bolstad found that delayed memory and encoding speed measures correlated with SA scores.

In summary, the existing evidence in applied research suggests that WM mechanism may play a critical role in acquiring and maintaining SA in a dynamic environment. In particular, the evidence suggests that spatial and verbal WM systems support SA in addition to some sort of executive control allowing operator to deal with multiple tasks.

2.9 WM: A cognitive model for SA

To date, theoretical and empirical developments relating WM to SA have been weak. Partly the problem is conceptual, that is, it has been unclear whether SA should be understood in terms of what the operator knows or how the operator gets there. On Endsley's view for example, WM is seen as a temporary store for "the product" SA. However, given the evidence reviewed above, there seems to be a good reason to explore the potentially more complex relationship between SA and WM.

Various models have been proposed to explain the WM mechanism (see Miyake & Shah, 1999). In the SA literature, researchers have focused on the LT-WM model by Ericsson and Kintsh (1995) as the best suited WM model to explain the cognitive

processes underlying SA (Durso & Gronlund, 1999). However, there is no evidence linking LT-WM to SA. A different model is chosen in this dissertation to build up a conceptual link between SA and WM. This model is the multi-component model proposed originally by Baddeley and Hitch (1974) and revised by Baddeley (1986, 2000) and Logie (1995). The multi-component model proposes that the WM mechanism is comprised of a verbal and spatial system for maintaining information and central executive control for coordinating tasks and attentional control. This model therefore fits well with the existing evidence reviewed in the section above. The multi-component model is also well established for generating hypothesis and has been studied extensively in the context of complex cognitive tasks.

The goal of the present dissertation is to use the multi-component model of WM to build up a theoretical foundation for exploring the role of WM in maintaining SA. This will provide the necessary basis for testing if and how WM and SA are related. An empirically tested relationship will provide foundation for further theoretical development in the field. If WM does in fact play a role in supporting SA, then understanding the relationship will facilitate our understanding of SA, providing researchers with tools to predict when and how SA can be affected.

2.10 Summary

An important distinction is made between SA as a *product* or what the operator knows and can report, versus *how* SA is generated and supported. This distinction has not been discussed explicitly in the SA literature. In fact, researchers have been generally silent on the existing difference in theoretical approaches or what they might mean for

further theoretical and empirical development. As summarized above, understanding the relationship between SA and underlying cognitive mechanisms is important if researchers are to generate a predictive model of SA. The goal of this dissertation therefore is to examine the role of WM in supporting SA.

There are however, some foundational issues in SA research that cannot be overlooked for further theoretical and empirical work to take place. In particular, the assumption that SA is what the operator is conscious of knowing is troublesome for the field. On this assumption, the criteria for what is or is not an example of SA, is determined by what the operator can report. This however, seems at minimum a limited approach for studying complex cognition in a dynamic environment. A particular concern is that there has been no discussion of why the term “awareness” in SA should necessarily mean conscious report. Arguably, people can be aware of a situation in that they perceive, comprehend and anticipate events, in the absence of the ability to report on their awareness. In fact, it is conceivable that an operator could have good SA, but is unaware of what he or she knows. Accordingly, it is important both for theory and methodology to discuss what the term “awareness” means in the context of SA. This is addressed in the following chapter.

Chapter 3: The meaning of “awareness of the situation”

3.1 What is SA awareness of?

A troublesome foundational issue that has not been dealt with in SA research is the meaning of awareness in situation awareness. The concept has been measured as the ability to report locations of surrounding aircraft (Endsley & Bolstad, 1994) and as the ability to maintain a defined flight path (Williams, 2002). But are both examples of SA? On the product view, SA is assumed to be what the operator is conscious of knowing and can report. This can clearly be seen in how SA is measured by having operators report information from memory. But if SA is what the operator is conscious of knowing, then awareness in situation awareness really means being aware of what one knows about the situation, rather than just being aware of the situation. Therefore, maintaining the flight path does not necessarily reflect SA; a pilot may rely on knowledge that is not easily accessible for verbal report. But there is something deeply troubling about this assumption. Ace pilots, for example, seemed to have better awareness of the world compared to the average pilot but they are not necessarily more aware of what they know about the world (the task situation). The aces have been found to be quick and to have better strategies than average pilots. Aces also know when to quit and when to pursue an engagement (Spick, 1988). There are, however, no reports suggesting that their reported knowledge about the situation was better than that of the average pilot. This chapter will review some evidence from cognitive psychology and neuroscience that contradict the view that SA is determined by what the operator consciously knows and can report.

3.2 Explicit vs. implicit knowledge

The association frequently made between higher order cognition, consciousness and verbal report can be traced to a distinction made in the cognitive literature between explicit and declarative knowledge on the one hand and implicit and procedural knowledge on the other. Explicit or declarative knowledge refers to facts, memories of events and episodes. It is knowledge that people can access and frequently report. This knowledge is traditionally associated with complex cognitive activity such as problem solving, reasoning and decision-making. Furthermore, explicit knowledge is often thought of and studied in isolation from other processes and mechanisms. For example, people are often aware of how they solve problems. That is they are aware of their mental states or at least that is the common assumption. Therefore, having them explain how they are solving the problem or talk aloud while performing the task is in many cases considered a reliable source of data.

Contrary to explicit and declarative knowledge, procedural knowledge refers to learned skills, and automatic performance. Implicit knowledge has been defined with reference to performance facilitation due to prior experience or learning on tasks that do not require that people are conscious of those experiences (Schacter, 1987). Both procedural and implicit representations have in the past been thought of as categorically different than explicit and declarative knowledge. They have also been thought of as some form of representation of the world that people are not aware of having, at least in some sense of the concept. Furthermore, implicit knowledge is thought to be simpler than

explicit knowledge and not to play any direct role in complex cognition. It is therefore not surprising that SA is associated with explicit knowledge that the operator can report.

Recent evidence in cognitive psychology however, poses a problem for many assumptions about explicit and implicit knowledge. In particular, it seems that people can learn and represent complex information and use it to guide performance without explicit awareness of it. Chun and Jiang (2003) used a paradigm referred to as contextual cuing to study implicit representation of information. Participants were presented with various displays where a rotated T (the target) was presented among Ls (distractors). Several displays would repeat the same layout, where targets appeared in the same relative location in the overall spatial configuration. This gave the participants the opportunity to learn the overall spatial layout and use it as a cue in their visual search for the target. As expected when participants were presented with the same spatial layout they responded more quickly compared to when they were presented with a new layout (Chun & Jiang, 2003). Furthermore, Chun and Jiang demonstrated that implicit representations were coded in LTS because participants showed the same enhancement in performance when searching for targets in the learned configuration after one week. However, these implicit representations did not affect participant's explicit knowledge. Following the visual search learning, Chun and Jiang had participants predict where the target would appear, showing them either the same displays as they had used before or new displays. Participants performed at chance in predicting the target location for both new and old displays. This suggests that although implicitly they represented and used the information on the overall spatial layout, explicitly they had no awareness of having this information.

Even when participants were instructed to pay attention to the repeated layout, during the learning phase, their performance in predicting the target location was not enhanced.

What is interesting about the study by Chun and Jiang (2003) is that the contextual cuing paradigm requires fairly complex coding and representation of information. The evidence therefore is against the claim that implicit knowledge is categorically different from explicit knowledge. The main difference may be that participants did not have access to the represented information and therefore could not articulate it. It is possible that with extensive training, participants would be able to explicitly identify the different configurations. The key here however, is that explicit awareness did not seem to be necessary for the complex representation of information that guided performance in the study by Chun and Jiang. The evidence suggests that whether or not an operator is aware of what he knows may not make any difference in the underlying knowledge and how it affects performance. Rather, sometimes people have access to what they know and sometimes not. For example, there are cases where procedural knowledge (how to) is accessible (part of what a person is aware of knowing). This is often the case with novices as will be discussed in the next section. Similarly, it is possible that complex declarative knowledge cannot be articulated. A driver in a heavy traffic may be making complex inferences; choosing strategies and evaluating alternatives (e.g. alternative routes). But the driver may or may not be aware of having that knowledge or be able to articulate it.

As pointed out by Banburry, Andre, and Croft (in press) maintaining SA in a dynamic environment may involve both implicit and explicit knowledge. It is in fact

unlikely that all information in a multitasked environment is explicitly coded and maintained. Therefore, focusing exclusively on measuring explicit knowledge may not provide a complete understanding of the SA phenomenon. In particular, it is erroneous to assume that an operator is always able to report on what he or she knows about the situation. In the study by Chun and Jiang, participants “knew” the target location once they had learned it and they used this knowledge to guide their performance. But participants were not aware of what they knew. Furthermore, as demonstrated by Nisbett and Wilson (1977), people are often unaware of how stimuli affect their behaviour or even that the stimuli are there. Thus it seems unlikely that operators will be aware of everything they “know” when performing more complex tasks. .

3.3 SA and Expertise

An area where the explicit-implicit distinction becomes a particular concern for the product view of SA is the area of expertise. Endsley (1995a) argued that expert operators use fewer resources in gaining SA as they quickly recognize patterns in the situation by activating stored mental models in LTS. Therefore, the process of gaining knowledge is automatic and not consciously available to the operators, but the product or what they know is available to the operator’s conscious mind. Hence, SA of the expert operator can always be measured by having the operator report the status of the system and the task situation. But is this really the case? For example, novice drivers start by utilizing some abstract rules such as shifting into second gear when the speedometer points to 10 mph (Dreyfus, 1997). They are also likely to remind themselves to look

around, to think about how close they can be to other cars and still be safe and so on. The novice drivers therefore are able to report their knowledge about the driving situation. They are aware of what they know. In contrast, expert drivers know more about the driving task than the novices and are more likely to have superior SA of the driving situation. However, expert drivers are also less likely than novice drivers to access this knowledge and to articulate it. This is the paradox of expertise (Johnson, 1983, cited in Banbury, Andre & Croft, in press). As people gain more knowledge in certain areas, they also start to lose awareness of what they know.

The real paradox here though is that clearly in most cases expert operators have superior SA than novice operators. The previous analysis, however, suggests that experts are often less able to articulate what they know about the task situation compared to novices. Therefore, on the default assumption in the literature, the experts really have poorer SA than the novices. Obviously there is something wrong here. The problem with the product view is the clear-cut distinction between the conscious product and everything else. With experience much of what the operator knows about the task gets transferred into skilled and automatic performance. It becomes harder to distinguish the product (what the operator knows) from the processing of information. Stoke et al (1997) for example, found that expert pilots demonstrated superior performance in a cockpit simulation compared to novice pilots. The experts seemed to have superior association between stored knowledge and cues in the environment that made them better than the novices at ignoring irrelevant cues and attending to relevant ones. However, the expert pilots did not have more explicit knowledge of the task situation than the novice pilots.

As demonstrated in the study by Stoke et al. experts' SA may often have little to do with their explicit knowledge of the situation. In fact, as expertise increases, SA may be more comprised of implicit than explicit knowledge. Hence, the accepted methodology in the field of asking the operators to report what they know may be largely inadequate.

3.4 Blindsight, inattentional blindness and other evidence from neuroscience

There is a different angle to the explicit/implicit distinction. Namely, when can a person be said to be aware of the environment or the task stimuli? What is acceptable criteria for awareness? This issue is recently being raised in neuroscience. One criteria everyone seems comfortable with is verbal report. But there is no one to one relationship between verbal report and awareness. Therefore, most researchers would rather rely on performance measures as an indication of whether a participant is aware of task stimuli. Yet, when awareness breaks down, as in the case of blindsight, the absence of verbal report has most researchers question the presence of awareness. This is true even though blindsight patients demonstrate complex cognitive processing in the absence of reported awareness.

In blindsight, damage to the visual cortex leaves people with a 'blind spot' in a part of their visual field. Research has shown that for blindsight patients, some visual abilities are preserved for stimuli presented in the blind spot. They are, for example, able to detect movement of stimuli and to discriminate orientation and shapes although they report not seeing (i.e., not being aware of) the stimuli (see Weisenkrantz, 1986, 1990 for

reviews). In fact, blindsight patients are still remarkably accurate in reaching for and making judgments about stimuli presented in the blind spot. When pushed, some blindsight patients can report that they feel the onset of a stimuli or that they sense the difference in two different shapes presented in the blind spot (Weisenkrantz, 1986).

Kentridge and Heywood (2000) discuss a previous study by Kentridge, Heywood, and Weiskrantz (1997) where a blindsight patient reported trying to direct his attention to a spatial location within his blind field. Kentridge and Heywood did further testing on this patient and found that the patient was in fact able to direct his attention to a target following a cue that was presented either outside the blind field or within the blind field. The cue was an arrow that either pointed toward or away from the target. Research has shown that for normal participants, responses are faster when a cue is valid compared to when a cue is neutral or invalid. Also, invalid cues generally cause slower responses compared to neutral cues. The same method was used for estimating whether the blindsight patient was in fact aware of the cues presented in the blind field. The cue was accompanied by an auditory tone and the patient's task was to guess as quickly as possible whether presentation of a target followed the auditory tone. On half of the trials there was no target present (neutral). On the remaining trials, the cue was valid (pointed to the target) on 68.75% of the time and invalid (did not point to the target) on 31.25% of the time. After guessing, the patient was also asked to indicate if he had awareness of the target. Kentridge and Heywood found that the blindsight patient was able to use the cue to respond to the target. Similar to normal participants, he was significantly faster at responding to the target with valid than invalid cues, suggesting that the orientation of the

cue was processed. For both types of cues, he was equally accurate at detecting the presence of a target. However, the patient never reported *awareness* of the target. Kentridge and Heywood were concerned that the results reflected simply the effect of energy summation between the target and the cue and not selective attention in the absence of reported awareness. To address this possibility, they tested performance when the target appeared most of the time in reverse relationship with the cue. The experimenters informed the patient that after the first three blocks of trials, the target would be more likely to appear on the opposite side of the location that was indicated by the cue. For the first three blocks of trials, the patient was faster to respond to the target when it occurred in the cued location than when it occurred in the opposite location. However, this changed after the first three blocks and the patient became faster to respond to the target when it appeared in the opposite location of what the cue indicated than when it did not.

The Kentridge and Heywood (2000) results suggest that blindsight patients can not only discriminate targets presented in their blind field, but can also switch schemas and attentional focus without being aware of having switched focus or schema. The conclusion that blindsight patients have *awareness* of stimuli presented in their blind spot is not universally accepted. It is hard to argue, however, why this would not be awareness unless awareness necessarily is accompanied by verbal report. There are however, cases where researchers readily refer to awareness in the absence of verbal report. For example, commissurotomized patients show deep processing of stimuli presented to their left visual field but report having no awareness of the stimuli (see Rugg, 1992). It is

commonly accepted that commissurotomy patients are aware of the stimuli but that the disconnections between the two hemispheres means that what is content of awareness in the right hemisphere is not accessible to the left lateralized speech system. It is however, not clear why this is an accepted case of awareness and not for example, the blindsight phenomenon (Rugg, 1992).

Inattentional blindness makes it even harder to deny awareness in the absence of verbal report than the blindsight phenomenon. In inattentional blindness studies (e.g., Mack & Rock, 1998), participants are presented with fixation followed by a target stimulus and asked to make a judgement (e.g., which of two lines in a cross is longer). The target stimulus is presented for 200 ms followed by a mask. In some trials the target stimulus is presented unexpectedly along with another stimulus (called the critical stimulus) followed by the mask.

Most participants (up to 75%) fail to see the critical stimulus even when it is presented at fixation with the target presented in the fovea region. When the critical stimulus is something highly meaningful to the participants such as their name, they never fail to report the stimulus. However, if the participants' name is replaced by another name with only a single letter difference, participants again show high inattentional blindness. This suggests semantic processing of the critical stimuli in the absence of explicit awareness. Furthermore, when participants are asked following the inattentional blindness study to complete word stems, they are much more likely to complete the stems according to words used as critical stimuli in the inattentional blindness trials compared to words that were not used.

Inattentional blindness demonstrates a deep processing of stimuli in the absence of reported awareness. Furthermore, the processing of stimuli is such that if it is relevant to the participants like their name they do become explicitly aware of it. Using techniques such as fMRI and ERPs researchers have found that in fact, neural activity is the same when awareness of the stimuli is reported and when no awareness of the stimuli is reported (for a review see, Dehaene & Naccache, 2001; Kanwisher, 2001; Merikle, Smilek & Eastwood, 2001). Luck, Vogel and Shapiro (1996) for example, found that accessing the meaning of a word shows the same neural activity regardless of whether the person is aware of knowing the meaning or not. Similar results were found by Whalen et al. (1998) for emotional expressions of faces that were masked by other stimuli.

Is it possible that perhaps awareness is not a unitary phenomenon and that a person can be aware of the world without necessarily having access to that awareness and being able to report it? Both Allport (1988) and Young (1994) argued that it is not that straightforward to decide when someone is aware and when not. A good example of this is described by Allport where a patient with a right hemisphere lesion was asked to put a midpoint on a horizontal line. These patients tend not to see the left side of an object. As expected, the patient put the midpoint about three quarters to the right. However, in a different task, the same patient was asked to treat a horizontal line as the base of an imaginary square and then draw the right-hand side of it (see Allport, 1988, p. 164). This time the patient drew a vertical line of the right length. Was the patient unaware of the line in the first task and aware of it in the second task? Or was the patient unaware of the line in both tasks?

3.4.1 Two separate pathways

One way in which the data from neuroscience can be explained is by claiming that there are two processing pathways, one providing access to the world and the other providing subjective experience and report of the world. There is in fact, evidence suggesting such a division in processing information (Allport, 1987; Goodale & Milner, 1992; Milner & Goodale, 1998). The visuo-motor system has been associated with the dorsal pathway from the primary visual cortex (V1) to the posterior parietal cortex. Information here is not accessible to the person and cannot be used for verbal statements of the world. This system is used to control action in the world. Another processing pathway has been associated with the ventral pathway from V1 to the inferotemporal cortex. Information here is conscious and can at least partly be accessed by the person and used in statements about the perceived world (Dienes & Perner, 1999).

If there are two systems it means that information can be represented and used by the cognitive system without the person being subjectively aware of it. That is, the cognitive system may represent properties of objects and this information can be used for action control without there being any factual representation of this information that the person is aware of having. For example, Bridgeman (1999) showed that for a human observer a stationary dot within a moving rectangular frame appears to move opposite to the movement of the frame. After the apparent movement took place the display disappeared and participants were asked to indicate either verbally or by pointing, at which of five marked locations the dot had been. Interestingly, only when the participants

reported verbally where the dot had moved, were they susceptible to the illusion. When participants pointed, they were able to point to the accurate position of the dot, that is, showing that it did not move.

It is interesting to speculate what the results from the study by Bridgemen (1999) and the two separate processing systems mean for SA. It is likely that in a dynamic environment like driving, performance is guided by both dorsal and ventral channels. This means that the driver's subjective experience does not capture (all of) how the driver interacts with the task environment. There is in fact evidence to support this claim. Kirsh and Maglio (1994) had participants play the highly interactive video game of Tetris. What they found was that the participants would start to manipulate the zoids (the shapes) as they appeared on the screen long before any subjective experience (and decision-making) of where and how to fit the zoids took place. Rather, the participants seemed to be interacting with the environment first, and their subjective experience of what happened came later. When asked, the participants described various strategies, but their descriptions never reflected an awareness of how early manipulation of the blocks influenced performance on the shape-fitting task.

The explicit channel (the ventral pathway) may in fact require more cognitive resources than the implicit channel (the dorsal pathway). Therefore, in a dynamic task environment, performance may be largely guided by the implicit channel. This would be a problem for the product view on SA. In fact, the explicit channel may encode events after performance takes place rather than encode information in order to guide performance. Where the explicit channel might become more involved for the

maintenance of SA is in situations such as where the driver gets lost and needs to figure out the instructions given earlier. But on this view a complete assessment of SA would have to include information the operator cannot report.

3.5 Some philosophical interpretations

As pointed out by Brook and Raymont (in press), there is much inconsistency in how terms like consciousness and awareness are being used. For example, compare the sentence, *the dog suddenly became aware of the cat* with *John is aware that he is less stressed today than yesterday* (examples taken from similar examples in Brook and Raymont, in press). Clearly awareness in the first sentence is not the same type of awareness as referred to in the second sentence. Or more importantly, what would account as criteria for awareness in the second sentence might not work for defining awareness in the first sentence. Consider the third example; *I am vaguely aware of something behind me*. Is being vaguely aware of something in the environment the same as being aware of one's own psychological states?

In order to deal with the apparent inconsistency in using the terms awareness and consciousness, some philosophers are proposing a distinction between being aware and having access to that awareness (e.g. Brook & Raymont, in press; Dretske, 1995). For example, Block (1997) distinguished between phenomenal consciousness and what he refers to as access consciousness. On this distinction, one can be aware in the sense of having representations of the world that are "...poised for free use in reasoning and for direct 'rational' control of action and speech" (p. 382). This is access consciousness,

which can exist without phenomenal consciousness or the feeling of it being like something to be in some state. A slightly different distinction is made by Brook and Raymont between awareness of the world and awareness of oneself and one's mental states. Awareness of the world refers to "central, cognitively-controlled, goal-directed information processing" (p. 13). It is a conscious access to the world. On this view, people represent information about the world even at semantic level and use it to guide action in the world. As such, people have conscious access to the world. Brook contrasts this to information access; the iris closing down due sudden increase in light hitting the eye or an organism leaning towards a sudden wind gust. The main difference between the distinction made by Block on one hand and Brook and Raymont on the other is that for Block, people can represent the world and use the representation to control action without it being like anything to represent the world. For Brook and Raymont however, awareness of the world is an example of a full-blown consciousness – it is like something to represent the world. But it does not necessarily follow that the person has access to this awareness as in being able to describe it or report it.

Brook and Raymont argue that there is a need for the concept of awareness of the world that is different from being aware of what one knows about the world. Making this distinction allows researchers to explain the dissociation between what is reported and what is represented and used to control performance. People are clearly able under certain circumstances to represent stimuli even at a semantic level and using that information to control performance without subjective experience of representing the stimuli. The distinction between awareness of the world and being aware of what one knows about the

world also fits better with cognitive theories. WM may hold the coordinated representation of the world that supports awareness but awareness here does not necessarily mean that the underlying representation can be accessed by the person and reported.

It seems reasonable to accept that an operator in a dynamic environment can maintain coordinated representation in WM that supports awareness of the situation but that it does not necessarily follow that the operator can report it. A good example of this comes from the study by Kirsch and Maglio (1994). The participants playing the Tetris game clearly had awareness of the world. They were aware of the game and the game environment, the goal of the game and so on. But they had no insight into this awareness. In a dynamic environment where people must react quickly to things changing in time and space, being conscious of what one knows might involve a too slow process to be practical. The distinction between awareness of the world and awareness of ones own mental states is important for SA. Clearly it is the latter sense of awareness that researchers have been focused on. But the former, awareness of the world, seems to capture better the meaning of SA in the case of expertise. The expert driver for example seems to maintain SA by a superior access to the world rather than or at least in addition to awareness of what he or she knows about the driving scene.

Dretske (1995) makes a similar distinction to Brook and Raymont between being aware and having access to that awareness. He argues that researchers are too prone to view mental states as either being conscious or unconscious. However, making this type of distinction is meaningless. Rather, mental states are in nature the same but sometimes

we are aware of having them and sometimes not. But whether we are aware or not of our mental states does not make them more or less relevant to explain behaviour. There may be cases where being conscious of what one knows is relevant, for example, when performing a new task. This brings up the issue again of experts and novices. The novice operators for example, may be more aware of their knowledge and how it is related to the task than expert operators. But being aware of one's mental state does not alter the relevance of the mental state as such. This is important for SA. Being aware of the situation can be supported by mental states that integrate information and facilitate projection regardless of whether the operator is aware of what he knows about the situation. Dretske argues that it is a common misconception to think that if an experience or mental state is conscious then we are necessarily conscious of it. In fact, it is a dangerous assumption to make. It leads further to the assumption that one cannot see a tree without being aware, not just of the tree, but our experience of it (Dretske, 1995). Therefore, rather than thinking of mental states in terms of either being conscious or unconscious, a better approach is to accept that representing the world is fundamentally the same, sometimes we are aware of what we represent, sometimes not.

3.6 Back to SA

The evidence reviewed here does not determine the meaning of awareness. Much more conceptual and empirical work is needed to define what it means to be aware of the environment. Clearly, however, the product view of SA does not work. Each of the reviewed studies raises a problem for the assumption that SA is necessarily explicit,

conscious knowledge that can be reported. Researchers in the field have largely been silent on the issue. Durso and Gronlund (1999) briefly mention the lack of discussion on what awareness might mean. Sarter and Wood (1991) argue that awareness in situation awareness must include knowledge currently inactive but available for recall. Again, the ability to report seems to be the default assumption in SA research. But much of cognition takes place in the absence of reported awareness of it. Expert operators for example, know more about the situation but they are not more aware of what they know compared to novice operators. In a dynamic environment knowledge may be more supported by the implicit rather than the explicit channel. In general, the view of SA as conscious comprehension that is separate from information gathering and performance does not seem to fit well with the existing data.

One way to deal with this problem in the field is to go with the alternative approach at least as a general framework. On this view as proposed by Adams et al. SA is both the product and the processes. Furthermore, there is no reference made to conscious report. In fact, Adams et al. argue that online assessment of performance is the most direct method to assess SA. The emphasis here is on understanding how the operator perceives, comprehends and anticipates events by specifying the operations of the underlying mechanism of WM. It is therefore proposed for now that if the operator perceives, comprehends and anticipates events in a dynamic environment –the operator is aware of the environment. Hence, how perception, comprehension and anticipation are operationalized determines what is or is not a case of SA. But whether or not the operators are able to verbally report what they know is not necessary or sufficient criteria

for SA. For example, expert operators may comprehend the task situation and anticipate future events but not be able to clearly articulate their knowledge.

The important point here is that reportable knowledge can be relevant but it does not determine SA nor does it make certain cognitive operations more relevant. For instance, a pilot who is landing a helicopter in a difficult terrain must keep track of multiple factors, like the closeness to the ground, rate of descent, development of attitude, pitch and roll, any wires or other objects in the way and so on. The pilot must keep track of those elements as they change over time, integrate them into a coherent picture of the situation and anticipate future states. Also, the pilot must evaluate different strategies and their outcome in the context of the situation. The pilot has SA but is not necessarily aware of what he or she knows about the task situation. Contrast the helicopter pilot to a commercial pilot who is faced with some anomaly in the system and needs to interpret what it means and make a decision accordingly. That pilot needs to access prior knowledge about *what to do when x happens*. Therefore, the pilot in this case very likely not only represents the situation but is aware of what is being represented.

In the case of the commercial pilot being aware of what the operator knows is relevant to making an accurate decision. In the case of the helicopter pilot however, there is no need to be aware of what he or she knows about the situation. But the difference between the helicopter pilot and the commercial pilot is not that one pilot consciously comprehends the situation and the other one does not (hence one pilot has SA the other one doesn't). Rather, one pilot is aware of what he or she knows about the environment, and the other one is aware of the situation but may not have much insight into that

awareness. There is however, no fundamental difference in their cognitive states or how it contributes to performance and understanding. Both pilots rely on an integrated representation that allows them to comprehend the situation and to anticipate future events. Therefore, both pilots have SA but only in the latter case does it make sense to have the operator report what he or she knows.

The example of the two pilots and the evidence reviewed in this chapter are difficult to account for on the product view. On a product view, the commercial pilot would necessarily be a better example of SA. Being aware of what one knows is assumed to determine whether the operator actually comprehends the situation. Clearly, however, maintaining awareness in a dynamic situation may or may not be accompanied by verbal report. Therefore, a natural step forward in the field is to put the emphasis on not what the operator is aware of knowing but rather how the operator perceives, comprehends and anticipates in a dynamic, complex-task environment. This is an important shift both methodologically and theoretically. Methodologically, a wider range of measures can be used to assess the operator's SA. Furthermore, the emphasis is not on what the operator knows but rather on how the operator perceives, comprehends and anticipates. Understanding SA therefore must involve studying the underlying mechanism of WM. Theoretically, the relationship between SA and WM is not limited to explicit, conscious knowledge. This is important as in fact, only a subset of represented information in WM is thought to be what the person is aware of and is able to articulate. The following chapter will use a well-established WM model to build up a theoretical foundation for exploring the role of WM in maintaining SA. This model will furthermore generate

hypotheses on how the operator comprehends and anticipates events in a dynamic environment that then can be tested empirically.

Chapter 4. Working memory: outlining the role of the underlying mechanism

4.1 Working Memory and SA

WM is a theoretical construct that refers to the on-line processing and maintenance of information (Baddeley, 1986). This cognitive mechanism is what allows people to keep track of and integrate information both temporally and spatially, to control perceptuomotor coordination, plan movements and represent spatial layout (Baddeley, 1986, 2001; Garden, Cornoldi and Logie, 2002; Logie, 1995; Logie, Mane, Donchin, & Sheptak, 1989; Robbins et al., 1996). WM has been shown to underlie the ability to solve problems (Carpenter, Just & Shell, 1990), perform mental arithmetic (Baddeley & Logie, 1999; Logie, Gilhooly, & Whynn, 1994), play chess (Robbins et al., 1996) and to engage in spatial reasoning (Shah & Miyake, 1996) to name but a few tasks. WM has also been linked to the ability to anticipate future events (Posada, Franck, Georgieff & Jeannerod, 2001).

It is expected that the WM mechanism plays a role in maintaining SA. As reviewed in chapter 2 there is evidence in applied research supporting this claim. The WM mechanism also has been identified as the “place” where SA is maintained. There is however, no theoretical model that links the two concepts in a way that would guide empirical research. The following sections therefore will review the architecture of WM and the existing evidence in order to link the concepts. The potential relationship between WM and SA will then be tested empirically.

4.2 Structure and processing: The multi-component model

The proposed models on WM may vary from one model to another in terms of the structure of the system and the nature of processing of information. It is generally agreed, however, that WM is not a unitary system, rather it involves different representational codes or subsystems and some sort of a central control (see Miyake & Shah, 1999; Kintsch et al., 1999). The multi-component model of WM (Baddeley & Hitch, 1974; Baddley, 1986) consists of three main parts; verbal and visuospatial subsystems, an episodic buffer (see Baddeley, 2000) and a central executive control system (see Figure 3). This model has extensive empirical support. The evidence supporting the multi-component model ranges from selective (verbal versus visual/spatial) interference effects found in normal adults using dual-task paradigms, through to patterns of selective impairment in brain-damaged patients and differences in development observed in children (Baddeley & Logie, 1999). It also fits well with the data reviewed in chapter two that indicates that dealing with complex multitask environments may involve visuospatial and verbal WM resources in addition to central executive control (see for example, Aretz, 1991; Carreta et al., 1996; Gugerty & Tirre, 1995; 2000; Perez et al., 2000).

Figure 3. The Multi-Component Model of WM (Baddeley, 2002)

The verbal and visuospatial WM subsystems are responsible for maintaining task-relevant information such as intermediate solutions and subgoals in order to carry out cognitive tasks. There may be other storage systems within WM for representing information in other modalities but the visuospatial and the verbal systems are well established through research. Neuroimaging research has for example, found support for distinct spatial and verbal WM systems (Smith & Jonides, 1996).

The episodic buffer is a recent addition to the multi-component model and refers to a limited capacity buffer that represents coordinated information from the subsystems and from stored knowledge in LTS. The episodic buffer therefore is the connection between processing and representation in WM, and stored knowledge in LTS. Finally, the central executive is a special, possibly multi-dimensional, control system that is responsible for coordinating information from the various WM systems. The central executive also handles attention switching and controls both encoding and retrieval of information (Baddeley & Logie, 1999).

4.2.1 The phonological loop or the verbal subsystem

The verbal subsystem has been characterized as a phonological loop that consists of a phonological store for speech perception and an articulatory rehearsal system for speech production (Baddeley, 1986). The rehearsal system utilizes some sort of a subvocalization process to store verbal information within the phonological loop. This process is linked to a speech motor program rather than being a capacity to articulate per se (Baddeley, 2000a). There is also evidence suggesting that information within the loop is stored in a phonological code and is limited by a temporal duration (Baddeley, 1986). For example, research has shown that participants find it easier to recall a sequence of short words (e.g. wit, sun, harm, bag) compared to longer words (e.g. university, opportunity, constitutional, auditorium). One explanation is that with longer words it takes longer to rehearse each word thus allowing for more time for memory traces to decay. This phenomenon is called the “word-length effect” (Baddeley, 1986). Consistent with this idea is the fact that when participants are prevented from rehearsing an item, by continuously repeating an unrelated word, recall performance is significantly decreased. This task has been referred to as articulatory suppression and has been shown to impair participants’ ability to perform various verbal tasks where maintenance of verbal information is required (Baddeley, Lewis, & Vallar, 1984; Garden, Cornoldi, & Logie, 2002; Logie et al, 1989).

The phonological loop has been studied extensively, but usually in the context of relatively simple cognitive tasks. Hence, the role of the loop in complex cognitive tasks is not well understood. However, there is recent evidence suggesting that the

phonological loop may consist of a fairly complex system involved in many higher-order cognitive tasks (Baddeley et al., 2001; Miyake & Shah, 1999; Shah & Miyake, 1996). Therefore, many researchers in the WM field use the broader term *verbal subsystem*, rather than the phonological loop. Baddeley, Gathercole and Papagno (1998) argued that such a system may have evolved to facilitate language learning and that it plays an important role both in language acquisition for children and in second language learning for adults. This claim is further supported by studies in neuroscience. Patients with phonological deficits can learn word pairs in their own language but are unable to learn any words from an unknown language. Furthermore, children that have problems in sounding and repeating words are known to have general language difficulties suggesting some problems with the verbal subsystem (Baddeley, 2002). Baddeley et al. (2001) found evidence suggesting that the verbal subsystem may utilize some form of inner speech to monitor and control action. Furthermore, Shah and Miyake (1996) linked the verbal subsystem to the ability to make complex causal inferences during text comprehension.

The recent evidence concerning the role of the verbal subsystem in complex cognitive tasks makes it interesting to examine in the context of SA. There are many complex tasks for which SA is of concern and where verbal information needs to be processed. For example, a pilot may need to interpret an incoming radio message while simultaneously performing a number of other complex operations. Also making causal inferences or using prior knowledge may depend on the verbal subsystem for maintenance and temporary storage.

4.2.2 The visuospatial subsystem

The visuospatial subsystem is known to play an important role in tasks such as spatial reasoning (Shah & Miyake, 1996), chess playing (Robbins et al., 1996) and navigating (Garden et al., 2002). The visuospatial subsystem is believed to be less automatic and less practised than the verbal system, hence relying more on central executive resources (Baddeley, 2002). In an early study by Baddeley (1986), the visuospatial system was examined by having participants perform a spatial imagery task where they imagined a 4 x 4 matrix. Their task was to encode a set of sentences by travelling through the matrix. The first sentence might ask them to put a number in the starting square. The next sentence might ask them to put another number in the square below or next to the first one and so on. Participants' ability to perform on the imagery task was impaired by concurrent performance on another visuospatial task such as tracking but not (or less so) by another concurrent performance on a verbal task. It was also found that having participants point to a sound sources when blindfolded interfered with performance on the matrix task but judging brightness on a screen did not. These findings suggested that the visuospatial subsystem can be divided into separate spatial and visual systems.

It is now generally accepted that the visuospatial system can be further distinguished into a visual cache and an inner scribe (Logie, 1995). The visual cache is a temporary visual store that represents information such as shape and colour. The inner scribe is a spatial movement-based system that involves planning and executing physical movements as well as representing the path between objects or target sequences. It is hard to find tasks that reflect only on spatial or visual components. However, there is extensive evidence suggesting that tapping tasks where the

participants tap out a pattern according to a rhythm interferes selectively with the inner scribe. Conversely, presenting participants with tasks that involve judging or working with patterns and colours selectively interferes with the temporary visual store.

Logie and Marchetti (1991) found that when they presented participants with a sequence of squares appearing one after the other in a random location on a computer screen, the ability of the participants to recall the various locations was disrupted when they were asked to tap out a regular pattern during the retention interval (10 sec). No disruption was found when the retention interval was filled with presentation of random sequences of line drawings in the same location. However, the presentation of the line drawing disrupted recall of different hues of a single colour. Studies in neuroscience further support this distinction between a visual based and spatial based system. Patients suffering from brain damage show a selective deficit in visual or spatial WM (Logie, Engelkamp, Dehn, & Rudkin, 2001).

Research also suggests that movement-memory and movement-execution both rely on spatial coding in the inner scribe (Pickering, et al., 2001; Quinn, 1991; Salway & Logie, 1995; Smyth, & Scholey, 1994). Quinn and Ralston (1986) found that arm movements that are incompatible with spatial material can make it difficult to accurately represent spatial information. In this experiment participants imagined a path through a matrix by placing a number in the different squares of the matrix. While participants performed the primary task they concurrently moved their arm in a way that was either compatible, or incompatible, with their path through the matrix. The incompatible arm movement significantly impaired the participants' ability to represent the path through the matrix. Other studies have confirmed this phenomenon

by showing that tapping out a pattern on a keypad also interferes with the participants' ability to represent a path between objects in the environment (Logie et al, 2001).

It has been suggested that spatial coding of information is more important for many real-life tasks than visual coding (Logie et al, 2001). Encoding and maintenance of spatial information and the planning and execution of movement would be particularly important for the maintenance of SA, as spatial structures are a critical part of a dynamic environment.

4.2.3 Executive control

In many WM models, it is assumed that there is some sort of a control system that regulates and controls performance and information processing. The idea of a central control, however, is problematic as it raises the issue of homunculi doing the controlling. But it is generally not easy to get away from the idea of a system that coordinates and manages resources. Baddeley (1996) suggests that rather than thinking about the central executive as a unitary, central controller, focus should be placed on specifying the various sub functions of the system. This obviously does not avoid the problem of the homunculi, but allows us to test the various functions of the central executive, and as such, to gain some understanding of this system. It is only recently that more research has been devoted to the central executive system (Baddeley, 1996; Shah & Miyke, 1999).

Kane, Bleckley, Conway, and Engle (2001) found that people with high working-memory span are better than people with low working-memory span at using attention to resist interference from other items in the task environment while coding and or retrieving information. They argued that executive control serves an important

function in most cognitive tasks, allowing people to selectively attend to information despite distraction and to maintain goals. The central executive system might therefore play an important role in maintaining SA in complex environments by allowing the operator to prioritize and coordinate tasks. Research in neuroscience suggests that lack of central control function impairs people's ability to coordinate tasks and to distribute attention between stimuli (see Baddeley, 1996). Patients suffering from Alzheimer's disease can perform on individual tasks (spatial and verbal) at a similar level as young and older control participants. However, when they have to perform two tasks simultaneously their performance is significantly impaired compared to control groups. In general, damage to the frontal lobe seems to generate difficulties in coordinating tasks. However, frontal lobe and central executive functions are not the same thing. In fact, Baddeley (1996) warns that we should not think of the central executive in terms of a structural place in the brain.

The central executive is often measured by having people randomly generate numbers or letters or by randomly tapping on a keypad (Baddeley et al., 1998). The assumption is that random generation is a demanding task that requires the participants to maintain an internal task program allowing them to monitor what letters have been reported. It is assumed that the random generation task relies heavily on resources from the central executive. In fact, various studies have shown that participants' ability to solve problems, to reason, or to make complex decisions is impaired when they are required to concurrently generate letters randomly (see Baddeley, 1986; 1996). For example, Robbins et al. (1996), found that when intermediate and expert chess players randomly generated letters, their ability to recall and solve tactical chess positions was significantly impaired. Similarly, Baddeley et

al. (2001) found that the random generation task impaired the participants' ability to correctly alternate between adding and subtracting a series of numbers.

Hence, the central executive is assumed to play a role in coordinating and switching between tasks and in monitoring and evaluating performance/strategies. There is also some indication that the central executive works to activate LTS knowledge and integrating knowledge from LTS and knowledge represented in the subsystems (Baddeley, 1996).

4.2.4 The episodic buffer: Connection to long-term memory

The episodic buffer is a multimodal storage system (Baddeley, 2000; 2002). It binds information from the subsystems and LTS into a unitary multidimensional representation. This system is episodic in that it represents integrated scenes. Buffer refers to the limited capacity of the system that utilizes different codes and integrates them into a coherent representation (Baddeley, 2002). The episodic buffer is assumed to bind information necessary for establishing a coherent awareness of the task environment. This further "allows multiple sources of information to be considered simultaneously, creating a model of environment that may be manipulated to solve problems and plan future behaviour" (Baddeley, 2000, p. 93).

It is now generally agreed that WM has strong relationship with LTS and that a large part of processing information and forming a mental model of the world is based on activated information in LTS. For some time the only existing model that accounted for the relationship between WM and LTS was the LT-WM proposed by Ericsson and Kintsh (1995). However, the idea of the episodic buffer is in many ways similar to the LT-WM model. The main difference is that the LT-WM model assumes

that WM is a subset of LTS and carries pointers to stored knowledge and skill in LTS. This expands the limited capacity of WM and explains the difference in performance between a novice and an expert. In Baddeley's model however, the episodic buffer is not an extension of LTS, rather it is functionally separate system. Therefore, through the central executive, information or stored mental models can be activated and brought into the episodic buffer as a part of the current representation being built. But it is important here that the episodic buffer forms a new representation based on both incoming information and stored mental models in LTS rather than adding to an existing representation in LTS that has been activated. Note, however, that research on the verbal system suggests a direct link between the verbal system and LTS in addition to the link through the episodic buffer. For example, people tend to have a longer memory span for familiar words than unfamiliar words (Hulme, Maughan, & Brown, 1991; Hulme, Roodenrys, Brown, & Mercer, 1995). This suggests that for familiar words the verbal system may activate stored knowledge in LTS. It is also assumed that learned information on colours, shapes or familiar routes may be activated directly by cues in the visuospatial system

As Baddeley and Logie (1999) suggest, the relationship between WM and LTS knowledge is complicated. WM is more than a subsystem of LTS. Even for the expert operator, WM plays a vital role in temporary representation of information in a complex environment and in guiding the operator's actions. For example, Robbins et al. (1996) found that both medium-level and expert-level chess players rely on spatial WM for planning moves ahead of time. Similarly, Logie et al. (1989) had participants play the same videogame repeatedly until they gained some expertise. It was found that the novice participants tended to rely on general WM resources. However, as the

participants gained more expertise they started to rely more on different WM subsystems for different aspects of the task. In both studies, there was no indication of WM playing a lesser role in performing a complex cognitive activity as the participants gained more experience. Baddeley (1986) points out that as people gain more experience in a particular task context, more of their representation can be activated as a meaningful chunk from LTS. However, the maintenance of that representation along with incoming information from the present environment is still dependent on WM resources.

Although the episodic buffer has not been extensively studied, there is some evidence supporting such a system. For example, some densely amnesic patients with well-preserved intelligence have good immediate recall of prose paragraphs (Baddeley & Wilson, 2002). This cannot be explained by referring to activated information in LTS. In order to recall the prose the patients must store information into coherent chunks and episodes. Because these patients do not have access to LTS, there must be a system that coordinates information into coherent episodes. There are also amnesics who are able to run a store or play bridge (Baddeley, 2000, 2002). The only way to explain how they managed to carry out such complex tasks is by referring to a “place” where information can be integrated into a coherent whole.

The episodic buffer also provides some explanation on how information from the subsystems is brought together. There are various studies suggesting combined visual and verbal coding of information (Logie et al, 2000). Also a recent study on the connection between WM and awareness suggests a system like the episodic buffer. Baddeley and Andrade (2000) found that when participants were asked to form images from stored knowledge in LTS their judgements on whether the images were

vivid or not depended on combined WM and LTS factors. That is, they relied on information such as whether the image was active or passive and conventional or bizarre. Baddeley and Andrade concluded that the images are activated from LTS but maintained in the episodic buffer. The fact that people are able to generate and manipulate images they have never seen suggests that these images are maintained in the episodic buffer as a new representation.

Putting this into a context for SA, the episodic buffer might be the “place” where the underlying representation is formed and maintained to support SA. The underlying representation would depend on integrated information from the subsystems and from LTS. It provides the operator with a coherent understanding and awareness of the task situation.

4.3 The role of WM in complex cognition

WM has taken a central role in explaining complex cognitive activity. This is the mechanism that allows people to monitor information moment to moment and to maintain task-relevant information which is required for carrying out many cognitive tasks (Baddeley & Logie, 1999). Tasks that have been used to examine the role of WM include problem solving on standard intelligence tasks (Carpenter, Just, & Shell, 1990), mental arithmetic (Baddeley & Logie, 1999; Logie, Gilhooly, & Wynn, 1994), chess (Robbins et al., 1996) and text comprehension. Only recently has WM been tested in dynamic environments such as video games, tracking and driving. Although, basic cognitive research has not measured SA, research on text comprehension, chess and in particular dynamic environments is informative for generating hypothesis regarding the potential role of WM in maintaining SA.

4.3.1 Playing chess

Robbins et al. (1996) studied the role of WM in chess. They found that for both expert and intermediate players, the ability to remember chess positions was significantly affected by a concurrent visuospatial task but not by a concurrent verbal task. Robbins et al.'s findings suggest that remembering chess positions depends on representing patterns in the visuospatial system. The ability to remember chess position was also affected by a concurrent central executive task (random tapping). Robbins et al. suggested that an important component for playing chess is to access stored structures in LTS. The central executive therefore provides access to LTS knowledge about known chess patterns that facilitate memory of displayed positions. Robbins et al. also found that both the visuospatial subsystem and the central executive are important for analyzing chess positions and for projecting future positions.

What is particularly interesting about the results from Robbins et al. (1996) study is that there was little evidence suggesting that differences in chess skill depend on qualitative or quantitative differences in WM. The relative decrement in performance from single- to dual-task conditions was the same for experts and intermediate players. Also, the players all relied on the visuospatial subsystem suggesting that rather than simply utilizing activated representations from LTS, players construct a representation in the episodic buffer using information from both LTS and the visuospatial subsystem. It seems, therefore, that an expert will maintain task-relevant information in the subsystems along with activated mental models possibly in the episodic buffer. The difference between experts and novices may lie in

the demand put on the subsystems. The expert may need fewer links to be maintained in the subsystems than the novice. Importantly, expertise does not undermine the role of WM as is often assumed in the SA literature. Instead, expertise may change the way WM is used.

The Robbins et al. study suggests that representing spatial layouts and movements involves the spatial system and to some extent the central executive. Representing spatial layout and movements is an important part of maintaining SA while navigating and for monitoring traffic. Furthermore, representing spatial layout is particularly important for maintaining SA in air traffic control. This suggests a role for the spatial system in maintaining SA.

4.3.2 Inner speech and action control

Not much work has been done on linking the verbal system to complex cognitive tasks. Recently, however, the potential role of inner speech in controlling and monitoring behaviour has linked the verbal system to complex activity (see Baddeley et al, 2001; Emerson & Miyake, 2003; Miyake & Shah, 1999). The general idea is that the verbal subsystem may contain pointers to a larger action control program possibly in the episodic buffer (Baddeley et al, 2001). The notion of an inner speech mechanism for mediating thinking and guiding action was first brought up and studied by Vygotsky (1986; cited in Emerson & Miyake, 2003) and Luria (1961; cited in Emerson & Miyake, 2003). Vygotsky emphasized that inner speech works as a medium for thought and is important in the cognitive development of children. Luria (1961) went further, looking at the role of inner speech in controlling action. Recent studies in WM are pointing back to this idea of an inner speech mechanism that

controls behaviour. This idea of inner speech and the role of verbal WM is particularly interesting to consider in the light of complex task environments where operators have to remind themselves to monitor different systems or go back and forth between carrying out different tasks.

Baddeley et al. (2001) had participants perform simple counting tasks where they either added or subtracted the number 1 from a list of numbers. On a different trial however, participants had to alternate between adding and subtracting. The ability to switch between adding and subtracting when the signs were externally present was affected by a concurrent central executive task. But concurrent articulatory suppression had no effect. These findings are consistent with previous research suggesting that task switching depends on the central executive (Rubinstein, Meyer, & Evans, 2001). However, when the signs were not present to guide participants in alternating between adding and subtracting both a concurrent central executive task and a concurrent articulatory suppression affected participants' ability to switch. When the signs (plus and minus) were not visible, participants had to control their own behaviour reminding themselves to alternate between adding and subtracting. Baddeley et al. suggested that with the signs absent the participants utilized the verbal system as a self-cuing device to remind them which action had to be performed. They further speculated that the verbal system only supports a larger action control program possibly represented in the episodic buffer.

Emerson and Miyake (2003) extended the Baddeley et al. (2001) study using similar task procedures but different manipulations of the task. In one experiment it was found that even arbitrary colour cues will reduce the switching cost compared to no external signs. However, having participants switch between three task operations

rather than two made no difference in terms of the cost of switching. Similarly, having participants add or subtract one, three and six to the list of numbers did not alter the switching cost. In accord with Baddeley et al., Emerson and Miyake concluded that in the absence of signs, the verbal system facilitates switching. The inner speech mechanism is not directly responsible for switching which depends primarily on executive control processes. Rather, the inner speech mechanism indirectly supports the task of switching by activating a larger action or goal plan. It is possible that some sort of a verbal cueing mechanism or inner speech plays a role in many real life tasks. For example, the driver may maintain verbal cues such as *turn left at the first intersection* that activates a larger navigational plan maintained in the episodic buffer.

4.3.3 Text comprehension

What makes research on text comprehension particularly relevant to SA is the empirical link established between situation models and WM. It has been suggested in SA research that a situation model is the underlying representational structure. Although, text comprehension differs from a dynamic task environment such as driving in significant ways there are many similarities. Both tasks (driving and comprehending a text) require that information is kept track of, spatially, temporally and causally and integrated into an event, allowing comprehension and anticipation of further events. Research shows that in comprehending a text people seem to simultaneously monitor and integrate information along different dimensions (spatial, causal etc.) (Zwaan, Langston & Grasser, 1995; Zwaan & Radvansky, 1998). More importantly recent research has found evidence suggesting that causal and spatial information is processed and maintained in the verbal and spatial subsystems

respectively. Friedman and Miyake (2000) for example, found that increased spatial demands reduced the readers' ability to construct and maintain spatial information during comprehension of a text. Similarly, increased causal demands reduced the readers' ability to maintain and construct causal information. Furthermore, the ability to construct spatial structure correlated with spatial span measures whereas the ability to construct causal information correlated with verbal measures. The results therefore relate WM subsystems to the comprehension of texts. Other studies support the claim that the readers' ability to make certain inferences is based on the verbal system (Singer & Richot, 1996; Whitney, Ritchie & Clark, 1991). There is also some evidence suggesting that the visuospatial WM plays a role in the maintenance of spatial structure described in the text (Denis & Cocude, 1989). Shah and Miyake (1996) also found that a reading span task predicted performance on a comprehension task and conversely a spatial span task predicted performance on a spatial reasoning task. They argued that the two subsystems serve important and differentiated functions in higher-order cognition. Maintaining SA may therefore depend on utilizing both subsystems simultaneously.

4.3.4 Navigating

Important work on navigation actually supports the claim that in information rich environments where SA is an issue both subsystems are involved (Barshi & Healy, 2002; Garden et al., 2002). For example, Garden et al. (2002) found that learning a route through map segments depends on the visuospatial subsystem and does not involve the verbal subsystem: participants' ability to learn the route was significantly affected by a concurrent tapping task but not by a concurrent articulatory

suppression. Interestingly however, when participants were required to learn routes through a real city they relied on both visuospatial and verbal representations. Garden et al. suggested that in a complex multitask environment where there are multiple cues and information sources to be processed people involve both subsystems. In related research, Barshi and Healy (2002) also found that participants relied on both verbal and spatial representations when they were required to follow verbal instructions to navigate. This provides further support to the study by Aretz (1991). But in his study, when pilots had to navigate and control the aircraft and the navigation task became harder, they switched from using spatial resources to navigate to use verbal resources.

4.3.5 Dynamic environments

Studies on chess playing and navigating have important implications for SA. However, only in dynamic task environments where the participants are performing multiple tasks is the issue of WM and SA truly being addressed. Unfortunately, relatively few studies have actually looked at WM in a dynamic environment. Furthermore, studies that use dynamic environment often are addressing issues related to SA but do not include proper manipulation of WM. Or those studies specifically measure the impact of WM on performance but do not address the issue of SA. Reviewing the evidence however, provides important insight into how the fundamental mechanisms of WM support SA.

The most commonly used paradigm for studying WM in a dynamic environment is tracking of some sort. Research has shown that tracking interferes with visuospatial tasks such as keeping track of spatial locations, suggesting that it depends on the spatial system (Baddeley & Logie, 1999; Morris, Craik, & Gick,

1990). Conversely, there is a minimum interference on performance when people simultaneously perform tracking and a verbal task (Baddeley et al, 1991; Logie, Zucco & Baddeley, 1990). In a more complex tracking task paradigm however, participants seem to rely on both systems. Logie et al. (1989) for example, examined the role of WM in playing a complex computer game known as Space Fortress. The game involved a complex perceptuomotor control where the participant manoeuvred a space ship around the computer screen using a joystick. Early in training participants' performance was equally impaired by concurrent verbal and visuospatial secondary tasks. However, when participants had reached a higher level of expertise a different pattern of disruption was seen. The concurrent visuospatial task disrupted only aspects of the task that had to do with perceptuomotor control. Similarly, concurrent verbal tasks only impaired verbal parameters. Furthermore, a concurrent spatial tapping task, which had previously shown little effect, now impaired the participants' ability to control the joystick and perform other motor responses required by the game. What is interesting about the results from the Logie et al. study is that while participants were getting used to the game, general cognitive load was high. As they got more used to playing the game, participants started to utilize the two subsystems for performing different aspects of the task.

The space fortress task has been used in aviation to train pilots in attention management in multitask environments. In particular, the task has proven to be useful to train pilots in time-sharing attention when dealing with multiple tasks (Gopher, Weil, & Barket, 1994). When playing the game, participants must control and position the space ship, battle the fortress, collect bonus points and detect and respond to mines (Shebilske et al., 2000). The game therefore requires that participants prioritize

tasks, manage resources, and time share attention. As such, it involves SA. The results from Logie et al. study therefore suggest that maintaining SA may involve both the verbal and the spatial subsystem as well as the central executive. Similarly, work on text comprehension and navigation suggests that participants require both subsystems to carry out the tasks.

In an even more complex tracking environment, research has shown that driving also involves the WM mechanism. Lansdown et al. (2004) for example, found that introducing a secondary task seems to have negative impact on driving performance. Furthermore, performing multiple tasks rather than a single task affects people's ability to drive. In fact, the ability to switch attention (Kahneman et al., 1973) or selectively attend to tasks (Arthur et al., 1990) is correlated with accident involvement. This research suggests that the central executive is involved in driving for managing multiple tasks and controlling attention.

Recarte and Nunes (2000) had people perform either a verbal task or a spatial-imagery task while driving an instrumented car with advanced head-free eye tracking system. For the spatial-imagery task, participants judged whether imagined letters in the alphabet fulfilled certain conditions such as remaining unchanged when rotated vertically and horizontally. For the verbal task, participants repeated words starting with a certain letter that was provided by the researcher. Both tasks were found to increase pupil dilation suggesting that they consumed similar amount of attentional resources. The tasks however, selectively affected visual search during driving. When performing the spatial task, participants' fixation duration increased significantly compared to when they performed the verbal task. Similarly, the spatial task caused a decreased visual inspection window or 40% horizontal and 60% vertical reduction. In

comparison the verbal task caused 25% horizontal and 40% vertical reduction in visual inspection window. Although, SA was not measured, decreased visual search should reduce SA.

Recarte and Nunes (2003) further tested the relation among mental tasks, visual search, and detection and decision-making. As they pointed out, very little is known about how impaired visual search affects drivers' ability to detect things in the environment and make decisions. In order to test the link between visual behavior and detection and decision-making they had participants perform mental tasks while driving. Recarte and Nunes used a similar setup to the earlier experiment with the addition of having participants detect targets and make decisions on how to respond to the targets. Similar to their previous research (Recarte & Nunes, 2000) they measured pupil size as an indication of mental effort, spatial gaze variability as an indication of visual search behaviour, and the frequency of rear view mirrors and speedometer inspection. The visual detection task consisted of flashing spotlights that appeared at random location in space and at random time intervals. The participants' task was to distinguish between flashing rates (high and low) by pressing either left or right button on the steering wheel. The two mental (secondary tasks) consisted of two acquisition tasks and two production tasks. In each case one task had abstract content (verbal) and the other one concrete or description of spatial references to concrete objects (spatial). In the acquisition tasks the participants listened to an audio message for 2 minutes. They were informed that they needed to maintain the message in memory as they needed the information to perform a subsequent task. In the production tasks the participants listened to an audio message but this time had to reproduce the message they listened to. The dependent variables therefore included

both the difference between spatial and verbal tasks and between production and acquisition.

The results for the visual search behaviour were the same as in Recarte and Nunes (2000). All mental tasks increased pupil size, spatial gaze concentration, and reduced rear view mirror and speedometer inspection. Reproducing the message significantly decreased participants' visual search. Interestingly however, the acquisition tasks did not produce remarkable visual search changes. Probably participants were able to only memorize the main content. Therefore, the acquisition task did not cause much cognitive load. It was interesting that memorizing the spatial description did not have any impact on measured behavior. If participants used spatial resources for representing the described spatial relations it should have affected visual search similar to what Recarte and Nunes found in their earlier experiment. It is possible though that participants used verbal representations rather than spatial for the acquisition task regardless of the content. As Recarte and Nunes pointed out, driving with the visual detection was highly reliant on visuospatial resources and hence the participants had extra verbal resources to perform the acquisition tasks. This would be in accord with the WM model and with other findings.

Both target detection and correct responses were decreased by all the mental tasks. In particular, the production task affected participants' ability to detect and make correct decisions. Furthermore, detection was not caused by cognitive tunnelling as participants' ability to detect the target decreased regardless of whether the target was presented in the fovea or in the periphery. Therefore Recarte and Nunes concluded that mental tasks produce general interference in building up a sufficient SA to detect targets and respond to them. The nature of the mental task (verbal or

spatial) made no difference in the interference on any of the measured variables. It is possible that the task format, listening to certain instructions or verbal description, caused participants to rely more on verbal than spatial resources. The main shortcoming of the Recarte and Nune (2003) study is the emphasis on visual resources and the similarity of the mental tasks. To better understand the role of WM in SA while driving it is important to use spatial, verbal, and central executive load tasks and measure different aspects of maintaining SA while driving.

Gugerty (1997) measured SA in a driving simulation by having participants monitoring surrounding cars and responding to potential hazards such as approaching car or a car that suddenly entered the lane in front of the participants. In his experiment, participants did not actually drive but monitored traffic on a PC-based driving simulator showing 3-dimensional animated driving scenes on a computer screen. The participants watched the driving scenarios from the drivers seat and had full use of all mirrors for traffic behind them. They were instructed to imagine that their car was on autopilot hence their task did not include driving but was mainly based on monitoring traffic. On few trials the participants were required to override the autopilot mode in order to avoid hazards from surrounding cars by using the arrow keys on the keyboard. The 'up' and 'down' arrows meant that the participants' car would speed up or slow down. The 'right' and 'left' arrows meant that the participants' car would move to the right or the left lane respectively.

Gugerty used both implicit and explicit measures. As an implicit, performance-based measure, participants' response time to hazards from other cars was measured. The explicit recall method was based on the SAGAT technique (Endsley, 1995b). At the end of the trial, which lasted for 18 to 35 seconds, the screen

went blank and the participants were asked to locate the surrounding cars on a grid. The main finding here was that the more cars in the scene, decreased accuracy of recall and how well participants avoided hazards from surrounding cars. Furthermore, with heavier surrounding traffic, participants tended to allocate attention specifically to cars closer to their own vehicle. This was possibly done in order to attend to cars that were potentially hazardous. Gugerty concluded based on the results that SA of surrounding traffic depends on dynamic spatial memory.

The main problem with Gugerty's experiment is that without the right manipulation it is impossible to conclude whether SA of surrounding traffic depends on the spatial or the verbal system or involves the central executive system. In general, the studies reviewed in this section provide further support that WM plays a complex role in maintaining SA. It is still an empirical question as to how WM plays a role. None of the studies reviewed included both sufficient manipulation of WM tasks and sufficient measures of SA. Therefore, the relationship between WM and SA in a dynamic environment like driving has not yet been properly addressed empirically.

4.4 Potential role of WM in maintaining SA

Based on the evidence it is expected that in a complex task environment where SA is an issue, keeping track of, comprehending and anticipating events requires the processing and maintenance of information in both subsystems. The evidence reviewed in the section above suggests that although the spatial and the verbal system may be selectively involved in performing certain tasks, in complex task environments, both systems must be used. The central executive should also play an

important role in coordinating tasks, maintaining task relevant information and controlling attention. The central executive may also play a role in coordinating information from the subsystems and LTS.

Chapter 5. Empirical research

5.1 Overview

Four experiments were conducted to examine the relationship between SA and WM. In all four experiments a dual-task paradigm was used in which tracking tasks were combined with different WM load tasks. Tracking tasks are ideally suited for examining the relationship between WM and SA: To perform well at tracking, the operator must have good SA in that they are required to calibrate their inputs against the current and future status of the tracked object. The operator's SA is thus dependent on representing, updating and manipulating information in WM. The use of specific WM subsystems during tracking was assessed by introducing different WM load tasks.

Experiments 1 and 2 were designed to provide an initial link between tracking and the WM subsystems. In these experiments, two rectangles were presented on a display: a target rectangle and a participant-controlled rectangle. The target rectangle moved according to a computer-controlled algorithm and participants were required to keep the controlled rectangle on top of the moving target. The role of the WM subsystems was measured by introducing verbal, visuospatial, and central executive load tasks. These experiments showed that tracking of a visual object relies primarily on the visuospatial subsystem.

In Experiment 3 the tracking paradigm was extended to provide a more direct link between the WM subsystems and SA. In addition to tracking, this experiment measured the role of the WM subsystems in (a) predicting the future location of a tracked target and (b) searching for a tracked target. Experiment 3 provided strong support for the role of WM in maintaining SA. As in Experiments 1 and 2, tracking depended primarily on the

spatial subsystem. In addition, for the prediction and search components, the verbal and the spatial subsystems were involved, but for different aspects of performance.

Experiment 4 examined the involvement of the WM subsystems in (a) controlling a vehicle and (b) maintaining SA for information around the vehicle. In this experiment tracking involved maintaining a lane position (lateral tracking) in a driving simulation. The results showed that both the spatial subsystem and the central executive are used in maintaining lane position during driving. In addition, Experiment 4 showed that driver's SA for cars that are positioned around their vehicle differentially requires the spatial versus verbal subsystems in WM. Specifically, SA for cars positioned in front of a driver's vehicle relies on the spatial subsystem. In contrast, driver's SA for cars positioned behind their vehicle relies on the verbal subsystem.

5.2 Experiment 1

The goal of Experiment 1 was to provide an initial link between tracking and the verbal versus visuospatial WM subsystems. In this experiment, participants tracked a target rectangle on a display by controlling a second rectangle with a mouse. Their task was to keep the controlled rectangle on top of the target rectangle. Participants performed the tracking task on its own (single-task tracking condition) or concurrently (dual-task conditions) with either a verbal load or a visuospatial load.

For the verbal load, participants were required to indicate whether a number that was presented on the display increased or decreased in value. For the visuospatial load, participants indicated whether an arrow moved up or down against a graduated vertical

bar. Research suggests that monitoring numeric values primarily involves the verbal subsystem (Baddeley & Logie, 1999; Logie, Gilhooly & Wynn, 1994). In contrast, working with visual forms primarily involves the visuospatial system (Baddeley, 1986; Logie, 1995).

It was expected that there would be a concurrence cost where tracking performance would be adversely affected by the general resource demands of having to perform two tasks. Hence, tracking performance was expected to be generally worse in the dual-task conditions than in the single-task tracking condition. More importantly, on the assumption that tracking primarily involves the visuospatial subsystem in WM (Logie, Zucco & Baddeley, 1990; Morris, Craik & Gick, 1990), concurrent performance on the visuospatial load task should impair tracking significantly more than concurrent performance on the verbal load task.

5.2.1 Method

Participants

A total of 10 participants volunteered for this experiment, six females and four males. All participants were graduate students at Carleton University. Participants were paid \$10. Ages ranged from 24 to 40 years.

Apparatus, stimuli, and procedure

The experiment was controlled using E-prime software. The visual display was presented on a 17-inch high-resolution colour monitor. To provide enhanced realism to

the task, a map of the greater-Ottawa area was presented as background on the display (see Appendix B). This map was shown in grey colour/tones and moved vertically up the screen.

The tracking task included two stimuli: a pink target rectangle and a red controlled rectangle, 1 cm x 2 cm in size. These stimuli were superimposed over the moving map: the red and pink colours were easily distinguished from the background and from each other. The pink rectangle moved along both x-and y-axis according to a pre-defined loop which was independent of the background map movement. The red rectangle was controlled by the participant using a mouse: participants were instructed to use the mouse to keep the red rectangle on the top of the pink rectangle. Data for the X and Y position of the two rectangles was collected at 10 Hz (i.e., every 100 ms).

For the verbal-load task, five 3-digit numbers were presented one at a time. The digits used were: 500, 510, 520, 530, and 540. These numbers changed randomly every 2 seconds. Participants were instructed to monitor value changes in the number and to respond by using the “up” and “down” arrow keys on the keyboard. For example, if the number changed from 520 to 530, the correct response was to push the “up” arrow on the keyboard. If the number changed from 550 to 540, the participant should have pushed the “down” arrow.

The stimulus for the visuospatial-load task was a vertical bar. The vertical bar contained five values (lines). An arrow pointing towards the lines on the bar moved up and down randomly every 2 seconds. The task was configured such that most of the time the arrow moved between the three centre lines on the bar. This was done to reduce the

predictability of the task: when the arrow pointed towards the upper and lower lines it was obvious whether the arrow would move up or down next. Participants were instructed to monitor the vertical bar (located on the right-hand side of the display) and to indicate whether the arrow moved down or up by using the “up” and “down” arrow keys on the keyboard.

Task stimuli for both the verbal-load task and the visuospatial-load task were presented inside the red rectangle that the participants controlled. This was done to maximize the likelihood that the stimuli were within participants’ foveal view and within attentional focus at all times. Participants were to respond as quickly and accurately as possible to both tasks.

Participants performed tracking without a load task (single-task tracking) and tracking with each of the two load tasks (dual-task conditions). They also performed each load task without tracking (single-task verbal and single-task visuospatial load). Hence, there were a total of five conditions in the experiment. Each condition had ten 30-second trials. The five conditions were presented in a random order to each participant. In the dual-task conditions, participants were instructed to place equal emphasis on the tracking and the load task.

A one-minute practice trial was presented at the start of each condition. Participants initiated each condition by pressing the space bar. Once the condition started, the ten 30-second trials appeared one after the other with a 2-second pause between trials. When the trials for a condition were completed, instructions for the next condition were

presented on the display and the participant pushed the space bar when they were ready to continue.

5.2.2 Results

An alpha level of .05 was adopted throughout this research. For comparisons between individual conditions, 95% confidence intervals were used (Loftus & Masson, 1994). Rather than recalculating the difference between conditions based on planned comparison, confidence intervals allow for a simple (visual) heuristics where a difference between two conditions is judged to be significant when the confidence intervals of two conditions overlap by $\frac{1}{4}$ or less of the total interval. A significant difference between two conditions is also referred to as a critical difference and can be calculated by multiplying the confidence interval by the square root of 2.

RMSE tracking

Root means square error (RMSE) captures the difference between the position of the tracked rectangle and the position of the controlled rectangle at set points in time (every 100 ms). The values of the positions of the two rectangles are squared and the square root is taken of the mean difference between the two rectangles over time. RMSE scores for the x-and y-axis were combined and the data subjected to a one-way ANOVA with 3 levels of condition (single-task tracking, dual-task verbal and dual-task visuospatial) as the within-subject factor. RMSE varied with condition, $F(2, 18) = 15.62$, $MSE = 5.08$. As shown in Figure 4, tracking performance was worse in both of the dual-

task conditions compared to the single-task tracking condition. More importantly, tracking was impaired significantly more by concurrent performance on the visuospatial-load task than by concurrent performance on the verbal-load task. This supports the conclusion that tracking depends primarily on the visuospatial subsystem in WM.

Figure 4. RMSE for baseline (tracking only) and for tracking with verbal and visuospatial load tasks. The numbers presented in the bars represent the mean RMSE for each condition.

Load tasks

The latency and accuracy data for the two load tasks were analyzed in separate 2 (task: visuospatial vs. verbal) x 2 (condition: single vs. dual) ANOVAs with repeated measures on both factors. Accuracy rates were generally high (average of 95%), and did not differ significantly across the tasks or conditions, and thus are not discussed further.

Latencies were analyzed for correct responses only. For the latency data (see Figure 5), there was a main effect of task $F(1, 9) = 6.18$, $MSE = 6,777.67$, $\eta^2 = .29$ and of condition, $F(1, 9) = 89.32$, $MSE = 2046.51$, $\eta^2 = .63$. Follow up comparison using the 95% confidence intervals showed that participants responded more quickly in the single-task condition compared to the dual-task condition. This was true for both the verbal-load and the visuospatial-load tasks. Furthermore, participants responded more quickly to the visuospatial-load task than the verbal-load task. This finding suggests that the visuospatial load task was not more difficult than the verbal load task: accordingly, the greater impairment that was observed in tracking performance when combined with the visuospatial load cannot be simply attributed to differential difficulty across the two load tasks. Instead, it is concluded, that tracking requires the visuospatial subsystem in WM. The task by condition interaction was not significant.

Figure 5. Response time for verbal and spatial load tasks for single (load task only) and dual (load task with tracking) conditions. For the verbal task, participants responded to changes in a number. For the spatial task, participants responded to changes in a vertical bar.

5.2.3 Discussion

The primary finding in Experiment 1 was that performance on the tracking task was impaired significantly more by the visuospatial-load task than by the verbal-load task. This pattern occurred even though participants were generally faster at responding to the visuospatial-load task. The results support the conclusion that tracking involves the visuospatial subsystem to a greater extent than the verbal subsystem.

5.3 Experiment 2

Experiment 2 was designed to delimitate the role of the spatial versus the visual store component of the visuospatial subsystem in tracking. This is an important step

toward fully understanding the role of the WM subsystems in maintaining SA in tracking situations. Experiment 2 also examined whether tracking involves central executive resources: previous research has suggested a role for the central executive in tracking (Morris, et al., 1990).

Experiment 2 used the same tracking task as in Experiment 1. This task was combined with four WM load tasks: verbal, visual, spatial, and central executive. For the verbal-load task, participants were required to discriminate between consonant-vowel-consonant (CVC) combinations. The visual-load task involved discrimination of geometric shapes, a task that has been shown to depend on the visual but not the spatial system (Logie, 1995). The spatial-load task involved tapping a defined pattern on a keypad (Logie, 1995; Logie & Marchetti, 1991). The central executive load involved tapping a set of keys randomly: the central executive task was assumed to utilize spatial resources similar to tapping a defined pattern as in the spatial load task. In addition, however, random tapping uses central executive resources as the participant must monitor their own performance in order to avoid generating patterns (Baddeley et al., 1998).

The load tasks were expected to result in a general concurrence costs whereby tracking performance was worse in the dual-task conditions relative to the single-task tracking condition. Of primary interest, however, was whether specific and differential patterns of interference would be found. Based on the results of Experiment 1, it was expected that tracking performance would decrement more when combined with the spatial load than the verbal load. The central executive load should, at minimum, result in tracking decrements that are similar to those found with the spatial-load task because the

central executive and spatial load tasks have similar spatial processing requirements. An additional role for the central executive in tracking above and beyond the spatial subsystem would be reflected in greater interference than that observed with the spatial-load task. The extent to which the visual-load task would interfere with tracking was unknown. However, a different amount of tracking interference from the visual- versus the spatial-load tasks would support the notion that these are independent components of the visuospatial WM system.

5.3.1 Method

Participants

A total of 13 participants volunteered for this experiment, seven females and four males. All participants were graduate students at Carleton University. Participants were paid \$10 for their participation. Ages ranged from 26 to 40 years.

Apparatus, stimuli, and procedure

Participants performed the same tracking task as in Experiment 1 combined with one of 4 load tasks: discrimination of shapes (visual), rhyming (verbal), tapping a defined pattern on a keypad (spatial) and tapping randomly on a keypad (central executive). The tracking and load tasks were performed alone (single-task) and together (dual-task conditions). Hence, there were nine conditions for the experiment. Each condition consisted of ten 30-second trials. The nine conditions were presented in a random order to each participant. Data for x-and y-axis for tracking was collected at 10 Hz.

Visual load

For the visual load task, the stimuli consisted of visual forms selected from the SPSS marker set available in Microsoft Word . Prime symbols were chosen from the set and different variations made for each list. The SPSS marker set was used because it provides distinct symbols that do not resemble well-known forms such as a triangle or a square. Using more familiar symbols runs the risk of having the participants using verbal rehearsal when comparing the symbols. Furthermore, using symbols from the SPSS marker set made the discrimination task more challenging than if familiar symbols had been used.

The visual stimuli were presented inside the red rectangle that the participants controlled, in order to maximize the likelihood that the stimuli were within participants' foveal view and within attentional focus at all times. The stimuli were presented randomly every 2 seconds. There were five lists of fifteen symbols and each list appeared four times during the experiment; twice during the single-task condition (visual discrimination task without tracking) and twice during the dual-task condition (visual discrimination task and tracking). Participants were required to compare a series of symbols against the first symbol that appeared on each trial. Their task was to judge if each symbol was the same as or different from the first symbol presented. Participants responded by pressing on a keyboard '1' for same or '2' for different.

Verbal load

Consonant-vowel-consonant (CVC) syllables were presented visually and participants were required to determine whether the CVCs rhymed with the first CVC

presented on that trial. Responses were indicated by pressing on a keyboard '1' for rhyme and '2' for does not rhyme. The response requirements of this task were similar to those in the other load tasks in that the participants responded manually rather than verbally to the target stimuli.

Short CVC combinations (some words and some non-words) were used. There were five lists of fifteen CVC combinations each. As in the visual discrimination task, each list appeared four times; twice during the single-task condition (visual discrimination task without tracking) and twice during the dual-task condition (visual discrimination task and tracking). Each list had rhyming as well as non-rhyming CVCs. Approximately 15 to 20 rhyming CVC and five to ten non-rhyming CVC were chosen for each of the five lists. Fifteen CVC combinations were then randomly chosen for each list by using a computer program that randomly selected 15 CVCs for each list out of the available CVC combinations. Following the random selection, each list was checked to make sure that it included both rhyming and non-rhyming CVC combinations and alternated between the two at least twice per list.

Spatial load

Tapping out a defined pattern has been used successfully to engage the spatial WM subsystem (Logie et al., 2001; Robbins, et al., 1996). Accordingly, in the present experiment the spatial load task consisted of having participants tap a predefined pattern using the four arrow keys on a computer keyboard. Participants were instructed to tap the keys in a steady rhythm at approximately 1 key per second.

Central Executive load

The central executive is most often measured by random letter or number generation or by having people randomly generate days of the week or months of the year. Other methods that have been used to measure the central executive, include random tapping of keys (Baddeley et al., 1998) and random time intervals (Vandierendonck, De Vooght, & Van der Gotten, 1998; Vandierendonck, 2000). Random tapping of keys was chosen for the present experiment, but this method is well established in the WM literature as a mean of engaging the central executive system (Baddeley, et al., 1998). The same keys were used as for the spatial tapping task. Hence, contrary to common practice, only four keys were included in the random tapping task (it is common to use 10 keys). The reason for why so few keys were used in the central executive task is because participants were highly engaged visually with the primary task and instructed not to look at the keys they were tapping. Randomly tapping ten keys with no visual cues would be quite demanding (Baddeley, et al., 1998). Furthermore, the attempt was to keep the spatial tapping and the random tapping tasks as similar as possible. The central executive task therefore utilized some spatial resources. In addition, however, the task utilized resources from the central executive by requiring conscious monitoring to avoid repeating patterns.

*Overview of data scoring and analyses**Tracking*

For the tracking task, positions for both rectangles (computer controlled and the one controlled by the participants) were collected along both x-and y-axis. RMSE was then calculated for the difference between the tracked and the controlled rectangle on x- and y-axis combined. RMSE scores for tracking were subjected to a one-way ANOVA with load task (no task, visual, verbal, spatial and random tapping) as the within subject factor.

Load tasks

The data for the four load tasks were analyzed in two groups: the two discrimination tasks (visual and verbal) and the two tapping tasks (spatial and random). This was done so that the load tasks could be compared against each other as much as possible. Clearly the data for the tapping tasks was categorically different from the data obtained from the two discrimination tasks. Therefore, the visual and the verbal discrimination tasks were paired and the spatial and the random tapping tasks were paired for further analysis.

Comparing the two sets of load tasks allowed for a more complete interpretation of the overall data. In a dual task study, it is hard to control completely for the variability of the concurrent tasks. That is, one load task may interfere more with the primary task than another load task due to complexity rather than shared resources. Furthermore, a selective interference caused by one load task and not by another may be apparent by a

decrement in performance on either the tracking or the load task. Accordingly, comparing the load tasks (where feasible) in a common analysis may show whether the two variables (task and condition) interact. An interaction may indicate that (a) one load task is more difficult than the other load task and/or (b) selective interference occurs with the load task(s) per se, rather than or in addition to the tracking task.

Visual and verbal discrimination load tasks. Latency and accuracy data for the verbal and visual discrimination tasks was analyzed in a 2 (task: verbal vs. visual) x 2 (condition: single-task vs. dual-task) repeated measures ANOVA.

Spatial and central executive tapping tasks. For the spatial and central executive tapping tasks, the average time between key presses was measured and analyzed in a 2 (task: spatial vs. central executive) x 2 (condition: single-task vs. dual-task) repeated measures ANOVA. Accuracy for both the spatial and the central executive tasks was measured by calculating the balanced score for each key given the total number of key presses and comparing to the actual number of key presses for each key. The combined number of deviation from balanced count for each key was then calculated as a percentage of total number of key presses. This was done to control for variations in time and number of key presses across different conditions.

The procedure of comparing the number of key presses to a balanced count was obvious for spatial tapping where the participant's task was to follow a pattern where each key should have been tapped an equal number of times. Measuring accuracy for the central executive (random tapping) task was less straightforward. It is difficult to measure randomness. Vandierendonck (2000) suggests that rather than measuring the extent to

which participants are generating stimuli randomly, it is better to measure the generation of random time intervals. Random time intervals allow for a better empirical control of the executing task. Another solution to the problem of measuring randomness has been to view randomness as the opposite of generating patterns and sequences. On this approach, randomness is assessed by measuring the extent of patterns and sequences generated (Towse & Valentine, 1997). Viewing the present data however, it was obvious that with only four keys to tap, participants tended to repeatedly tap a single key or two keys when the task became more demanding (dual-task condition) as opposed to a more even distribution in a single task condition. That is, as task demand increased participants seemed to deviate more from even distribution. Therefore, it was assumed, that in avoiding repeating patterns participants should have tapped each of the four keys closed to equally over time. Hence, lack of randomness was determined by looking at how much participants deviated from a balanced score for each key.

5.3.2 Results

As in Experiment 1, an alpha of .05 was adopted and 95% confidence intervals were used for specific comparisons.

RMSE tracking

RMSE for lateral and vertical deviation in the tracking data was analyzed with a one-way ANOVA with condition (single-task, dual-task visual, dual-task verbal, dual-task spatial and dual-task central executive) as the within-subjects factor. The effect of

condition was significant, $F(4, 48) = 14.977$, $MSE = 2.516$. As shown in Figure 6, when combined with any of the four load tasks, tracking performance was impaired relative to tracking in the single-task condition. More interestingly, decrements in tracking performance were significantly greater when tracking was performed concurrently with the spatial-load and the central executive load tasks than with the visual- or verbal-load tasks. This suggests that tracking primarily relies on the spatial system rather than on the verbal subsystem or the temporary visual store. The fact that the central executive load had a similar impact on tracking as the spatial-load task indicates that the tracking task did not place a heavy demand on central executive resources.

Participants in experiment 2 were generally better at tracking compared to experiment 1. Hence, the RMSE scores are lower in experiment 2 compared to experiment 1. But more importantly, the pattern is the same in that the spatial stimuli caused greater decrement in tracking compared to the verbal stimuli in both experiments. Furthermore, both verbal and spatial stimuli in experiments 1 and 2 caused significantly greater decrement in performance compared to baseline condition (tracking only).

Figure 6. RMSE for baseline (tracking only) and for tracking with different load tasks

Load tasks

Visual and verbal discrimination

Accuracy rates for visual and verbal discrimination load tasks were generally high (average of 90%). A 2(task: visual vs. verbal) x 2(condition: single vs. dual) ANOVA revealed no significant differences in accuracy between the tasks. Latency data for correct responses was analyzed in a similar 2 x 2 ANOVA. There was only a significant main effect for condition, $F(1, 12) = 20.36$, $p < .05$, $MSE = 5421.6$, $\eta^2 = .63$. For both the visual and the verbal task, responses were faster in the single-task condition than in the dual-task condition. Importantly, the change in response latencies from single- to dual-task conditions was undifferentiated across the two load tasks.

Figure 7. Response time for verbal and visual discrimination tasks for single (discrimination task only) and dual (discrimination and tracking) conditions.

Spatial and random tapping

Time between each keystroke was analyzed in a 2 (task: spatial vs. central executive) x 2 (condition: single vs. dual) repeated measures ANOVA. A significant main effect of task, $F(1, 12) = 6.21$, $MSE = 4221.6$, $\eta^2 = .34$, showed that tapping a defined pattern on the keypad was significantly faster than tapping a random pattern. No other effects were significant, however, further analysis compared the individual conditions using the 95% confidence interval (see Figure 8). As can be seen in Figure 8, for the spatial task, there is a significant decrement in average tapping time from single- to dual-task conditions. However, for the central executive task, the difference in average tapping time between single- and dual-task conditions did not differ significantly. This might be due to the fact that random tapping in the single-task condition was already

challenging for the participants. Hence, adding the tracking caused no further decrement in performance than what was seen in the RMSE scores for tracking.

Figure 8. Time between keystrokes for spatial and random tapping tasks for single (tapping only) and dual (tapping and tracking) conditions.

Accuracy in the spatial and central executive tasks was analyzed in a 2(task: spatial vs. central executive) by 2(condition: single-task vs. dual-task) repeated measures ANOVA. There was a main effect for task, $F(1, 12) = 169.5$, $MSE = 18.89$. As expected there were generally more deviations from balanced count for random compared to spatial tapping. No other effects were significant.

5.3.3 Discussion

Experiment 1 established the role of the visuospatial system in tracking and ruled out the possibility that tracking merely utilizes general WM resources. Experiment 2 shows that tracking largely relies on processing and maintaining information in the inner scribe, rather than on the temporary visual store, the verbal system or on central executive resources.

5.4 Experiment 3

Experiment 3 extended the tracking paradigm used in Experiment 1 and 2 to provide a more direct link between the WM subsystems and SA. Specifically, Experiment 3 assessed the role of the WM subsystems in (a) predicting the future location of a tracked target and (b) searching for a tracked target. As noted above, prediction is considered to be an important characteristic of SA: in many tasks, an operator must rely on being able to anticipate future occurrences in order to react quickly and accurately as well as to reduce cognitive load.

The same tracking task from Experiments 1 and 2 was used. However, in addition to tracking, a search condition and a prediction condition were introduced. In these conditions the pink tracked rectangle periodically changed colour. A change to blue meant that the rectangle would disappear and then reappear in the lower right corner of the display. This is referred to as the prediction condition because the change from pink to blue was a consistent cue that would allow the participant to predict the future location of the tracked target. A change from pink to yellow meant that the tracked object would

disappear and then reappear in one of the four corners of the display (randomly determined). This is referred to as the search condition because the change to yellow indicated that the target would appear in a corner: the participant was required to search the corners of the display to find the target. Experiment 3 therefore, represented a task environment where fundamental aspects of SA were important. Specifically, participants were required to maintain a representation of task-relevant information in order to quickly activate predict or search activities. In addition, participants were required to use knowledge to predict the future location of the target and to coordinate the tracking task.

As in Experiments 1 and 2, tracking was expected to depend primarily on the spatial component of the visuospatial subsystem. An important aspect of Experiment 3, however, was that while tracking, participants were also required to maintain task-relevant information in WM (i.e., information regarding the colours and their meaning) that would enable them to engage either a prediction or a search strategy. It was expected that the spatial subsystem would play a strong role in the search condition because effective search requires a spatial representation of the display layout and an understanding of the relevant distances between the controlled rectangle and each corner of the display. This spatial representation would presumably enable the participant to quickly spot the target when it reappeared and importantly, to elicit the correct control input to quickly and accurately get to the display corner where the target appeared. Spatial layout and representation of movement have both been associated with the spatial subsystem (Logie et al., 2001; Quinn, 1991; Salway & Logie, 1995).

For the prediction condition it was expected that the verbal subsystem would play a strong role. Evidence suggests that the verbal subsystem plays a role in making causal inferences (Shah & Miyake, 1996) and representing information for task switching (Emerson & Miyake, 2003). Therefore, it was hypothesized that in order to predict the target location participants would maintain a cue in the verbal system that activates the relevant knowledge (if blue go to the lower right corner).

In general, it was expected that responses to shifts in target location would be faster in the prediction than in the search condition. Undifferentiated response times between the prediction and search conditions would indicate that the participants were unable to use the predictive information and stored action plans. It was also expected that the introduction of the prediction and search conditions would involve the central executive system in addition to the subsystems. In order to coordinate task-related knowledge and switch from tracking to predicting or searching for the tracked object, the central executive must be involved. In sum, the following hypotheses were examined in Experiment 3:

1. Tracking performance will be impaired more by spatial and central executive tasks than by visual and verbal tasks.
2. Participants will be generally faster to move to the correct corner when they can predict where the object will appear.
3. The ability to predict the target location will be impaired more by concurrent performance on the verbal task than the visual or spatial tasks.

4. The ability to search for the target will be impaired more by concurrent performance on the spatial task than the verbal or visual tasks.
5. The central executive will impair ability to search and predict in addition to the spatial and verbal tasks.

5.4.1 Method

Participants

A total of 17 undergraduate students (10 females, 7 males) from Carleton University volunteered for this experiment. Participants received 2 experimental credits for their participation.

Apparatus, Stimuli, Design and Procedure

The same apparatus and stimuli were used as in Experiment 2. The design and procedure differed from Experiment 2 in that participants were required to either search for or predict the occurrence of the target rectangle. During the tracking task, the tracked pink rectangle randomly changed colour. The colour change occurred once or twice during each 30-second trial. A change from pink to blue indicated that the rectangle would reappear in the lower right corner (prediction). A change from pink to yellow indicated that the rectangle would reappear in one of the four corners (search). Participants were instructed to move (as quickly as possible) the tracking (controlled) rectangle to the corner where the rectangle reappeared. They then followed the (blue or yellow) rectangle as it joined the defined tracking loop at which time the rectangle

became pink again and the normal tracking task continued for the remainder of the trial or until next colour change occurred. Otherwise, the instructions for tracking were the same as in Experiment 2. The time it took participants to move the controlled rectangle to the corner of the display where the tracked rectangle reappeared was measured

Statistical analysis

Statistical analysis was the same as in Experiment 2 except for the analysis of search and prediction. RMSE were calculated for the tracking of the pink rectangle but not for the period when participants were responding to the blue and yellow rectangles. Mean latencies associated with moving to the corner of the display where the target object appeared were analyzed in a 2 (search vs. prediction) by 5 (load tasks) repeated measures ANOVA.

The load tasks were analyzed as in Experiment 2 except for levels of condition. In Experiment 2, load tasks were compared across single-task and dual-task conditions. In Experiment 3, there was a single-task and three dual-task conditions. The three dual-task conditions included load task with tracking, load task with prediction and load task with search. Hence a 2 (task) by 4 (condition) ANOVA was used. As in Experiment 2, the visual and verbal discrimination tasks were compared and the two tapping tasks were compared.

5.4.2 Results

RMSE tracking

RMSE tracking was analyzed in a one-way ANOVA with load task (baseline, verbal, visual, spatial and central executive) as a within-subject factor. There was a significant effect of load tasks on tracking $F(4, 64) = 14.917$, $MSE = 5.051$, indicating that tracking was generally worse in the load task conditions than in the single-task tracking condition. In contrast to Experiment 2, the visual and the verbal tasks did not result in significantly greater decrement in tracking performance than the baseline condition. It is possible, that the additional demands of the search and prediction requirements made the tracking task harder such that performing the visual and verbal load tasks did not cause further decrement in tracking performance. More importantly, the spatial task resulted in significantly larger tracking RMSE as compared to the single-task condition as well as the verbal and the visual conditions. This result provides further support for the notion that tracking involves the spatial, movement-based subsystem of WM. The central executive task (random tapping) resulted in similar decrements in tracking performance to those in the spatial task condition, suggesting that tracking in Experiment 3 did not involve additional central executive resources.

Figure 9. RMSE for baseline (tracking only) and tracking with the different load tasks.

Prediction and Search

The time it took participants to move the controlled rectangle to the correct corner where the tracked rectangle reappeared was analyzed in 2 (task: prediction vs. search) by 5 (condition: single-task, dual-task visual, verbal, spatial and random) repeated-measures ANOVA. One subject was removed from the analysis because of a data collection loss in the verbal condition. As expected, there was a significant main effect of task, $F(1, 15) = 58.44$, $MSE = 39805.53$, and of condition, $F(4, 60) = 2.72$, $MSE = 10869.96$. The analysis also revealed a significant task by condition interaction $F(4, 60) = 7.24$, $MSE = 8511.12$.

Table 1: ANOVA for prediction and search

	<i>df</i>	<i>F</i>	η^2
Task	1	58.44*	.80
Error	15	(39805.51)	
Condition	4	2.72*	.15
Error	60	(10869.96)	
Task x Condition	4	7.24*	.33
Error	60	(8511.12)	

Note. Values enclosed in parentheses represent mean square errors.

* $p < .05$

A follow up comparison, using the 95% confidence intervals (Figure 10), shows that participants were quicker to respond when they could predict where the target rectangle was going to be compared to when they had to search for it. Concurrent performance on all but the central executive load tasks impaired participants' ability to predict the location of the tracked rectangle as compared to single-task tracking. This result suggests that general cognitive load affects peoples' ability to use knowledge to predict. Of primary interest was that the verbal task caused significantly more impairment in predicting the target location than the spatial task. This finding supports the hypothesis that predicting involves the verbal subsystem in WM. In contrast, the difference between the visual task and the verbal task was not significant, suggesting that in addition to the verbal resources prediction relied on some visual resources for representing colour related information. For the search task, only the spatial load task resulted in significant impairments relative to the single-task condition. This supports the hypothesis that searching for a target location depends on the spatial subsystem in WM.

Figure 10. RTs for moving the controlled rectangle to the corner where the tracked rectangle reappeared. Prediction: The tracked rectangle reappeared in the lower right corner. Search: The tracked rectangle reappeared in any of the four corners.

Load tasks

Visual and verbal discrimination

There were no significant decrements in accuracy rates for the verbal and visual tasks across the different conditions. Latencies for correct responses were analyzed with a 2 (task: visual vs. verbal) by 4 (condition: single vs. dual) repeated-measures ANOVA. As in Experiment 2, a main effect of task, $F(1, 16) = 13.759$, $MSE = 35974.68$, $\eta^2 = .46$ and condition, $F(3, 48) = 59.654$, $MSE = 7221.41$, $\eta^2 = .79$. Figure 10 shows that participants were significantly slower to respond to the verbal targets than to the visual targets in all conditions. Furthermore, participants were slower in responding to the target in all three dual-task conditions as compared to the single-task condition. The task by condition interaction was not significant.

Figure 11. Response times for verbal and visual discrimination tasks for different primary task conditions. Single conditions: Discrimination task without the tracking. Tracking: Discrimination task performed with tracking. Prediction: Discrimination task perform.

Tapping tasks (spatial and random)

A 2 (task: search vs. prediction) by 4 (condition: single-task tapping, tracking, dual-search, prediction) repeated measures ANOVA of the time between taps revealed a main effect of condition, $F(3, 48) = 14.338$, $MSE = 7410.33$, $\eta^2 := .47$. Compared to single-task tapping, the requirement to perform the search and prediction tasks resulted in greater decrements in tapping performance (spatial and random) than when tapping was performed with just tracking. Furthermore, spatial tapping was impaired more in the prediction condition than in the search condition. It is possible that in the prediction task the requirement to move the controlled rectangle to the bottom right location of the

display interfered with the specific spatial location(s) being tapped and patterned in the spatial tapping task.

Figure 12. Response times for spatial and random tapping tasks for different primary task conditions. Single conditions: Tapping without the tracking. Tracking: Tapping performed with tracking. Prediction: Tapping performed with the prediction task. Search: Tapping performed with the search task.

Accuracy of tapping was analyzed with a 2 (task: search vs. prediction) by 4 (condition: single-task tapping, tracking, dual-search, prediction) ANOVA. There was a main effect of task, $F(1,16) = 37.483$, $MSE = 234.10$, $\eta^2 = .04$. As expected there was less deviation from a balanced count for spatial than for random tapping (see Figure 13). There was also a main effect for condition, $F(3, 48) = 17.973$, $MSE = 40.52$, $\eta^2 = .47$. As shown in Figure 12 there was a significant increase in errors from single- to dual-task conditions for both the spatial and random tapping tasks. For spatial tapping, prediction caused significantly more decrements in accuracy as compared to the single-task

condition (tapping only) and to dual-task tapping with tracking. For random tapping there was a significant deviation from a balanced output of taps per key for all of the dual-task conditions as compared to the single-task condition. Furthermore, prediction and search both resulted in significantly more errors (11% and 12%, respectively) in tapping as compared to when tapping was combined with just tracking. This suggests that participants may have shed the task of randomly tapping in order to protect the task of searching for or predicting the tracked object. The data therefore supports the hypothesis that the central executive plays a role in search and prediction.

The difference between spatial and random tapping tasks as shown in Figure 13 suggests, as expected, that participants were not using the four keys equally in the random tapping task. Rather, as the overall task demand increased, participants seemed to tap specific keys more frequently than others.

Figure 13. Percentage of total key presses that deviate from balanced count (equal number of key presses for each key) for spatial and random tapping and different primary task conditions.

Overall, the results showed that prediction of a target location involves mainly the verbal subsystem whereas search for a target location primarily involves the spatial subsystem. In general, participants were faster to go to the correct corner when they could predict where the object was going to be compared to when they had to monitor all four corners. The results therefore suggest that even in a complex task situation, people can use knowledge to predict target location. Search and prediction interfered more with load task performance than did tracking alone. Presumably, both search and prediction add to the cognitive demands of tracking. In particular, simultaneously searching for the object or predicting its future location increased deviation from balanced count for the spatial and the central executive tasks. It is possible that the central executive is involved in

coordinating the tasks and allowing participants to switch from tracking to prediction or search.

5.4.3 Discussion

Experiment 3 supports the notion that WM is required to maintain SA in a complex task environment. Participants relied on the different components of WM to carry out the dynamic task. In particular, they used the verbal and the spatial subsystems selectively to predict versus search for the target rectangle. The results therefore provide further support to the WM literature where it has been shown that for tasks such as spatial reasoning and making causal inferences in text comprehension, participants use the spatial and the verbal subsystems, respectively (Shah & Miyake, 1996; Friedman & Miyake, 2000). Furthermore, Garden et al. (2002) found that when participants navigated in a real city, they used both visuospatial and verbal resources.

As expected, searching for the tracked object relied primarily on the spatial subsystem. This finding is in line with other WM research showing that representing location, movement and spatial layout depends on the inner scribe (Pickering, et al., 2001; Quinn, 1991; Quinn & Ralston, 1986; Salway & Logie, 1995). In contrast, prediction primarily involved the verbal subsystem. Recent evidence in the WM literature shows that the verbal subsystem plays a role in many complex tasks by maintaining cues that activate task-relevant knowledge in LTS (Baddeley et al., 2001; Emerson & Miyake, 2002). On this view, it is possible that participants in Experiment 3 maintained active cues in the verbal system that allowed them to quickly retrieve the task-relevant

knowledge (e.g., if blue move to the lower right corner). Quickly retrieving the relevant knowledge from LTS would require less cognitive resources than rehearsing the rule (if blue go to lower right corner) in verbal memory. The results from experiment 3 provide support to the claim that in a multi-task environment SA is supported by processing and maintaining information in the different WM systems. The results also suggest that the verbal system and verbal representation of information play a critical role in maintaining SA.

In general, participants were faster when they could predict where the object was going to be compared to when they had to search for it, suggesting that people are able to use knowledge to enhance performance even when they are engaged in other tasks. Unlike the search condition, the ability to predict was generally impaired by the load tasks. The results therefore provide some support to a theoretical claim made by Wickens (1995) who argued that the ability to predict future occurrences depends on constantly updating information in WM and calculating possible future scenarios. Hence, the ability to predict should be affected by general workload, as other tasks would compete for WM resources making it hard to reach level three SA (predicting).

The results showed that when predicting or searching for the target rectangle, participants generated significantly more errors in the random tapping task compared to a single-task condition (tapping only) or when they tapped with tracking only. This finding suggests that the ability to search or to predict effectively involves central executive resources. The central executive might play a role in allowing participants to maintain and access representations regarding what the target colours indicate. The central

executive system would also facilitate coordination allowing participants to quickly and readily switch from tracking to either prediction or search.

In sum, the results of Experiment 3 showed that the WM subsystems play different roles in maintaining SA in complex task environments. The results therefore provide support to the claim that there is a complex relationship between SA and the WM mechanism. Rather than functioning as a temporary store, WM supports SA by active and selective processing and maintenance of information. As such, Experiment 3 brings us closer to understanding the nature of the relationship between SA and WM.

5.5 Experiment 4

In experiment 4, the relationship between SA and WM was examined in a real-world driving task. The involvement of the WM subsystems were related to (a) controlling a simulated vehicle and (b) maintaining SA for information around the vehicle. As in Experiments 1-3, tracking was involved in that participants were required to maintain a lane position (lateral tracking) while driving. In addition, participants were required to maintain SA of the surrounding cars. This SA was assessed using a post-trial question (position grid) which the participants used to indicate the position of cars relative to their own vehicle at the time each trial ended.

Along with the primary driving task, participants performed one of three load tasks that tapped into the WM subsystems. To shorten the time of the experiment, only three of the four load tasks from in Experiments 2 and 3 were used. Based on the results of Experiments 2 and 3, Experiment 4 was focused on examining the role of the spatial

and verbal WM subsystems and the central executive. The visual discrimination task was not included in Experiment 4.

Based on the results of Experiments 1-3, it was hypothesized that maintaining lane position (lateral tracking) would depend primarily on processing and maintaining information in the spatial WM subsystem. In addition, the integration of several sources of information that is requisite to driving (e.g., lane position, speed, surrounding traffic) should implicate the central executive. It was expected, therefore, that maintaining a lane position would also be affected by a concurrent performance on the central executive task.

It was further hypothesized that a driver's SA for surrounding traffic will require both the verbal and the spatial WM subsystems, but for different aspects of performance. To maintain SA of surrounding traffic, participants must represent spatial locations and the location of other cars relative to their own vehicle. This suggests that maintaining SA of surrounding traffic requires spatial resources (Gugerty, 1997; Logie et al., 2001). However, the use of spatial resources is most relevant for SA of cars that are located in front of a driver's vehicle. A driver must monitor cars in front of their vehicle in order to optimize control and to react (e.g., brake) to changes in the relative positioning of cars in front. This monitoring involves frequent sampling of the spatial positions in the front view, and as such, logically implicates the spatial WM subsystem. The demands are quite different for SA of cars behind the driver's vehicle. Cars in the rear are not in direct view most of the time, as these must be intermittently sampled using the mirrors. Indeed, most drivers must remind themselves to regularly switch attention away from the front to view

the cars that are behind them. Thus, it was hypothesized that SA for cars in the rear of the driver relies primarily on the verbal subsystem. On this view, drivers presumably use verbal cues (a) to activate action control programs (Baddeley et al., 2001; Emerson & Miyake, 2003) that allows them to monitor the mirrors and cars behind their vehicle and (b) to generally represent the number and location of cars to the rear and whether the cars are close or far from their own vehicle. There is evidence that this type of location information can be represented verbally (Aretz, 1991; Barshi & Healy, 2002).

SA for cars in front and to the rear is expected to involve central executive resources in addition to the subsystems. This is because maintaining SA of surrounding cars depends on coordinated information of the task situation. Controlling the vehicle, deciding whether to speed up or slow down, is relevant to monitoring the surrounding cars. Furthermore, cars in the rear may be expected to overtake the driver's vehicle and as such, information regarding cars in the rear is relevant to what is happening in front and vice versa. Accessing a coordinated representation is controlled and facilitated by the central executive. Another role for the central executive is in facilitating the switch between monitoring cars in front and back. The task switch would be mediated by accessing larger action control program in LTS. In sum, the following hypotheses were tested in Experiment 4.

1. Maintaining a lane position (lateral tracking) will depend on the spatial subsystem.

2. Maintaining SA of surrounding traffic depends on both the verbal and the spatial systems. For cars in front, the spatial subsystem is primarily required. For cars to the rear, the verbal subsystem is primarily required.
3. The central executive will be implicated in all of the driving tasks.

5.5.1 Method

Participants

A total of 18 participants (9 females, 9 males) volunteered for this experiment. All participants were from the Carleton University community, had a valid drivers licence and a minimum of two years of driving experience. Participants were paid \$20 for their participation. The participants ranged in age from 26 to 41 years and had an average driving experience of 10.1 years.

Apparatus and Stimuli

Driving simulator

The experiment was conducted using a high fidelity, fully configured, DriveSafetyTM 500c fixed platform driving simulator. The simulator consisted of a quarter of a compact car located in front of five projection screens providing 21.8 degrees vertical and 150 degrees horizontal field of view. Participants had full use of the rear view mirror as well as both side mirrors providing them with a complete view of the traffic behind them. The simulator included computer-generated engine noise and external noise of passing traffic. The driving scenes were scripted using a TCL scripting

language that was executed under a PC-based Linux platform. The scenes simulated a freeway in an industrial area with two lanes going in each direction. Surrounding traffic was located on the two lanes driving in the same direction as the participants' car. There were either three or five cars placed around the participants' car. Each scene was scripted to stop at a random location after 30 and 50 seconds of run time. The scenes varied only in number and location of the surrounding cars and the type of objects included in the scene such as pedestrians and signs. The scenes were updated at a rate of 30 to 60 Hz.

The scenes were scripted to collect data regarding the actual location of the cars relative to the participants' vehicle at the end of each trial. The computerized measure (measured in car lengths) was compared to the subjective assessment of car location.

Load tasks

A keypad with four 1 cm x 1 cm keys placed in a row was fastened horizontally onto each participant's right leg. The keypad was used for responses to the three load tasks (verbal rhyming, spatial and random tapping). The verbal stimuli were the same as used in Experiments 2 and 3. The rhyming CVCs were projected onto the centre of the front screen. The distance between the verbal stimuli and cars in front was about 4 degrees: the rear view mirror was located about 6 degrees above and 2 degrees to the right of the verbal stimuli. The participant had to move their eyes away from the stimuli in order to monitor traffic in front and back of their vehicle.

Design and overview

Participants drove a series of short driving trials. Their task was to maintain a steady lane position in the right lane, maintain a steady speed of 60 km/h and monitor surrounding traffic. Participants performed driving only (baseline) as well as driving with one of the three load tasks: verbal, spatial and central executive. They also performed each load task without driving. Hence, there were 7 conditions in total. Each driving condition (baseline, verbal, spatial and central executive) had 18 trials, each lasting 30-50-seconds. For tapping, single task condition, participants tapped for 2, two and a half minute sessions. For the verbal task, the five lists of rhyming and non-rhyming CVC from Experiments 2 and 3 were used. Participants received the five lists of CVCs at a rate of one CVC every two seconds. The lists were then repeated. There was a short break between each list indicating that the list had finished and a new list was starting.

Adding load tasks made the driving task quite demanding. Thus, the seven conditions were divided into 3 experimental conditions. Each experimental condition started with 6 scenarios of driving only, followed by a load task (single task condition) and concluded with 18 scenarios of driving with the same load task. This ensured that participants always drove without a load task before driving while performing a load task. The order of the three experimental conditions was randomized between the 18 participants.

Procedure

Upon arrival participants were given an overview of the nature and aim of the experiment. They then practiced driving the simulator for 5 to 10 minutes. Once they were comfortable with driving the simulator, participants drove a single trial that simulated the experimental trials. Following the practice session the experiment started and participants completed the first of three parts. The following day, participants completed the other two parts of the experiment.

At the beginning of each experimental session participants completed 6 trials of driving only. There was then a short break while the experimenter explained the following load task, which was performed first on its own and then with the driving. At the start of each trial, participants put the car into drive and started driving. They were instructed to stay in the far right lane and to maintain a speed of 60 km/h. They were also told to monitor the surrounding traffic and be prepared to fill in the relative locations of the other cars at the end of each trial.

There were three load tasks. For the verbal rhyming task, the CVC combinations appeared on the screen in the participant's forward field of view (FFOV). Participants responded to the CVCs by indicating whether they rhymed or not with the first CVC that appeared. A keypad with four keys in a row was placed horizontally on participants' right leg and used to record their yes/no responses. The same keypad in the same position was used for the two tapping tasks. For spatial tapping subjects tapped out a defined pattern, from left to right and then in the reverse order. For the random tapping participants were instructed to tap the keys randomly, avoiding repeating any patterns.

During the single-task condition (load task without driving) the participants sat in the drivers seat with their left hand on the steering wheel and performed the load tasks using their right hand. In the dual-task condition (load task with driving) participants used their left hand to drive and their right hand to perform the load task. Participants were also instructed to use the left hand only when they drove without performing a load task to maintain consistency between conditions.

Each driving trial lasted for 30 to 50 seconds. At random intervals after 29 seconds, the simulation was frozen and the screen was blanked. After putting the car into park, participants were handed a grid map showing the four lanes and surrounding buildings (see Appendix A). Each grid represented one car length (4-5 meters) and the participants had to mark the location of the surrounding cars relative to their own car, which was shown on the grid map. Once the participants had filled out the grid map the next trial started.

Statistical analysis

RMSE for tracking and speed

Data for lane position was collected every at 30 Hz. RMSE scores were calculated based on deviation in centimetres from the centre of the lane. Data for speed was also collected at 30 Hz. RMSE speed was calculated as a deviation from 60 km/h.

The lane position and speed data were analyzed in separate 2 (load: 3 or 5 cars surrounding cars) by 4 (load task: baseline, verbal, spatial and central executive) repeated-measures ANOVAs.

SA data

Both x-and y-coordinates for each car were collected when the simulation was frozen at the end of each trial. The x-coordinate measured the distance in car lengths between each car and the participants' car. The y-coordinate measured whether the cars were positioned in the right or the left lane. This data was compared to the participant's evaluation using the grid positioning sheets. Scores were calculated as the difference between the computer value and the participants' estimation of car distance.

Estimation errors (difference between the computer value and participants estimation) ranged between 0 and 10 to 15 car lengths. Missed cars or added cars were therefore given the value of 15 (maximum error). Problems in the programming occasionally caused one of the cars not to start at the initiation of a trial. This resulted in extremely high distance errors (18-27) affecting the overall average. Due to the fact that these high values occurred rarely and that the total SA score for each condition for each participant was low, they were taken out of the analysis. It should be noted that missed cars (cars not noticed by the participants) were not among the values that were taken out. Furthermore, the first trial for each subject was not included in the analysis as the participants always started with the baseline (driving only) condition. Due to the fact that participants did not practice using the grid map during the practice session, the first trial often resulted in missed values. The missed values in the first trial caused the overall average of the baseline condition to be high relative to the other conditions. Therefore, for each participant the first trial (driving only) in the initial experimental condition served as a practice trial for using the grid map.

Data for distance error (difference between participants estimation and the actual distance of cars relative to own car) was analyzed in a 2 (load: 3 vs. 5 cars) x 2 (SA type: distance estimation for cars in front vs. behind the participants vehicle) x 4 (load tasks: baseline, verbal, spatial, central executive) repeated measures ANOVA.

Load tasks

Data was collected for the load tasks as described in Experiments 2 and 3. Due to problems in programming, the data for the three load tasks was only collected every 200 ms. Eight participants (4 females and 4 males) were therefore asked to repeat the experiment, this time with reaction time collected at 60 Hz or every 16.67 ms. The data for the 8 subjects was compared to the original data. The comparison revealed the same patterns and hence the data for the 8 participants in the second round was analyzed for the load tasks. Although, the original data for the load tasks (collected every 200 ms) revealed the expected pattern, it was decided to include only the data collected at 16.67 ms in the analysis for the load tasks. The original data was used for analyzing all other variables.

Data for the verbal rhyming was subjected to one-way repeated-measures ANOVA of condition (single-task, dual with 3 cars, dual with 5 cars). The tapping data was analyzed in a 2(task: spatial vs. central executive) x 3 (condition: single, dual with 3 cars dual with 5 cars) ANOVA.

5.5.2 Results

RMSE for tracking

The analysis revealed a main effect of task, $F(3, 33) = 3.464$, $MSE = 60.30$. No other effects were significant. A follow up comparison of the conditions (Figure 14) shows that performing the verbal task did not cause significantly greater deviations in lane position compared to driving alone. However, as predicted, concurrent performance on both the spatial and the central executive tasks significantly increased deviations in lane position as compared to the baseline driving condition (and compared to driving while performing the verbal task). This result therefore supports the hypothesis that the complex perceptuomotor tracking task involved in controlling a vehicle depends on processing information in the spatial WM subsystem. The central executive task (random tapping) resulted in significantly greater deviations in lane position than the spatial task, indicating that controlling the lane position of a vehicle involves not only spatial resources but also central executive resources.

Figure 14. RMSE for baseline (driving only) and for driving with different load tasks.

Speed control

The analysis of RMSE scores for controlling speed did not show any significant effects.

Load tasks

Verbal task

The analysis revealed a significant effect of condition, $F(2, 14) = 5.765$, $MSE = 2020.77$, $\eta^2 = .46$, where participants were faster to respond to the CVC combinations in the single-task (607 ms) than in the dual-task conditions (659 ms for three car and 682 ms for five car load). As shown in Figure 15, latencies to the verbal task were significantly slower in both of the dual-tasks conditions (3 cars and 5 cars) compared to the single-task

(verbal task only) condition. There was no significant difference in latencies to the verbal task across load (3 or 5 cars).

Figure 15. Response time for verbal discrimination task for baseline (discrimination task only) and for different load conditions (3 cars vs. 5 cars).

Tapping tasks

Time between taps was analyzed in a 2(task: spatial vs. central executive) by 3(condition: single-task, dual-tasks with 3 cars and dual-task with 5 cars) ANOVA. There were no significant effects in time between taps across the spatial task versus central executive tasks or between single and dual conditions. Participants maintained a steady speed in tapping throughout.

The analysis of the accuracy data revealed a main effect for task, $F(1, 7) = 37.260$, $MSE = 1101.56$, $\eta^2 = .04$. As shown in Figure 16, participants deviated more from a balanced count for the random tapping compared to the spatial tapping. No other effects

were significant. The follow up comparison using the 95% confidence intervals showed that for both spatial and random tapping, deviations from balanced count were significantly increased in the dual-task condition compared to the single- task condition.

Figure 16. Percentage of total key presses that deviate from balanced count (equal number of key presses for each key) for spatial and random tapping for single (tapping only) and dual (tapping and driving) conditions.

SA measure

The data for mean deviation in distance estimation was analyzed in a 2 (SA type: cars in front versus cars behind) x 2 (load: 3 cars vs. 5 cars) x 4 (load task: baseline, verbal, spatial, random) repeated-measures ANOVA. A main effect of load, $F(1,17) = 59.51$, $MSE = 1.33$, showed that estimated positioning of cars deviated more in the 5-car than the 3-car condition. There were significant type x load, $F(1, 17) = 122.19$, $MSE = 1.04$ and type x task, $F(3, 51) = 0.73$, $MSE = 3.05$ interactions. To further probe these

interactions, separate 2 (load) x 4 (load task) analyses were performed on the two types of SA data: cars in front and cars behind.

Table 2: ANOVA for SA of cars in front and back

<i>Source</i>	<i>df</i>	<i>F</i>	η^2
Type	1	.49	.03
Error	17	(1.86)	
Load	1	59.51*	.78
Error	17	(1.33)	
Task	3	2.37	.12
Error	51	(.56)	
Type x Load	1	122.19*	.88
Error	17	(1.04)	
Type x Task	3	.73*	.15
Error	51	(3.05)	
Load x Task	3	1.09	.06
Error	51	(.44)	
Type x Task x Load	3	1.38	.08
Error	51	(.39)	

Note. Values enclosed in parentheses represent mean square errors.

* $p < .05$

The analysis of SA for cars located behind the driver's vehicle revealed a main effect for load, $F(1, 17) = 102.63$, $MSE = 1.98$, where participants were significantly worse at estimating distances for cars in back when there were 5 cars compared to when there were 3 cars. Hence, workload (more cars) affected participants' ability to accurately estimate car positions for cars that were located behind the driver. There was also a main

effect of load task, $F(3, 51) = 2.87$, $MSE = 1,01$. Participants were significantly worse at estimating car locations in the rear when concurrently performing the verbal task and the central executive task compared to driving only and the spatial task

Table 3: ANOVA for SA of cars in back

<i>Source</i>	<i>df</i>	<i>F</i>	<i>η^2</i>
Task	3	2.87*	.14
Error	51	(1.01)	
Load	1	102.63*	.86
Error	17	(1.98)	
Task x Load	3	1.16	.06
Error	51	(.55)	

Note. Values enclosed in parentheses represent mean square errors.

* $p < .05$

The results therefore support the hypothesis that SA for cars located behind the driver's vehicle relies on the verbal WM subsystem. In addition, SA for cars in the rear requires central executive resources, suggesting that drivers were activating an action control program to monitor and guide their behaviour.

Figure 17. Deviation in estimating positions of cars behind the participants' car measured in car lengths for baseline (driving only) and for driving with different load tasks.

Analysis of SA for cars in front showed a main effect for load, $F(1, 17) = 7.14$, $MSE = .39$, where participants were significantly worse at estimating car distances when there were 3 cars compared to when there were 5 cars. There was also a main effect of load task, $F(3, 51) = 2.31$, $MSE = .28$, where participants were significantly worse at estimating car location when performing the spatial task as compared to either no load task (baseline) or the verbal or central executive tasks.

Table 4: ANOVA for SA of cars in front

<i>Source</i>	<i>df</i>	<i>F</i>	η^2
Task	3	2.31*	.12
Error	51	(.27)	
Load	1	7.14*	.30
Error	17	(.39)	
Task x Load	3	1.34	.07
Error	51	(.28)	

Note. Values enclosed in parentheses represent mean square errors.

* $p < .05$

The verbal and central executive load tasks did not significantly affect SA relative to the baseline driving condition. As shown in Figure 18, participants were somewhat worse at estimating distances when there were 3 cars compared to 5 cars, but only in the verbal load-task condition. This may be an artefact of the study insofar as the overall distance between surrounding cars and the participants' car was greater for cars in front when there were 3 cars compared to 5 cars.

Figure 18. Deviation in estimating positions of cars in front of the participants' car measured in car lengths for baseline (driving only) and for driving with different load tasks.

5.5.3 Discussion

Experiment 4 shows that the WM subsystems play specific roles in maintaining SA in a complex environment such as driving. As hypothesized, maintaining lane position relies on the spatial WM subsystem, thereby linking the task of controlling the position of a car to the perceptuomotor tracking task performed in previous experiments. Experiment 4 shows that maintenance of lane position also involves the central executive system. This is expected: driving is more cognitively demanding compared to the perceptuomotor task used in the previous experiments. The driver needs to utilize information regarding speed and distance of the other cars, conditions of the road and

possibly prior knowledge regarding driving. Controlling the car therefore depends on activation and scheduling of information in the central executive.

An unique contribution of Experiment 4 is in showing that SA for cars located behind the driver requires the verbal subsystem, whereas SA for cars located in front requires the spatial subsystem. This finding provides an important extension to previous research where SA for surrounding traffic in realistic tasks such as flying (Aretz, 1991; Endsley, 2000; Wickens, 2000) or driving (Gugerty, 1997) has been assumed to depend on spatial resources only. The present evidence showing an involvement of the verbal subsystem for representing cars behind the driver is particularly interesting. WM research is recently starting to connect the verbal system to more complex processing and representation of information (see Miyake & Shah, 1999). Shah and Miyake (1996) found that the verbal system maintains fairly complex information for causal inferences. Also, research has linked the verbal system to inner speech mechanisms that allows participants to switch between tasks and represent information that is not externally available (Baddeley et al., 2001). Cars in back are not in view most of the time. Rather, participants must periodically monitor the mirrors to keep track of cars behind them. It is possible that participants used verbal cues that activated larger action control program allowing them to switch between tasks and monitor their own behaviour. Baddeley et al. (2001) found that participants used the verbal system to monitor which tasks had been performed and also to represent verbally the information of the different task requirements. Activating a larger action control program may allow participants to regularly guide their monitoring activity and switch from monitoring cars in front to monitoring cars in back. It is also

possible for cars in the rear, information such as how many cars and whether they were close is maintained in the verbal system.

Experiment 4 also showed that the central executive played a role at all task levels, including lane positioning, monitoring cars in back and in front. The results therefore support the claim that maintaining SA while driving depends not only on the WM subsystems, but also on the central executive for scheduling and coordinating task-relevant information.

5.6 General Discussion

In a series of four experiments, I examined the relation between SA and WM. All four experiments used a dual-task paradigm in which tracking tasks were combined with different WM load tasks. Experiment 1 established the role of the visuospatial WM system in the perceptuomotor tracking task used in the current research and ruled out the possibility that tracking utilizes general WM resources. Participants' ability to track a moving object against a dynamic background was affected significantly more by a concurrent visuospatial task compared to a concurrent verbal task. This was true even though participants were faster at responding to the visuospatial load task than the verbal load task. Experiment 2 showed that tracking largely relies on processing and maintaining information in the spatial system, rather than on the temporary visual store, the verbal system or on central executive resources.

In Experiment 3 the tracking paradigm was extended to provide a more direct link between the WM subsystems and SA. In addition to tracking, participants had to predict

the future location of the tracked object and search for it. Overall, the results provided strong support for the role of WM in maintaining SA. Participants relied on the different components of WM to perform the tasks. Similar to Experiments 1 and 2, tracking required the spatial system. Furthermore, participants used the verbal system to predict where the tracked object was going to be and the spatial system to search for it. The central executive played a significant role in facilitating the transition from the tracking to prediction and search conditions. In general, participants were faster when they could use knowledge to predict where the target object was going to reappear than when they had to search for it. This finding suggests that people are able to use knowledge to enhance performance when performing other tasks.

In Experiment 4, the relationship between SA and WM subsystems was examined in a simulated driving task. Participants were required to maintain a lane position (tracking) and to maintain SA of surrounding traffic. Similar to Experiment 3 the results supported the role of WM in maintaining SA. In particular, maintaining SA of cars in front and behind the driver selectively involved the spatial and verbal subsystems in WM. Specifically, SA for cars positioned in front of a driver's vehicle relied on the spatial subsystem. In contrast, driver's SA for cars positioned behind their vehicle relied on the verbal subsystem. Tracking, as in controlling the simulated car, relied on the spatial system. Finally, the central executive was involved in all aspects of task performance.

Experiments 1 and 2 established the role of WM subsystems in tracking and therefore provided the initial link between WM and SA. There was an important parallel between tracking in Experiments 1-3 and Experiment 4 where tracking involved

controlling a simulated car. In all four experiments tracking required the spatial subsystem. These results therefore, further supports the link between simple perceptuomotor tracking and more complex control environments such as driving or controlling an aircraft (see for example, Wickens & Hollands, 1999). But tracking as in controlling the simulated car also involved the central executive. It is possible that controlling a car (driving) depends on a coordinated representation of related information such as movement and speeds of surrounding cars. Such representation would also provide information on road conditions, pedestrians or traffic lights and even stored knowledge from LTS.

The results of Experiments 3 and 4 showed that the WM subsystems play differential roles in maintaining SA and thus support the central argument of this dissertation that there is a complex relationship between SA and WM subsystems. Rather than functioning as a temporary store as suggested by Endsley (1995a), WM supports SA by active and selective processing and maintenance of information.

5.6.1 SA and the subsystems

A unique contribution of the present research was to show that the verbal subsystem plays an active role in supporting the ability to predict future location of a tracked object (Experiment 3) and maintain SA of cars positioned behind the driver (Experiment 4). Previous studies have not specified how or if the verbal system plays a role in maintaining SA (see for example, Aretz, 1991). The present research also showed that searching for a tracked object (Experiment 3) and monitoring cars positioned in front

of the driver (Experiment 4) requires the spatial subsystem. These findings provide further support to the WM literature where it is suggested that the verbal and the spatial subsystems play a selective role in many complex tasks (e.g. Garden et al., 2002; Friedman & Miyake, 2000; Robbins, 1996).

Adams et al. (1995) argued that one of the puzzles in SA research is how the operator is able to keep track of the multiple tasks that need to be performed and in what order they are to be performed. It is possible that the verbal system plays a critical role in maintaining SA by allowing operators in rich task environments to cue themselves to remember what to perform and when depending on a larger action program in the episodic buffer and/or LTS. This would allow the operators to remind themselves of monitoring the area and the system and in which order to perform the tasks. Therefore, maintaining SA of cars in back may have involved the verbal system allowing participants to regularly guide their monitoring activity and switch from monitoring cars in front to monitoring cars in back. Such a verbal cuing mechanism might be used to navigate and to predict future occurrences of a target object or event based on what the operator knows about the environment. In Experiment 3 for example, the ability to predict where the tracked object was going to be required the verbal subsystem. It was speculated that rather than rehearsing a rule, which would require considerable cognitive resources, it is more likely that the participants maintained active cues in the verbal system that allowed them to quickly retrieve the task relevant knowledge (e.g., if blue move to the lower right corner). The results from Experiment 3 therefore link the verbal system to high level SA (level 3). Although predicting future occurrences may be based

on the spatial system, Experiment 3 suggests that in cases where the prediction is based on prior knowledge about the task environment it depends on the verbal system.

Good SA relies largely on being able to anticipate future occurrences, allowing operators to react quickly and accurately to stimuli in the task environment. Wickens (1995) argued that predicting must depend on calculating and updating information in WM. Therefore, reaching level 3 SA (prediction) should require considerable cognitive resources. In contrast Endsley (1995a) argued that for the expert operator, predicting is not necessarily affected by cognitive load as the expert operator accesses stored mental models in LTS that largely provide the needed understanding to make a prediction. The results from Experiment 3 suggest that the ability to predict involves general cognitive load. Predicting was impaired by all three load tasks compared to baseline condition (tracking only). In contrast, only the spatial load task impaired participants' ability to search. The results therefore suggest more involvement of the WM mechanism in maintaining SA (predicting) than what is argued by Endsley.

The verbal system may also play a role in maintaining information such as how many cars and whether they are close and so on. The obvious role of the central executive for maintaining awareness for cars behind the driver however, suggests some sort of action control program is activated during driving. Both Baddeley et al. (2001) and Emerson and Miykae (2003) argued that an action control program activated by cues in the verbal system would necessarily require central executive resources. More work is needed to further understand the possible role of an action control program to support SA. For example, another way in which the verbal mechanism might play an important

role in supporting SA is by activating and maintaining task goals. As Emerson and Miyake (2003) argued, the verbal system may be involved in maintaining verbal cues that activate a larger goal program for the general task environment maintained in LTS.

In both Experiments 3 and 4, participants selectively used the verbal and the spatial system to perform different aspects of the task. In addition to the verbal system, the spatial system facilitated searching for the tracked object in Experiment 3 and maintaining SA of cars in front in Experiment 4. The focus in the field has been more on the role of spatial resources (Endsley, 1995a; Wickens, 2000) and spatial memory (Gugerty, 1997) than on verbal resources. In fact, spatial awareness is identified and measured in order to understand how operators orient themselves in space (Aretz, 1991; Williams, 2002). However, some studies have suggested a more complex role of WM in maintaining SA. Caretta et al. (1996) for example, found correlations between SA and spatial and verbal WM tasks as well as between SA and spatial reasoning tasks. Gugerty and Tirre (1995) also found correlations between awareness of surrounding traffic in driving simulation and various WM measures. Similarly, Aretz (1991) found that when pilots had to control a simulated aircraft and perform a difficult navigation task they switched from using spatial to using verbal resources for the navigation task. The present study extends those findings by showing that SA is supported by maintaining and updating information in both subsystems. The present research furthermore portrays the role of the subsystems in different SA tasks.

5.6.2 SA and the central executive

The involvement of the central executive in the four experiments has to be considered by looking at the overall results. The central executive for example, did not interfere more with tracking in Experiments 2 and 3 than the spatial tapping task. There is little evidence to support the role of the central executive in a pursuit tracking task (although see Morris et al., 1990). Very few experiments have in fact examined the role of WM in tracking. Those experiments tend to examine the spatial and the verbal subsystems in the context of tracking not the central executive (Baddeley, et al., 1991; Baddeley & Logie, 1999; Logie, Zucco, & Baddeley, 1990). But, given that a central executive task is generally harder to perform than either a spatial or a verbal task that do not involve the central executive resources it is reasonable to assume that the central executive task would impair tracking more than tasks involving the subsystems. However, tasks that involve the central executive vary in their cognitive demands. Research suggests that randomly generating words is more demanding than random tapping keys (Towse, 1998). Also, random tapping without visual cues is harder than random tapping where the participants can watch the keys as they tap (Baddeley et al., 1998). In all four experiments in the present research participants randomly tapped four keys without any visual cues that might facilitate the tapping.

It is possible that a pursuit tracking task does not involve the central executive as the results in Experiments 2 and 3 suggest. Or the random tapping task that was used in the present research may not have been demanding enough to have interfered with the tracking. The third possibility is that participants were not actually tapping randomly and

therefore the random tapping task was involving the spatial subsystem rather than the central executive. However, looking at the results from Experiments 3 and 4, it is unlikely that participants were not tapping randomly. For example, in Experiment 3 participants seemed to shed the random tapping task when they switched from tracking to prediction or search. While performing either the prediction or the search task participants seemed to tap specific keys more frequently than others indicating that when the primary task became more demanding the central executive task had more effect on performance than in less demanding conditions (e.g. tracking). Experiment 4, further supports the claim that participants were tapping randomly but that the random tapping task had more impact on performance when the overall task demand increased. When driving in the driving simulator (a more cognitively demanding task than tracking), the random tapping task significantly impaired participants' ability to maintaining central lane position compared to the spatial tapping task. The random tapping task also impaired the participants' ability to monitor surrounding traffic.

The results therefore suggest that the central executive may play an insignificant role in tracking but is important for supporting SA in more complex tracking task paradigm. Previous research has shown that the ability to perform multiple tasks is linked to better performance in the air (Damos, 1993) and superior performance of elite pilots (O'Hare, 1997). In particular, task switching and the ability to ignore distracting information (Tham & Kramer, 1994), both central executive functions, are important for superior piloting. The existing evidence therefore associates central executive control to flight performance but does not specify how the central executive is connected to SA.

Experiment 3 and 4, provide some insight into the potential role of the central executive in maintaining SA. In Experiment 3 for example, switching from tracking to prediction and search caused significant decrement in performance on the central executive task. It is possible that in order to perform the tasks participants involved the central executive to maintain and access representations of the meaning of each of the target colours. The central executive may also have been involved in coordinating the tasks, allowing the participants to quickly and readily switching from one task to another.

The central executive was involved in all aspects of the task in Experiment 4. This complex involvement of the central executive system may suggest that driving and maintaining SA depend on a coordinated representation of the task environment. Such coordinated representation would provide the participants with a coherent understanding of the driving scene and therefore facilitate performance on the individual tasks. For example, controlling the car, the driver may use information on surrounding traffic such as speed and distance of other cars. Similarly, for awareness of cars in front, information regarding cars in back (e.g., whether they are likely to overtake) is relevant.

As discussed in chapter 4, the central executive is assumed to be responsible for coordinating information from the subsystems and LTS (see for example, Baddeley, 1996). Baddeley, (2000) also proposed that the episodic buffer is the “place” where information is integrated and maintained. Very little work however, has been done in terms of further understanding how representations in the episodic buffer contribute to task performance. This is largely because tasks that are being used in WM research are often simple where only a single aspect of performance is being observed. In contrast the

task in experiment 4 had multiple dimensions to it that were all connected. Hence, performance on one aspect of the task (driving) affected performance on another aspect (SA). The coordinated representation in the episodic buffer is particularly relevant to SA. In fact, such representation may be what provides the operator with the coherent awareness and understanding of a task situation. The coordinated representation is constantly updated by processing information in the WM subsystems. However, only by integrating the information from the subsystems can the operator have access to a coherent representation of the task environment.

Although, the present work provides important insight into the potential role of the central executive in supporting SA much more work is needed to further specify the relation between SA and the central executive system. The evidence suggests that participants were involving the central executive in the random tapping task. However, using only four keys may not be demanding enough to fully understand the role of the central executive in dynamic task environment. Further research therefore could use different tasks to involve the central executive in the context of SA. In particular, it is difficult to evaluate randomness with only four keys. Participants were instructed to avoid generating any pattern. The instructions affected the participants' performance such that they tapped the four keys more equally in single-task conditions than in dual-task conditions where they tapped specific keys more frequently than others. Vandierendonck (2000) has suggested that a better methodology to measure randomness is to have participants tap at random time intervals.

The present work provides important information for both basic and applied research. The evidence certainly supports the approach of trying to understand SA in terms of how it is maintained by the underlying mechanism. This has some implications for SA theory that will be explored in the next chapter. The evidence also has some implications for WM research in dynamic environments that will be discussed briefly in the next section.

5.6.3 Implications for WM research and dynamic environments

More and more research is starting to study WM in a dynamic task environment like driving. This research however, often does not measure factors related to SA. Rather, WM has been tested in terms of task related performance such as visual search, braking time and so on (Lansdown et al., 2004; Recarte & Nune, 2000). In fact, Recarte and Nune (2003) argued that wider cognitive measures, including measures of SA, will provide a more complete understanding of the role of WM in driving and of the relevant cognitive factors underlying driving.

The present research contributes to WM research by linking the WM mechanism to the applied concept of SA. SA provides the needed conceptualization to study the role of WM in complex cognition in a dynamic environment. The present work also, provides some insight into the task-specific role of the WM subsystems during driving. In particular, researchers have not always been in a position to fully interpret the results within the multi component model of WM. Recarte and Nune (2000) for example, found that a spatial-imagery task affects visual search behaviour more than a verbal task. They

concluded that visual search while driving relies on spatial resources. In a follow up study where participants listen to concrete (spatial) versus abstract (verbal) content, there was no difference in the impact of spatial or verbal task on visual search, target detection or decision-making (Recarte & Nunes, 2003). However, differences between the spatial and verbal tasks in the later study was not enough to ensure that listening to the concrete (spatial) content for example, was not involving the verbal system. The study therefore did not allow for a complete interpretation of the potential role of WM in cognitive processing during driving.

Gugerty (1997) found that more cars presented in a driving simulation affected participants' ability to accurately recall car location. He concluded that maintaining SA of surrounding traffic relies on the spatial memory. However, lack of proper WM manipulation does not allow Gugerty to rule out the role of the verbal system and the central executive in driving or to properly understand the complex contribution of the two subsystems in performing different aspects of the driving task.

The present research is an important contribution to further understand the role of WM in dynamic task environments such as driving. The evidence showed that driving involves both the verbal and spatial subsystem as well as the central executive. In fact, it is possible that driving the car and maintaining SA relied on a coordinated representation of the driving scene in the episodic buffer. The role of the verbal WM is particularly interesting as driving research and applied research in general focuses almost entirely on spatial resources. Very little work has been done in understanding the possible contribution of the verbal system in comprehending and monitoring events while driving.

Furthermore, the concept of SA contributes to this research by allowing researchers to look at the role of WM in more complex cognitive processes during driving.

Chapter 6. Conclusions: Implications for SA theory and research

6.1 Summary of the dissertation work

The goal of this dissertation was to understand the relationship between SA and the fundamental mechanisms of WM. SA has turned out to be an important concept for applied research. It captures how the operator keeps track of, comprehends, and anticipates events in a dynamic environment. However, SA has proven to be a difficult concept to grasp both theoretically and empirically. Theoretically, there is much confusion in terms of what SA is as a formal psychological concept. Is it a phenomenon or the underlying representation, a process or a product? Researchers have also struggled with the meaning of awareness, leaning towards the operator's conscious report as equivalent to SA. This confusion has been carried over to empirical research. Hence, progress in SA research has been limited. The present work offers some clarifications on what SA is as a concept and how it can be further understood by a cognitive model of WM.

Current theoretical frameworks can broadly be distinguished into two approaches; *what* SA is as a product and *how* SA is acquired and maintained by the underlying cognitive mechanism. On the product view, the focus is on understanding SA as the operator's conscious comprehension. The alternative approach is to understand how the operator acquires and maintains knowledge in a dynamic, multitask environment. Ultimately, a complete model of SA must outline the functions of the underlying mechanism. Such a model will allow researchers to make predictions regarding why and how SA is maintained or breaks down. The product view provides no

information on how knowledge maps onto performance. There is also no specification of the cognitive processes required to maintain SA. Furthermore, treating SA as explicit conscious knowledge provides a very limited picture of the operator-machine interaction. The approach taken by Adams et al. (1995) puts the focus on outlining the functions of WM in order to understand how SA is maintained. On this approach it is not informative or practical to distinguish SA as a product from the underlying processes. Rather SA must be understood in terms of processing and representation of information in WM. Adams et al. show how a WM model is capable of explaining many critical SA issues such as task prioritization and control. However, their framework is still vague. It does for example, not provide a sufficient specification of the relationship between WM and SA. Therefore, their framework generates no testable hypothesis.

The goal of the present work was to use an established WM model to gain better understanding of the role of WM in maintaining SA. The multiple component WM model of Baddeley (1986, 2000) and Logie (1995) was chosen for two reasons; the model is well established for generating hypotheses about WM. The existing evidence suggests that SA might be supported by processing and maintaining information in the different subsystems proposed by the model. The model generated certain hypotheses that were then tested. The central argument of this dissertation, that WM subsystems play a complex role in maintaining SA was supported.

Before exploring the role of WM in maintaining SA, it was necessary to explore some of the foundational issues in the field. In chapter 3, the assumption that SA is necessarily what the operator is conscious of knowing and can report was questioned.

The discussion did not wander too far into the area of philosophy in offering a prominent solution for defining awareness. But it did bring to light some important evidence showing the limitation and perhaps the misconception of equating awareness with explicit, reportable knowledge. The issue had to be dealt with, both for methodological and theoretical development. Referring to SA as what the operator is aware of knowing and can report limits how SA is measured. Furthermore, how SA is conceptualized affects how the relationship between WM and SA is portrayed. Based on the evidence and some philosophical arguments it was proposed that rather than defining SA in terms of awareness, the concept should be determined in terms of the cognitive operations needed to maintain awareness. These are perception, comprehension and anticipation. The assumption is that if the operator perceives, comprehends and anticipates events, the operator is aware, not vice versa.

6.2 Implications for SA theory

Although, researchers have identified WM as the “place” where SA happens (Adams et al., 1995; Endsley, 1995a; Wickens, 1995) it was still an empirical question whether SA and WM are linked and what the nature of the relationship between those two concepts is like. As reviewed in Chapter 2 and Chapter 4, the evidence suggests a plausible role of WM in maintaining SA. But previous studies have lacked a proper manipulation for testing WM or alternatively, do not measure SA. The present research therefore expanded on the existing studies in an important way. The results suggest that SA requires active, online, processing and maintenance of task-relevant information in

the different WM systems. This supports a much more active role of the WM mechanism in maintaining SA than what has been proposed in the field.

Endsley (1995a) emphasizes LTS as the key mechanism supporting SA. On this view, WM serves to activate stored mental models. The processing limitations of the WM mechanism therefore can largely be bypassed. This view of SA and WM is in stark contrast to basic research on WM and the present work. Endsley's view is also limited in that it does not explain how the situation model is constructed and maintained in WM or how information is processed at the different stages of perception, comprehension and anticipation.

Adams et al. (1995) proposed a more active role of WM in maintaining SA. But they also emphasized connection to LTS. Their model is in line with the LT-WM model by Eriksson and Kintsch (1995). Durso and Gronlund (1999) have also argued that some sort of a LT-WM model would probably provide a good future model for SA theory and research. Obviously, the connection to LTS is important to SA. This is particularly true because SA is an issue in operational environment where experience and prior knowledge must be a factor in explaining SA. Until recently, the LT-WM model was the only model accounting for a relation between LTS and WM. However, with the addition of the episodic buffer, Baddeley's model also accounts for how WM and LTS are related (Baddeley, 2001, 2002). Of course, much more work is needed to further determine the relation among WM, LTS, and SA. The important point here however, that often gets overlooked, is that SA occurs in a dynamic environment where the operator is faced with an array of constantly changing information that needs to be integrated and updated

regardless of experience and prior knowledge. The driver for example, driving in heavy traffic, must constantly update the representation of the driving scene in WM. The WM mechanism therefore plays a much more central role in explaining SA than what is often proposed in SA research. Prior knowledge obviously will facilitate how information is processed and integrated but it does not mean that WM plays less of a role in maintaining SA. In fact, Perez et al. (2000), using fMRI found that when expert pilots maintained SA in tracking a projected flight path, they showed more processing in areas associated with WM than did novice pilots. Experts' also showed processing in more defined areas than the novices.

The present research has important implications for SA theory and research. For example, the distinction made in the field between SA as a product and the underlying processes seems meaningless when the complex role of WM in maintaining SA is considered. Even if researchers agree that SA is supported by a coordinated representation (product) in the episodic buffer, that representation is constantly being updated by processing and maintaining information in the WM subsystems. The information from the subsystems is then integrated and coordinated with prior knowledge. Therefore, a complete understanding of SA must necessarily involve both the product and the processes. In particular, portraying the functions of the underlying mechanism of WM must involve how information is processed in the different WM components. The tendency in the field has been to either refer to SA as a product or to replace SA with the concept of situation assessment which refers to the underlying processes (Fredrico, 1995). The present work however, is more in favour of the

alternative approach where SA is both the product and the processes (Adams et al., 1995; Dominguez, 1994; Tenney, et al., 1992). This approach seems to more accurately capture cognition in a dynamic environment.

The present work supports some of the claims made by Adams et al. regarding the role of WM in maintaining SA. For example, Adams et al. argued that an important part of understanding SA must involve understanding how an operator prioritizes tasks, knows what task to perform and when, and selectively attends to information. Adams et al. proposed that WM carries pointers to a larger task scheduling structure in LTS that guides the operators in switching between tasks and monitoring their own activity. Such task scheduling program would allow the operator to go back to performing a task if it is interrupted. Going back and forth between tasks might otherwise be hard to do as each task has sub steps, and therefore a lot of task-related information needs to be kept in active memory. The evidence in Experiment 4 supported the idea of a larger task scheduling program activated by pointers maintained in WM. In particular, it was found that in carrying out certain tasks such as monitoring traffic behind the car, participants seemed to use the verbal mechanism as a self-cuing device for activating an action control program in LTS. Furthermore, as speculated in Chapter 5, this mechanism might explain how operators follow and operate on task goals.

Adams et al. also argued that WM would function to control and select actions, adapt hypotheses and evaluate consequences. The present experiment found evidence supporting an active role of the central executive in coordinating tasks and facilitating task switching. Maintaining and processing information in the two subsystems also

guided performance on the different SA tasks. The present work therefore expands further on some of the claims made by Adams et al. explaining the selective role of the WM components in maintaining SA in a complex task environment.

The present work however, goes much further than Adams et al. and is guided by hypothesis generated by linking a theoretical model of WM to SA. Making the conceptual link between WM and SA and testing it, provides the foundation for defining the relation among SA, WM, and the underlying representation. By doing so the present work sheds some light on another foundational issue in the field. Namely, the confusion over whether SA is a phenomenon, the underlying representation or behavioural description (Durso & Gronlund, 1999). Endsley (1995a) for example, does not distinguish between SA and the situation model as the underlying representation. On her view, SA is the situation model formed in WM. This however, limits the nature of the phenomenon under study as SA may only reflect on an aspect of the representation that underlies the complex operator-machine interaction (Johannsdottir, LeFevre & Herdman, 2003). Furthermore, equating SA and the underlying representation leads to circular reasoning where SA is what is measured and also what causes the measured behaviour (Flach, 1995). In general, researches in the field do not take care in identifying if SA is the representation in WM or what reflects on that representation.

In Experiment 4, SA was measured as knowledge of car location. But participants' knowledge of the surrounding cars was determined by the representation and processing of information in WM. It is suggested that the coordinated representation in the episodic buffer is what provides the operator with a coherent awareness of the task

environment. This coordinated representation is constantly updated in the different WM components. The emphasis in SA research should therefore be on explaining how the coordinated representation is constructed and maintained in WM. The representation essentially is what causes SA. SA on the other hand, is what can be measured as a reflection of the representation in WM.

It is therefore suggested here that at the centre of explaining how the operator comprehends dynamic events is the WM mechanism and the coordinated representation. SA is what is being measured as a reflection of the represented information in WM. This marks an important shift in SA research and theory. As mentioned earlier, one of the confusion in the field is that SA is used as both what is being measured (the empirical concept) and what causes the observed behaviour (the theoretical construct). Furthermore, the emphasis has been more on correlating measured SA with system design rather than trying to understand how SA is maintained. The present work suggests that the representation in WM is the theoretical concept and SA is the empirical concepts. Also, maintaining the representation in WM requires considerable resources. Therefore, measuring workload in addition to measure SA may provide a more complete picture of the underlying representation. Adams et al. had already argued that the issue of resources cannot be clearly separated from the study of SA as maintaining SA by the underlying mechanism of WM in a complex task environment must involve considerable cognitive resources.

6.3 The CSE framework

The shift in the explanatory approach that is being suggested here has already been proposed in the Cognitive Systems Engineering (CSE) framework that has been developed within the ACE lab at Carleton University (Johannsdottir, LeFevre, & Herdman, 2003). The CSE framework was not developed to substitute existing theories or as a complete model of SA. Rather, it was developed to meet a demand for guidelines for the process of evaluation and testing of new cockpit technology. The CSE framework therefore responds to many of the issues raised in this dissertation. In particular, Johannsdottir et al. argued that the underlying representation must be distinguished from SA. Furthermore, SA is not sufficient to capture the complexity of the underlying representation. Therefore, in the CSE framework (see Figure 19) the theoretical concept at the centre is the dynamic mental model (DMM) which captures the constant representation and updating of the system and the task environment in WM. The framework also includes three empirical or measured constructs that reflect on the DMM, namely SA, workload, and task-relevant performance.

Figure 19. CSE Framework

The central argument here is that using converging measures of the three central concepts; SA, workload and task-relevant performance is more adequate for assessing the complex operator-machine interaction. Furthermore, the advantage of distinguishing between the DMM and the empirical concepts is that it allows researchers to more clearly operationalize the concepts for empirical purposes. The present work further supports that an adequate model of SA may place the DMM and WM at the centre, measuring SA, workload and performance as a reflection of that representation.

6.3.1 The DMM

The present research provides evidence for the construction of the DMM in different components of WM. Therefore, WM provides a cognitive model to study the DMM further. There is already evidence supporting the notion of a mental model when

carrying out many complex cognitive tasks such as problem solving, reasoning and text comprehension (Johnson-Laird, 1983; Van Dijk & Kintsch, 1983). The concept has also been used to explain cognition in a dynamic environment like playing football (Johnson-Laird, 1983). Mental models refer to a coherent representation that captures causal, spatial and temporal structures of the task environment. A similar concept of a situation model has well-established role in explaining how people comprehend texts (Van Dijk & Kintsch, 1983; Zwaan & Radvansky, 1998).

The main difference between a situation model and a mental model is that the situation model specifically refers to the current situation where as a mental model can refer to a more general knowledge based in LTS. The concept of a situation model is well suited as the underlying structure of SA. It accurately represents the structure of the task environment and is thought to combine incoming online information with stored information about similar situations or about the objects and persons described (Zwaan & Radvansky, 1998). There is evidence to suggest that the situation model is constructed separately in spatial and verbal WM systems (Friedman & Miyake, 2000). Information regarding spatial layout and the relationship between objects is processed in the spatial system whereas causal information is processed in the verbal system. Furthermore, Friedman and Miyake (2000) speculated that information processed in the two subsystems would be brought together at some point for LTS. Integrated representation in LTS would facilitate retrieval and also provide the reader with a coherent integrated understanding of the text.

The DMM bears many similarities with a situation model; it refers to the construction of situation-specific information in a dynamic environment. Furthermore, the present research suggests that similar to a situation model the DMM is constructed in the spatial and the verbal subsystems separately. This is in accord with basic research on the role of WM in information-rich environments. For example, Garden et al. (2002) found that the ability to navigate in a real city was interfered with by both spatial and verbal load tasks. In Experiment 3, participants relied on the verbal system for predicting where the target object was going to reappear and on the spatial system for searching for the target object. Similarly, in Experiment 4, participants relied on the verbal system for monitoring cars behind them and the spatial system for monitoring cars in front. But the present research also suggested that maintaining SA depends on integrated representation possibly in the episodic buffer. The role of the central executive in all aspects of the task in experiment 4 suggests that participants relied on coordinated representation of the driving scene. This would allow them to use information about other cars such as whether they are speeding up or slowing down or changing lanes, to control the car. Furthermore, knowing that a car positioned behind the driver is about to overtake the car affects the driver's SA of cars in front. There is evidence to suggest that information from the verbal and the spatial subsystem are integrated in a coordinated representation to support a performance on a complex task. Barshi and Healy (2002) for example, found that participants seemed to rely on an integrated spatial and verbal representation to carry out a navigation task. Such coordinated representation may be updated by processing

information in the two subsystems. The evidence therefore supports the role of different WM components in constructing the DMM.

Zwaan and Radvansky (1998) suggested that in text comprehension the situation model can be further divided into the current model, the integrated model and the complete model. The current model captures the current construction (coordination) of the information at t_n . That is, the current model captures the construction of the model for any particular clause or sentence. The integrated model integrates the various current models into a more coherent understanding of the text at times t_1 through t_{n-1} while the person reads through clauses c_1 to c_{n-1} . Finally, the complete model refers to the model that is stored in LTS after all the text has been comprehended and integrated. It is important to note that the model that gets stored in LTS is not necessarily the same as the models constructed throughout the reading process. People may add information and make additional inferences following the reading.

The DMM can also be thought of in terms of a current, integrated and complete model. At each moment while driving for example, the driver constructs a dynamic model of what is going on. The models at each moment in time are then integrated through the process of updating into a more coherent understanding of the situation. This allows the driver to have a more complete understanding of the overall situation and how things are evolving. The integrated model would therefore support prediction (level 3 SA) and comprehension (level 2 SA). In contrast, the current model would support level 1 SA –perception of the relevant elements that have not yet been integrated. It is important to note that the final model is not necessarily an accurate reflection of the

current or the integrated model. Therefore, again researchers should take care in relying on the operators' report after the task has been completed.

More work is needed to further understand the DMM and how it contributes to SA. For example, further research might focus on understanding how information is processed in the subsystems and what role the integrated representation plays in supporting SA in dynamic environments. It is important to note that the DMM refers to complex cognitive processing of information including selective attention, LTS and perceptuomotor coordination and control (Johannsdottir et al., 2003). Therefore, it is important to use SA along with the concept of workload and task-relevant performance to assess the DMM.

6.3.2 The measured constructs

Johannsdottir et al. argued that the three empirical concepts, SA, workload and task-relevant performance provide an inclusive assessment of the DMM. SA captures how the operator keeps track of, integrates and anticipates events in a dynamic environment. It is expected that in a complex operational environment, operators' performance is largely facilitated by SA. This is because in environments such as the automobile or the cockpit, most operations require integration of information and anticipation of future events. However, certain task-relevant performance will not directly reflect on SA and therefore needs to be understood along with SA measures. For example, measuring head and eye tracking provides important information on how new technology may alter the way in which the operator interacts with the system and

monitors the task environment. Furthermore, if a pilot or a driver spends a lot of time looking at a new system at the cost of monitoring the outside environment, his or her overall SA will be decreased. However, head tracking does not directly measure how the operator comprehends the task or the operator's ability to predict future occurrences. It is therefore important to measure both; the operator's SA and task-relevant performance that is not captured by the SA concept.

For example, Recarte and Nunes (2000) found that performing mental tasks while driving affected visual search, eye scanning and eye fixation. Knowing this however, did not provide them with understanding of how the mental tasks would affect SA and decision-making. Certain task performance obviously will reflect on comprehension and anticipation. For example, the highway in the sky (HITS) display was developed to project the flight path for the pilot and therefore enhance level 3 SA. How well the pilot performs in maintaining the path therefore is a direct reflection of his or her SA. But in other cases performance may not reflect how the information is understood. That is, good SA may not lead to good performance and good performance can occur even when the operator does not fully comprehend the situation. Therefore, there is great value in using multiple measures and concepts both for evaluating new technology and also for theoretical purposes. It provides more complete understanding of the phenomenon under study.

Another area, which is not covered by measuring SA, or task-related performance, is the issue of cognitive resources. In applied research SA has pushed aside another important concept; workload, which captures the cognitive resources needed to perform a

task. People have limited resources to process and integrate information and respond appropriately. It is therefore important to understand how a new technology affects the workload in addition to SA and performance. Adams et al. (1995) pointed out that maintaining SA or as proposed here, maintaining the DMM, requires cognitive resources and therefore understanding the DMM cannot be separated from the issue of workload. The present research showed that more demanding tasks such as prediction and search caused greater interference with load tasks than the tracking task. Furthermore, the ability to predict where the tracked object was going to be was affected by cognitive load. In general, updating and maintaining the DMM in particular in a multitask, dynamic environment is cognitively demanding. Therefore, implementing a new system into the car may enhance understanding but increase workload. If workload is too high, SA in turn will be diminished. For example, Gugerty (1997) found that increased number of surrounding cars mean that participants focused on cars close to their own car. Increased workload therefore forced them to focus their SA only on what was relevant or a potential danger. But it could also decrease the participants SA in that, cars further away may quickly become a danger depending on their speed and distance.

6.4 Conclusion

The CSE framework provides general guidelines for testing and evaluating the impact of a new system on the human operator. It also provides some important conceptual foundations for further theoretical development. In particular, the framework provides the foundation for defining the relationship among WM, DMM, and SA. It also

takes into consideration the complexity of the DMM. Therefore, it is suggested that researchers should measure SA, workload and task-relevant performance rather than testing only SA in order to get a more complete picture of the underlying representation. Johannsdottir et al. have also suggested that the three empirical constructs must be operationalized properly by including multiple measures both subjective and objective. Converging measures are more likely to give a complete evaluation of the complex operator-machine interaction (Gardner, Hake & Ericksen, 1956).

WM provides a cognitive model to further explore the relation between the theoretical and empirical concepts included in the CSE framework. In particular, the WM model allows researchers to further test the relationship between the WM mechanism and each of the empirical concepts of SA, workload and performance. The WM model also allows researchers to explore how the three empirical concepts relate to each other. A complete disconnection between SA, workload and task-relevant performance would be evidence against the CSE framework, as those concepts are a reflection of the construction and the maintenance of the DMM.

Maintaining the DMM requires considerable cognitive resources. In particular, integrating information in complex, information-rich environments in a way that allows the operator to predict future occurrences must be resource demanding. Therefore, a certain amount of workload may motivate the operator and as such enhance SA. However, high workload in many cases may decrease the operator's ability to integrate and manipulate the information in order to fully comprehend and anticipate events. The issue of workload is important and not easily separated from the issue of SA. Measuring

workload provides important information on how much cognitive work is required in using a new display in addition to whether or not it enhances the operators' SA. In general, using the three concepts of SA, workload and performance provides a comprehensive account of the DMM.

High workload and/or low SA may not always predict poor performance. In fact, the operator may perform better in high compared to low workload conditions. Similarly, performance can be poor even in situations where the operator comprehends the task situation. But the relationship between SA and performance and workload and performance needs to be further specified. Understanding the relation among SA, workload, and performance would provide researchers with further understanding of the boundaries of SA (i.e. what is and is not a case of SA). The status of performance measure is particularly confusing in SA research. On the one hand, performance measures are not accepted as a "good" measure of SA but on the other hand, performance measures are the most frequently used measures in the field (Vidulich, 2000). But it is not always clear when performance is reflecting SA and when it is not. The present work has attempted to provide some boundaries and definition for the meaning of SA. But future research needs to work out the relationship between SA and task-relevant performance.

6.5 Final words

The present work provides a link between the applied concept of SA and a basic cognitive mechanism WM. Linking SA and WM contributes independently to both basic and applied research. Although, WM is a well researched topic in basic research the

mechanism has not been studied in applied environment. Partly, the reason for why WM has not been studied in applied context is a lack of a conceptualization of cognitive processing in dynamic environments. Connecting WM to SA therefore expands our understanding of the WM mechanism. Linking WM and SA is also important for applied theory and research. The WM model provides guidance for further empirical work and may prove to be essential to understanding SA.

There is substantial evidence both in applied and basic research that supports the role of WM in maintaining SA. However, the relationship between WM and SA has not been well articulated. The problem is largely due to how SA has been conceptualized as a conscious product that is separate from the underlying processes. On the product view, researchers have focused on the content of SA, or what the operator knows and can report, rather than how knowledge in a dynamic environment is acquired and maintained. On an alternative approach Adams et al. (1995) suggested that SA must be understood by outlining the functions of WM. However, their framework is too vague to generate any testable hypothesis or to guide empirical work. The present work takes the alternative approach further by establishing theoretically and empirically the role of WM in supporting SA in dynamic environments. This is the only existing work in the field to my knowledge that actually tests the relationship between an established cognitive process and the applied concept of SA. The present research therefore is a significant contribution to further theoretical and empirical progress in the field.

The relation between WM and SA that has been established here empirically provides an important insight into SA and how it is maintained by the underlying

mechanism. It also starts to define what SA is as a formal psychological concept and where it fits within the more general explanatory framework. There is confusion in SA research and theory, where the concept is sometimes the phenomenon and sometimes the underlying representation. Therefore, SA as measured behaviour is also often assumed to be the explanation of breakdown in behaviour. However, as demonstrated in the present work, SA on its own does not explain why and how an operator understands a system or how it will affect the operator's performance. Rather, SA guided by the WM model and measured along with performance and workload provides an adequate approach to understand the operator-machine interaction in a dynamic environment. Further work may look at the relationship between SA and other cognitive concepts such as attention. Also, more research is needed to understand the relationship between WM, LTS and expert performance. Such research can be guided by the CSE framework and the multi component model of WM.

This dissertation is an example of how important it is to a) bring together fundamental cognitive processes and applied concepts and b) do interdisciplinary work. The WM mechanism provides an important insight into the nature of SA and how and why SA brakes down. Furthermore, by integrating evidence across different disciplines the present work sheds some light on the meaning of awareness in situation awareness. SA is assumed to necessarily mean what the operator is aware of knowing and able to report. This meaning however was questioned in Chapter 3. Unfortunately researchers have been silent on the issue but defining SA in terms of what the operator can report has affected both methodology and theory. However, integrating evidence from cognitive

psychology and neuroscience with some arguments from philosophy provided a strong case against defining SA in terms of explicit conscious knowledge the operator is able to articulate.

In general, the present work has important implications for further theoretical and empirical progress in the SA literature. The WM model provides the needed foundation for making predictions and guiding empirical work, allowing researchers to systematically study SA and build up a theoretical model that explains the concept.

References

Adams, M. J., & Pew, R. W. (1990). Situational awareness in the commercial aircraft cockpit: A cognitive perspective. In *Proceedings of the IEEE/AIAA/NASA 9th Digital Avionics Systems Conference* (pp. 519-524). New York: Institute of Electrical and Electronics Engineering.

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

Allport, A. (1988). What concept of consciousness? In A. J. Marcel & E. Bisiach (Eds.), *Consciousness in Contemporary Science*. Oxford: Clarendon Press.

Andre, A. D., & Wickens, C. D. (1991). Display formatting techniques for improving situation awareness in the aircraft cockpit. *The International journal of Aviation Psychology*, 1, 205-218.

Andre, A. D., Wickens, C. D., Moorman, L., & Boschelli, M. M. (1991). Display formatting techniques for improving situation awareness in the aircraft cockpit. *The International Journal of Aviation Psychology*, 1, 205-218.

Aretz, A. J. (1991). The design of electronic map display. *Human Factors*, 33, 85-101.

Arthur, W., Barrett, G. V., & Doverspike, D. (1990). Validation of an information-processing-based test battery for the prediction of handling accidents among petroleum-product transport drivers. *Journal of Applied Psychology*, 75, 621-628.

Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. In K. W. Spence, Ed., *The Psychology of Learning and Motivation: Advances in Research and Theory*. New York: Academic Press.

Baddeley, A. D. (1986). *Working Memory*. Oxford: Clarendon Press.

Baddeley, A. D. (1996). Exploring the central executive. *The Quarterly Journal of Experimental Psychology*, 49A, 5-28.

Baddeley, A. D. (2000). The episodic buffer: a new component of working memory? *Trends in Cognitive Science*, 4, 417-423.

Baddeley, A. D. (2002). Is working memory still working? *European Psychologist*, 7, 85-97.

Baddeley, A. D. (2003). Working memory and language: An overview. *Journal of Communication Disorder*, 36, 189-208.

Baddeley, A. D., & Andrade, J. (2000). Working memory and the vividness of imagery. *Journal of Experimental Psychology: General*, 129, 126-145.

Baddeley, A. D., Bressi, S., Della Sala, S., Logie, R. H., & Spinnler, H. (1991). The decline of working memory in Alzheimer's disease: A longitudinal study. *Brain*, 114, 2521-2542.

Baddeley, A. D., Chincotta, D., & Adlam, A. (2001). Working memory and the control of action: Evidence from task switching. *Journal of Experimental Psychology: General*, 130, 641-657.

Baddeley, A.D., Gathercole, S. E., & Papagno, C. (1998). The phonological loop as a language device. *Psychological Review*, 105, 158-173.

- Baddeley, A. D., Emslie, H., Kolodny, J., & Duncan, J. (1998). Random generation and the executive control of working memory. *The Quarterly Journal of Experimental Psychology*, 51A, 819-852.
- Baddeley, A. D., & Hitch, G. J. (1974). Working memory. In G. A. Bower (Ed.), *Recent Advances in Learning and Motivation* (Vol. 8, pp. 47-90). New York: Academic Press.
- Baddeley, A. D., Lewis, V. J., & Vallar, G. (1984). Exploring the articulatory loop. *Quarterly, Journal of Experimental Psychology: Human Experimental Psychology*, 36(A), 233-252.
- Baddeley, A. D., & Logie, R. H. (1999). Working memory: The multiple-component model. In A. Miyake & P. Shah (Eds.), *Models of Working Memory* (pp. 28-61). New York: Cambridge University Press.
- Baddeley, A. D., & Wilson, B. A. (2002). Prose recall and amnesia: Implications for the structure of working memory. *Neuropsychologia*, 40, 1737-1743.
- Banbury, S., Andre, A., and Croft, D. G. (in press). Do we "know" all that we "know"? The role of implicit knowledge in Situation Awareness. In *Proceedings of the 47th Annual Human Factors and Ergonomics Society Conference*, San Diego. USA.
- Barshi, I., & Healy, A. F. (2002). The effects of mental representation on performance in a navigation task. *Memory & Cognition*, 30, 1189-1203.
- Bass, E. J., Zenyuh, J. P., Small, R. L., & Fortin, S. T. (1997). A context-based approach to training situation awareness. In *Proceedings of HICS '96 -Third Annual*

Symposium on Human Interaction with Complex Systems. Los Alamitos, CA: IEEE Computer Society, (pp. 89-95).

Baxter, G. D., & Bass, E. J. (1998). Human error revisited: Some lessons for situation awareness. In *Proceedings of the Fourth Annual Symposium on Human Interaction with Complex systems*: Los Alamitos, CA: IEEE Computer Society Press, (pp. 81-87).

Billings, C. E. (1995). Situation awareness measurement and analysis: A commentary. In D. J. Garland and M. R. Endsley (Eds.), *Experimental Analysis and Measurement of Situation Awareness* (pp 295-304). Daytona Beach: Embry-Riddle Aeronautical University Press.

Block, N. (1997). On a confusion about a function of consciousness. In N. Block, O. Flanagan, & G. Guzeldere (Eds.), *Consciousness*. Cambridge, MA: MIT Press.

Bolstad, C. A. (1991). Individual pilot differences related to situation awareness. . In *Proceedings of the Human Factors Society 35th Annual Meeting* (Vol. 1, pp. 52-56). Santa Monica, CA: Human Factors Society.

Brickman, B. J., Hettinger, L. J., Roe, M. M., Stautberg, D., Vidulich, M. A., Hass, 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 and M. R. Endsley (Eds.), *Experimental Analysis and Measurement of Situation Awareness* (pp 295-304). Daytona Beach: Embry-Riddle Aeronautical University Press.

Bridgeman. B. (1999). Vertical modularity in the visual system. In H. Bradford & B. Velichkovsky (Eds.), *Stratification in Cognition and Consciousness. Advances in consciousness research*. Amsterdam, Netherlands: John Benjamins Publishing Company.

- Brook, A., & Raymont, P. (in press). *A Unified Theory of Consciousness*. Cambridge, MA: MIT Press.
- Carpenter, P. A., Just, M. A., & Shell, P. (1990). What one intelligence test measure: A theoretical account of the processing in the Raven Progressive Matrices Test. *Psychological Review*, 97, 404-431.
- Carretta, T. R., Perry, D. C., & Ree, M. J. (1996). Prediction of situational awareness in F-15 pilots. *The International Journal of Aviation Psychology*, 6, 21-41.
- Caretta, T. R., Perry, D. C., & Ree, M. J. (1994). The Ubiquitous Three in the Prediction of Situational Awareness: Round Up the Usual Suspect. In R. D. Gilson, D. J. Garland, & J. M. Koonce (Eds.). *Situational Awareness in Complex Systems*. Embry-Riddle Aeronautical University Press.
- Carpenter, P. A., Just, M. A., & Shell, P. (1990). What one intelligence test measures: A theoretical account of the processing in the raven progressive matrices test. *Psychological Review*, 97, 404-431.
- Charness, N. (1995). Expert performance and situation awareness. In D. J. Garland and M. R. Endsley (Eds.), *Experimental Analysis and Measurement of Situation Awareness* (pp 295-304). Daytona Beach: Embry-Riddle Aeronautical University Press.
- Chun, M. M., & Jiang, Y. (2003). Implicit, long-term spatial contextual memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 29, 224-234.
- Clark, A. (1997). *Being There: Putting Brain, Body and World Together Again*. Cambridge, MA: MIT Press.

Cocchini, G., Logie, R. H., Sala, S. D., & MacPherson, S. E. (2002). Concurrent performance of two memory tasks: Evidence for domain-specific working memory systems.

Damos, D. L. (1993). Using meta-analysis to compare the predictive validity of single- and multiple-task measures to flight performance. *Human Factors*, 35, 615-628.

Dehaene, S., & Naccache, L. (2001). Towards a cognitive neuroscience of consciousness: basic evidence and a workspace framework. *Cognition*, 79, 1-37.

Dehaene, S., Naccache, L., Le Clec'H, G., Koechlin, E., Mueller, M., Dehaene-Lambertz, G., van de Moortele, P.-F., & Le Bihan, D. (1998). Imaging unconscious semantic priming. *Nature*, 395, 597-600.

Denis, M., & Cocude, M. (1989). Scanning visual images generated from verbal descriptions. *European Journal of Cognitive Psychology*, 1, 293-307.

Dienes, Z., & Perner, J. (1999). A theory of implicit and explicit knowledge. *Behavioral and Brain Sciences*, 22, 735-808.

Dominguez, C (1994). Can SA be defined? In M. Vidulich, C Dominguez, E Vogl & G. McMillan (Eds.), *Situation Awareness: Papers and Annotated Bibliography* (pp. 5-15). AL/CF/TR-1994-0085, Armstrong Laboratory.

Dreyfus, H. L. (1997). Intuitive, deliberate, and calculative models of expert performance. In C. Zsombok & G. Klein (Eds.), *Naturalistic Decision Making*. Hillsdale, NJ, England: Lawrence Erlbaum Associates.

Dretske, F. (1995). *Naturalizing the Mind*. Cambridge MA: MIT Press.

- Durso, F. T. & Gronlund, S. D. (1999) Situation Awareness. In T. F. Durso, R. S. Nickerson, R. W. Schvaneveldt, S. T. Dumais, D. S. Lindsay and M. T. H. Chi (Eds.). *Handbook of Applied Cognition*, (pp. 283-313).
- 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.). *Experimental Analysis and Measurement of Situation Awareness*. Dayton Beach: Embry-Riddle Aeronautical University Press.
- Durso, F. T., Truitt, T. R., Hackworth, C. A., Crutchfield, J. M., & Manning, C. A. (1998). En route operational errors and situation awareness. *The International Journal of Aviation*, 8, 177-194.
- Eidelkind, M., Moffett, R., Ardendt, D., & McKee, C. (1994). Give Me Situation Awareness or Give Me Death. In R. D. Gilson, D. J. Garland, & J. M. Koonce (Eds.). *Situational Awareness in Complex Systems*. Embry-Riddle Aeronautical University Press.
- Emerson, M. J., & Miyake, A. (2003). The role of inner speech in task switching: A dual-task investigation. *Journal of Memory and Language*, 48, 148-168.
- Emerson, M. J., & Miyake, A. (2003). The role of inner speech in task switching: A dual-task investigation. *Journal of Memory and Language*, 48, 148-168.
- Endsley, M. R. (1988). Design and evaluation for situation awareness enhancement. In *Proceedings of the Human Factors Society 32nd Annual Meeting* (pp. 97-101). Santa Monica, CA: Human Factors Society.

- 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 and Ergonomics Society.
- Endsley, M. R. (1995a). Toward a theory of situation awareness in dynamic systems. *Human Factors*, 37, 65-84.
- Endsley, M. R. (1995b). Measurement of situation awareness in dynamic systems. *Human Factors*, 37, 65-84.
- Endsley, M. R. (1995c). Theoretical underpinnings of situation awareness: A critical review. In R. D. Gilson, D. J. Garland, & J. M. Koonce (Eds.). *Situational Awareness in Complex Systems*. Embry-Riddle Aeronautical University Press.
- Endsley, M. R. (1997). Supporting situation awareness in aviation systems. In *Proceedings of the IEEE International Conference on Systems, Man and Cybernetics*, (pp. 4177-4181). Piscataway, NJ: Institute of Electrical and Electronics Engineers.
- Endsley, M. R. (1998). A comparative analysis of SAGAT and SART for evaluations of situation awareness. In *Proceedings of Human Factors and Ergonomics Society 42nd Annual Meeting* (pp. 82-86). Santa Monica, CA: The Human Factors and Ergonomics Society.
- Endsley, M. R. (2000). Theoretical underpinnings of situation awareness: A critical review. In M. R. Endsley & D. J. Garland, (Eds.), *Situation Awareness Analysis and Measurement*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Endsley, M. R., & Bolstad, C. A. (1994). Individual differences in pilot situation awareness. *The International Journal of Aviation Psychology*, 4, 241-264.

Entin, B. E., Serfaty, D., & Entin, E. E. (1995). Modeling and measuring situation awareness for target-identification performance. In D. J. Garland and M. R. Endsley (Eds.), *Experimental Analysis and Measurement of Situation Awareness* (pp 295-304). Daytona Beach: Embry-Riddle Aeronautical University Press.

Ericsson, K. A., & Kintsh, W. (1995). Long-term working memory. *Psychological Review*, 102, 211-245.

Ericsson, K., & Simon, H. (1984). *Protocol analysis – verbal report as data*. Cambridge, MA: MIT Press.

Federico, P. (1995). Expert and novice recognition of similar situations. *Human Factors*, 37, 105-122.

Fracker, M. L. (1988). A theory of situation assessment: Implications for measuring situation awareness. In *Proceedings of the Human Factors Society 32nd Annual Meeting* (Vol. 1, pp. 102-106). Santa Monica, CA: Human Factors Society.

Fracker, M. L. (1989). Attention allocation in situation awareness. In *Proceedings of the Human Factors Society 33rd Annual Meeting*, (Vol. 2, pp. 1396-1400). Santa Monica, CA: Human Factors Society.

Fracker, M. L. (1991). Measures of situation awareness: Review and future directions. (AL-TR-1991-0128). Wright-Patterson AFB OH: Armstrong Laboratory, Crew Systems Directorate.

Fracker, M. L., & Davis, S. A. (1990). Measuring operator situation awareness and mental workload. In *Proceedings of the Fifth Mid-central Ergonomics/Human Factors Conference*. Dayton, OH: University of Dayton.

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

Friedman, N. P., & Miyake, A. (2000). Differential roles for visuospatial and verbal working memory in situation model construction. *Journal of Experimental Psychology: General*, 129, 61-83.

Gaba, D. M., & Howard, S. (1995). Situation awareness in anesthesiology. *Human Factors*, 37, 20-31.

Garden, S., Cornoldi, C., & Logie, R. H. (2002). Visuo-spatial working memory in navigation. *Applied cognitive psychology*, 16, 35-50.

Garland, D. J., Stein, E. S., Wise, J. A., & Blanchard, J. W. (1993). Situational awareness in air traffic control: A critical yet neglected phenomenon. In D. J. Garland & J. A. Wise (Eds.), *Human Factors and Advanced Aviation Technology*. Embury-Riddel Aeronautical University.

Garner, W. R., Hake, H. W., & Eriksen, C. W. (1965). Operationism and the concept of perception. *Psychological Review*, 63, 149-158.

Goettl, B. P. (1997). Situation awareness and executive control processes: Quot Homines, Tot Sententiae. In *Proceedings of the Human Factors and Ergonomics Society 41st Annual Meeting*, (Vol x, pp. x-x). Santa Monica: Human Factors and Ergonomics Society.

Goldman-Rakic, P. S. (1992). Circuitry of primate prefrontal cortex and regulation of behavior by representational memory. In J. M. Brookhart & V. B. Mountcastle (Eds.),

Handbooks of Physiology-The Nervous System V. Baltimore, MD: American Physiological Society.

Goodale, M. A., & Milner, A. D. (1992). Separate visual pathways for perception and action. *Trends in Neuroscience*, 15, 20-25.

Gopher, D. (1982). A selective attention test as a predictor of success in flight training. *Human Factors*, 24, 173-183.

Gopher, D., Weil, M., & Barket, T. (1994). Transfer of skill from a computer game trainer to flight. *Human Factors*, 36, 387-405.

Green, C. D. (1992). Of immortal mythological beasts: Operationism in psychology. *Theory & Psychology*, 2, 291-320.

Gugerty, L. C. (1997). Situation awareness during driving: Explicit and implicit knowledge in dynamic spatial memory. *Journal of Experimental Psychology: Applied*, 3, 1-26.

Gugerty, L. J. & Tirre, W. C. (1995a). Comparing explicit and implicit measures of situation awareness. In D. J. Garland & M. R. Endsley, (Eds.), *Experimental Analysis and Measurements in Situation Awareness*. Daytona Beach: Embry-Riddle Aeronautical University Press.

Gugerty, L. J. & Tirre, W. C. (1995b). Cognitive correlates of explicit and implicit measures of situation awareness. In D. J. Garland & M. R. Endsley, (Eds.), *Experimental Analysis and Measurements in Situation Awareness*. Daytona Beach: Embry-Riddle Aeronautical University Press.

- Gugerty, L. J. & W. C. Tirre, W. (2000). Individual differences in situation awareness. In M. R. Endsley & D. J. Garland, (Eds.), *Situation Awareness Analysis and Measurement*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Gutwin, C., & Greenberg, S. (2002). A descriptive framework of workspace awareness for real-time groupware. *Computer Supported Cooperative Work*, 11, 411-446.
- Hart, S. G., & Staveland, L. E. (1988). Development of a multi-dimensional workload rating scale: Results of empirical and theoretical research. In P. A. Hancock & N. Meshkati (Eds.), *Human Mental Workload*. Amsterdam. The Netherlands: Elsevier.
- Hartman, B. O. & Secrist, G. E. (1991). Situational awareness is more than exceptional vision. *Aviation, Space and Environmental Medicine*,
- Harwood, K., Barnett, B., & Wickens, C. (1988). Situational Awareness: A conceptual and methodological. In *Proceedings of the Psychology in the Department of defense Eleventh Symposium* (Tech. Report No USAFA-TR-88-1, pp. 316-320). Colorado Springs, CO: US Air Force Academy (AD-A198723).
- Haskell, I. D., & Wickens, C. D. (1993). Two- and three-dimensional displays for aviation: a theoretical and empirical comparison. *The International Journal of Aviation Psychology*, 3(2), 87-109.
- Herdman, C., & LeFevre, J. (2002). *Cognitive Systems Engineering Framework for Modeling and Simulation in the Acquisition Process*. Report prepared for DND Canada: 007SV-W7714-010S47.
- Hopkin, V. D. (1980). The measurement of the air traffic controller. *Human Factors*, 22, 547-560.

Hulme, C., Maughan, S., & Brown, G. D. (1991). Memory for familiar and unfamiliar words: Evidence for a long-term memory contribution to short-term memory span. *Journal of Memory and Language*, 30, 685-701.

Hulme, C., Roodenrys, S., Brown, G., & Mercer, R. (1995). The role of long-term memory mechanisms in memory span. *British Journal of Psychology*, 86, 527-536.

James, W. (1890). *The Principles of Psychology*. New York: Holt.

Johannsdottir, K. R., LeFevre, J.-A., & Herdman, C. (2003). Cognitive System Engineering (CSE) Framework for Evaluating Cockpit Interfaces. . In *Proceedings of The Twelfth International Symposium on Aviation Psychology*, Columbus, OH: The Ohio State University.

Johnson-Laird, P. N. (1983). *Mental Models*. Cambridge, MA: Harvard University Press.

Jones, D. G. (2000). Subjective measures of situation awareness. In M. R. Endsley & D. J. Garland, (Eds.), *Situation Awareness Analysis and Measurement*. Mahwah, NJ: Lawrence Erlbaum Associates.

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

Jonides, J., Smith, E. E., Koeppe, R. A., Awh, E., Minoshima, S., & Mintun, M. A. (1993). Spatial working memory in human as revealed by PET. *Nature*, 363, 623-625.

Kahneman, D., Ben-Ishai, R., & Lotan, M. (1973). Relation of attention to road accidents. *Journal of Applied Psychology*, 58, 113-115.

Kane, M. J., Bleckley, M. K., Conway, A. R. A., & Engle, R. W. (2001). A controlled-attention view of working-memory capacity. *Journal of Experimental Psychology: General*, 130, 169-183.

Kanwisher, N. (2001). Neural events and perceptual awareness. *Cognition*, 79, 89-113.

Kass, S. J., Herschler, D. A., & Companion, M. A. (1990). Are they shooting at me? An approach to training situational awareness. In Proceedings of the Human Factors Society 34th Annual Meeting (pp. 1352-1356). Santa Monica, CA: Human Factors and Ergonomics Society.

Kass, S. J., Herschler, D. A., & Companion, M. A. (1991). Training situational awareness through pattern recognition in a battlefield environment. *Military Psychology*, 3, 105-112.

Kentridge, R. W., & Heywood, C. A. (2000). Metacognition and awareness. *Consciousness and Cognition*, 9, 308-312.

Kentridge, R. W., Heywood, C. A., & Weiskrantz, L. (1997). Residual vision in multiple retinal locations within a scotoma: Implications for blindsight. *Journal of Cognitive Neuroscience*, 9, 191-202.

Kirsh, D., & Maglio, P. (1994). On distinguishing epistemic from pragmatic action. *Cognitive Science*, 18, 513-549.

Klein, G. (1997). The recognition-primed decision (RPD) model: Looking back, looking forward. In C. Zsombok & G. Klein (Eds.), *Naturalistic Decision Making*. Hillsdale, NJ, England: Lawrence Erlbaum Associates.

Lansdown, T. C., Brook-Carter, N., & Kersloot, T. (2004). Distraction from multiple in-vehicle secondary tasks: vehicle performance and mental workload implications. *Ergonomics*, 47, 91-104.

Loftus, G. R., & Masson, M. E. J. (1994). Using confidence intervals in within-subject designs. *Psychonomic Bulletin & Review*, 1, 476-490.

Logie, R. H. (1995). *Visuo-Spatial Working Memory*. Hove, UK: Erlbaum.

Logie, R. H., Baddeley, A. D., Mane, A., Donchin, E., & Sheptak, R. (1989). Working memory in the acquisition of complex cognitive skills. *Acta Psychologica*, 71, 53-87.

Logie, R. H., Engelkamp, J., Dehn, D., & Rudkin, S. (2001). Actions, mental actions, and working memory. In M. Denis, R. H. Logie, C. Cornoldi, J. Engelkamp & M. De Vega (Eds.), *Imagery, Language and Visuo-Spatial Thinking*. Hove, UK: Psychology Press.

Logie, R. H., Gilhooly, K. J., & Wynn, V. (1994). Counting on working memory in arithmetic problem solving. *Memory and Cognition*, 22, 395-410.

Logie, R. H., & Marchetti, C. (1991). Visuo-spatial working memory: Visual, spatial or central executive? In R. H. Logie & M. Denis (Eds.), *Mental Images in Human Cognition*. Amsterdam: Elsevier.

Logie, R. H., Zucco, G., & Baddeley, A. D. (1990). Interference with visual short-term memory. *Acta Psychologica*, 75, 55-74.

Luck, S. J., Vogel, E. K., & Shapiro, K. (1996). Word meaning can be accessed but not reported during attentional blink. *Nature*, 383, 616-619.

- Luria, A. R. (1961). *The Role of Speech in the Regulation of Normal and Abnormal Behaviour*. New York: Liveright Publishing Corporation.
- Mack, A., & Rock, I. (1998). *Inattentional Blindness*. Cambridge, MA: MIT Press.
- Merikle, P. M., Smilek, D., & Eastwood, J. D. (2001). Perception without awareness: perspectives from cognitive psychology. *Cognition*, 79, 115-134.
- Milner, A. D., & Goodale, M. A. (1998). *Precis of The Visual Brain in Action*. *Psyche*, 4(12), available online at <http://psyche.cs.monash.edu.au/v4/psyche-4-12-milner.html>.
- Miyake, A. & Shah, P. (1999). *Models of Working Memory: Mechanisms of Active Maintenance and Executive Control*. Cambridge: Cambridge University Press.
- Mogford, R. H. (1997). Mental models and situation awareness in air traffic control. *The International Journal of Aviation*, 7, 331-341.
- Morris, N. (1987). Exploring the visuo-spatial scratch pad. *The Quarterly Journal of Experimental Psychology*, 39A, 409-430.
- Morris, R. G., Craik, F. I. M., & Gick, M. L. (1990). Age differences in working memory tasks: The role of secondary memory and the central executive system. *Quarterly Journal of Experimental Psychology*, 42A, 67-86.
- Morrow, D., Leirer, V., Altieri, P., Fitzsimmons, C. (1994). When expertise reduces age differences in performance. *Psychology and Aging*, 9, 134-148.
- Navathe, P. Ds., & Singh, B. (1994). An operational definition for spatial disorientation. *Aviation, Space and Environmental Medicine*, 65, 1153-1155.

- Neisser, U. (1976). *Cognition and reality: Principles and implications of cognitive psychology*. San Francisco: W. H. Freeman.
- Nisbet, R., & Wilson, T. (1997). Telling more than we can know: Verbal reports on mental processes. *Psychological review*, 84, 231-259.
- O'Hare, D. (1997). Cognitive ability determinants of elite pilot performance. *Human Factors*, 39, 540-552.
- Olmos, O., Liang, C. & Wickens, C. D. (1995). Construct validity of situation awareness measurements related to display design. In D. J. Garland & M. R. Endsley, (Eds.), *Experimental Analysis and Measurements in Situation Awareness* (259-65). Daytona Beach: Embry-Riddle Aeronautical University Press.
- Peres, M., Van De Moortele, P. F., Pierard, C., Pharm, D., Lehericy, S., Satabin, P., Le Bihan, D., & Guezennec, C. (2000). Functional magnetic resonance imaging of mental strategy in a simulated aviation performance task. *Aviation, Space and Environmental Medicine*, 71, 1218-1231.
- Pew, R. W. (1992). An introduction to the concept of situation awareness. In R. D. Gilson, D. J. Garland, and . M. Koonce (Eds.), *Situational Awareness in Complex Systems*. Daytona Beach FL: Embry-Riddle Aeronautical University Press.
- Pew, R. W. (2000). The state of situation awareness measurement: Heading toward the next century. In M. R. Endsley & D. J. Garland, (Eds.), *Situation Awareness Analysis and Measurement*. Mahwah, NJ: Lawrence Erlbaum Associates.

Pickering, S. J., Gathercole, S. E., Hall, M., & Lloyd, S. A. (2001). Development of memory for pattern and path: Further evidence for the fractionation of visuo-spatial memory. *The Quarterly Journal of Experimental Psychology*, 54A, 397-420.

Posada, A., Franck, N., Georgieff, N., & Jeannerod, M. (2001). Anticipating incoming events: an impaired cognitive process in schizophrenia. *Cognition*, 81, 209-225.

Pritchett, A. R., Hansman, R. J., & Johnson, E. N. (1995). Use of testable responses for performance-based measurement of situation awareness. In D. J. Garland & M. R. Endsley, (Eds.), *Experimental Analysis and Measurements in Situation Awareness* (259-65). Daytona Beach: Embry-Riddle Aeronautical University Press.

Quinn, J. G. (1991). Encoding and maintenance of information in visual working memory. In R. H. Logie & M. Denis (Eds.). *Mental Images in Human Cognition*, pp. 105-115. Amsterdam: Elsevier.

Quinn, J. G., & Ralston, G. E. (1986). Movement and attention in visual working memory. *The Quarterly Journal of Experimental Psychology*, 38A, 689-703.

Recarte, M. A., & Nunes, L. M. (2000). Effects of verbal and spatial-imagery tasks on eye fixation while driving. *Journal of Experimental Psychology: Applied*, 6, 31-43.

Recarte, M. A., & Nunes, L. M. (2003). Mental workload while driving: effects on visual search, discrimination and decision making. *Journal of Experimental Psychology: Applied*, 9, 119-137.

Regal, D. M., William, H. R., Boucek, G. P. (1988). Situational awareness in the commercial flight deck – definition, measurement, and enhancement. In *Proceedings of the 7th SAE Aerospace Behavioral Technology Conference*, 65-69.

Robbins, T. W., Anderson, E. J., Barker, D. R., Bradley, A. C., Fearnlyhough, C., Henson, R., & Hudson, S. R. (1996). Working memory in chess. *Memory & Cognition*, 24, 83-93.

Roscoe, S. N. (1993). An aid in the selection process –WOMBAT. *Civil Aviation Training*, 4, 48-51.

Roscoe, S. N., & Corl, L. (1987). Wondrous original method for basic airmanship testing. In R. S. Jensen (Ed.), *Proceedings of the Fourth International Symposium on Aviation Psychology* (pp. 493-499). Columbus: Ohio State University, Department of Aviation.

Roscoe, S. N., Corl, L., & LaRoche, J. (1997). *Predicting Human Performance*. Canada: Helio Press.

Rubinstein, J. S., Meyer, D. E., & Evans, J. E. (2001). Executive control of cognitive processes in task switching. *Journal of Experimental Psychology: Human Perception and Performance*, 27, 763-797.

Rugg, M. D. (1992). Conscious and unconscious processes in language and memory-commentary. In A. D. Milner & M. D. Rugg (Eds.), *The Neuropsychology of Consciousness*. London, UK: Academic Press Limited.

Salas, E., Prince, C., Baker, D. P., & Shrestha, L. (1995). Situation awareness in team performance: Implications for measurement and training. *Human Factors*, 37, 123-136.

Salway, A. F. S., & Logie, R. H. (1995). Visuospatial working memory, movement control and executive demands. *British Journal of Psychology*, 86, 253-269.

- Sarter, N. B., & Wood, D. D. (1991). Situation awareness: A critical but ill-defined phenomenon. *The International Journal of Aviation psychology*, 1, 45-57.
- Sarter, N. B., & Wood, D. D. (1995). How in the world did we ever get into that mode? Mode error and awareness in supervisory control. *Human Factors*, 37, 5-19.
- Schacter, D. L. (1987). Implicit memory: History and current status. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 13, 501-518.
- Selcon, S. J., & Taylor, R. M. (1989). Evaluation of the situational awareness rating technique (SART) as a tool for aircrew systems design. In *Proceedings of the AGARD AMP Symposium on Situational Awareness in Aerospace Operations*, CP478. Seuilly-sur-Seine: NATO AGARD.
- Selcon, S. J., Taylor, R. M., & Koritsas, E. (1991). Workload or situation awareness? TLX vs. SART for aerospace system design evaluation. In *Proceedings of the Human Factors Society, 35th Annual Meeting* (pp. 57-61). Santa Monica, CA: Human Factors Society.
- Shah, P., & Miyake, A. (1996). The separability of working memory resources for spatial thinking and language processing: An individual differences approach. *Journal of Experimental Psychology*, 125, 4-27.
- Shebilske, W. L., Goettl, B. P., & Garland, D. L. (2000). Situation awareness, automaticity, and training. In M. R. Endsley & D. J. Garland, (Eds.), *Situation Awareness Analysis and Measurement*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Shepard, R. N., & Metzler, J. (1971). Mental rotation of three-dimensional objects. *Science*, 191, 701-703.

Singer, M., & Richot, K. F. M. (1996). The role of working memory capacity and knowledge access in text inference processing. *Memory & Cognition*, 24, 733-743.

Smith, K. & Hancock, P. A. (1995). Situational awareness is adaptive, externally directed consciousness. *Human Factors*, 37, 137-48.

Smith, E. E., Jonides, J., & Koeppe, R. A. (1996). Dissociating verbal and spatial working memory using PET. *Cognition*, 65, 167-196.

Smyth, M. M., & Pendleton, L. R. (1989). Working memory for movements. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, 41, 235-250.

Smyth, M. M., & Scholey, K. A. (1994). Interference in immediate spatial memory. *Memory and Cognition*, 22, 1-13.

Spick, M. (1988). *The Ace Factor: Air Combat and the Role of Situational Awareness*. Annapolis, MD: Naval Institute Press.

Stoke, A. F., Kemper, K., & Kite, K. (1997). Aeronautical decision making, cue recognition, and expertise under time pressure. In C. Zsombok & G. Klein (Eds.), *Naturalistic Decision Making*. Hillsdale, NJ, England: Lawrence Erlbaum Associates.

Sullivan, C., & Blackman, H. S. (1991). Insight into pilot situation awareness using verbal protocol analysis. In Processings of the Human Factors Society 35th Annual Meeting, (pp. 57-61). Santa Monica, CA: Human Factors and Ergonomics Society.

Taylor, R. M. (1989). Situational awareness rating technique (SART): The development of a tool for aircrew systems design. In *Proceedings of the NATO AGARD Conference on Situation Awareness in Aerospace Operations* (AGARD-CP-478). Springfield, VA: National Technical Information Service.

Taylor, R. M. (1990). Situation awareness rating techniques (SART): the development of a tool for aircrew system design. In *Situational Awareness in Aerospace Operations* (Chapter 3). France: Neuilly-sur-Seine, NATO-AGARD-CP-478.

Taylor, R. M. & Selcon, S. J. (1994). Situation in Mind: Theory, Application and Measurement of Situation Awareness.

Tenny, 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 (NASA Contractor Report 4451). Hampton, VA: NASA-Langley Research Center.

Tham, M. & Kramer, A. (1994). Attentional control and piloting experience. In *Proceedings of the Human Factors and Ergonomics Society 38th Annual Meeting*, (Vol. 1, pp. 31-35). Santa Monica, CA: Human Factors and Ergonomics Society.

Towse, J. N., & Valentine, J. D. (1997). Random generation of numbers: A search for underlying processes. *European Journal of Cognitive Psychology*, 9, 381-400.

Towse, J.N. (1998). On random generation and the central executive of working memory. *British Journal of Psychology*, 89, 77-101.

Uhlarik, J., & Comerford, D. A. (2002). A review of situation awareness literature relevant to pilot surveillance functions (Report DOT/FAA/AM-02/03). Washington DC: Office of Aerospace Medicine.

Vandierendonck, A. (2000). Analyzing human random time generation behaviour: A methodology and a computer program. *Behavior Research Methods, Instruments, & Computers*, 32, 555-565.

Vandierendonck, A., De Vooght G., & Van der Goten, Koen (1998). Does random time interval generation interfere with working memory executive functions? *European Journal of Cognitive Psychology*, 10, 413-442.

Van Dijk, T. A. & Kintsch, W. (1983). *Strategies of Discourse Comprehension*. New York: Academic Press.

Venturino, M., Hamilton, W. L., & Dvorchak, S. R. (1989). Performance-based measures of merit for tactical situation awareness. In *Situational Awareness in Aerospace Operations (AGARD-CP-478)*. (pp. 4/1-4/5). Neuilly Sur Seine, France: NATO-AGARD.

Vidulich, M. A. (1989). The use of judgement 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 and Ergonomics Society.

Vidulich, M. A. (2002). Testing the sensitivity of situation awareness metrics in interface evaluations. In M. R. Endsley & D. J. Garland, (Eds.), *Situation Awareness Analysis and Measurement*. Mahwah, NJ: Lawrence Erlbaum Associates.

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 and Ergonomics 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, A7-12.
- Vidulich, M. A., McCoy, A. L., & Crabtree, M. S. (1996). Attentional control and situational awareness in a complex air combat simulation. In *AGARD Conference Proceedings 575-Situation Awareness: Limitations and Enhancement in the Aviation Environment* (pp. 18-1-18-5).
- Vygotsky, L. (1934/1986). *Thought and Language* (A. Konzulín, Trans.). Cambridge, MA: MIT Press (Original work published in 1934).
- Waag, W. L., & Houck, M. R. (1994). Tools for assessing situational awareness in an operational fighter environment. *Aviation, Space and Environmental Medicine*, 65, A13-19.
- Weisenkrantz, L. (1986). *Blindsight: A Case Study and Implications*. Oxford: Clarendon Press.
- Weisenkrantz, L. (1990). Outlooks for blindsight: Explicit methodologies for implicit processes. *Proceedings of the Royal Society*, B239, 247-278.
- Whalen, P. J., Rauch, S. L., Etcoff, N. L., McInerney, S. C., Lee, M. B., & Jenike, M. A. (1998). Masked presentations of emotional facial expressions modulate amygdala activity without explicit knowledge. *The Journal of Neuroscience*, 18, 411-418.
- Whitaker, L. A. & Klein, G. A. (1988). Situation awareness in the virtual world: Situation assessment report. In *Proceedings of the 11th Symposium of Psychology in the Department of Defense*, April.

Whitney, P., Ritchie, B. G., & Clark, M. B. (1991). Working memory capacity and the use of elaborative inferences in text comprehension. *Discourse Processes*, 14, 133-145.

Wickens, C. D. (1992). Workload and situation awareness: An analogy of history and implications. *Insight: The Visual Performance and Technical Group Newsletter* (pp. 1-3).

Wickens, C. D. (2000). The trade-off of design for routine and unexpected performance: Implications of situation awareness. In M. R. Endsley & D. J. Garland, (Eds.), *Situation Awareness Analysis and Measurement*. Mahwah, NJ: Lawrence Erlbaum Associates.

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

Wickens, C. D., & Hollands, J. G. (1999). *Engineering Psychology and Human Performance*. New Jersey: Prentice Hall.

Williams, K. W. (2002). Impact of aviation highway-in-the-sky displays on pilot situation awareness. *Human Factors*, 44, 18-27.

Young, A. W. (1994). Neuropsychology of awareness. In, A. Revonsuo & M. Kampinen (Eds.), *Consciousness in Philosophy and Cognitive Neuroscience*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Zwaan, R. A., & Radvansky, G. A. (1998). Situation models in language comprehension and memory. *Psychological Bulletin*, 123, 162-185.

Zwaan, R. A., Langston, M. C., & Graesser, A. C. (1995). The construction of situation models in narrative comprehension: An event-indexing model. *Psychological Science*, 6, 292-297.

Appendix A

Appendix B