Examining the Relationship between Working Memory Capacity, Team Situational Awareness, and Team Performance in Extreme Work Teams

by

Garrett Morawiec

A thesis submitted to

the Faculty of Graduate and Postdoctoral Affairs

in partial fulfillment of the requirements for the degree of

Masters of Arts

in

Psychology

Carleton University

Ottawa, Canada

©2013 Garrett Morawiec

Abstract

Extreme work teams such as the ones found in the military, police, and firefighter domains rely heavily on team situational awareness (TSA). In these domains, a lack of TSA has been associated with the loss of life. While there are several factors which may contribute to TSA and team performance (e.g. communication), selecting the right individuals with the right cognitive abilities may also play an important role. This study, therefore, examined the relationship between a team's mean working memory capacity (WMC), TSA, and team performance. Results showed that while within-team WMC differences were not associated with TSA or team performance, teams with a higher mean WMC did have a positive relationship with TSA, team specific queries, and had higher overlapping correct SA. A positive relationship was also found between TSA and two separate forms of team way-finding performance.

Keywords: working memory capacity, team situational awareness, team performance

Acknowledgements

First and foremost, I would like to thank Dr. Avi Parush for his patience, guidance, and wisdom. His desire to see his students learn and succeed is very evident. I would also like to thank Sharmili Shanmugaratnam, Chunyun Ma, and Nadya Rustandjaja, all of whom helped me along the way by answering my numerous questions.

My sincere thanks also go to Dr. Logan and Dr. Whitehead who provided me with constructive feedback as members of my defence committee. To LCol Johnson, PhD, from Director General Military Personnel Research and Analysis (DGMPRA), thank you for your leadership while I was working at DGMPRA and to Dr. Jalbert (DGMPRA) for providing feedback on my prospectus.

Finally, I deeply appreciate the patience and support of my wife Jennifer, who is no doubt happy that I am done, as well as to my four year old son Liam who somehow understood why Daddy could not always play cars with him. To my newborn son Ethan, your ability to sleep through the night early on was greatly appreciated.

Table of Contents

List of Figures

List of Tables

List of Appendices

A blazing house fire, a hostage situation, the neutralization of enemy forces defending complex terrain, or even a government's reaction to natural or manmade disasters require the heroic efforts of extreme work teams. Jones and Hinds (2002) define extreme work teams as those "that come together for a single event, are highly interdependent, and whose performance can save or cost lives" (p. 1). Such work environments are extremely dangerous, very dynamic, stressful, and often unpredictable. In such conditions, one of the major factors that degrade team performance is the loss of team situational awareness (TSA). The following are two examples from different domains that depict such extreme situations.

On September 11, 2001, the world changed as terrorists flew planes into the north and south tower of the World Trade Center, as well as the Pentagon, while passengers of Flight 93 stopped the terrorists from accomplishing their mission but crashed in the process. At the World Trade Centre site, the scene was horrific, as both towers were engulfed in flames. The Fire Department of New York (FDNY) responded by deploying hundreds of on and off duty firefighters whose primary mission was to enter the towers and rescue civilians. In an analysis of the communication transcripts from the event, Garrity (2007) discovered many instances where situational awareness (SA) was degraded and at times, completely lost. From the firefighters within the buildings who were confused as to which buildings they were in, to receiving reports from the police department that an elevator was about to come down but not knowing which building, resulting in both buildings losing important access to the elevators, to complete SA breakdown when the south tower collapsed (Garrity, 2007). As the world witnessed the collapse of the towers live, orders to evacuate the firefighters might have come sooner had the fire department commanders received pertinent information from the police who had a helicopter in the air at the time (Lipton & Glanz, 2002). Ultimately, 343 firefighters lost their lives that day

(Lipton & Glanz, 2002). It should be stressed, however, that despite these shortcomings, there were many examples when the FDNY successfully fought to regain SA (see Garrity, 2007) and ultimately, the first responders reacted with "skill and intensity" (Cohen, Eimicke, & Horan, 2002).

While for firefighters this was an extraordinary event, it nonetheless demonstrates how TSA can breakdown within the firefighting domain. In the military domain, history is littered with examples in which TSA has broken down, resulting in the deaths of soldiers. On 18 April, 2002, members of the 3rd Battalion, Princess Patricia's Canadian Light Infantry Battle Group (3 PPCLI BG) were conducting live-fire training exercises at Tarnak Farms, Afghanistan, while deployed as part of Operation Enduring Freedom (OEF). While en-route back to base, two US F-16 fighter jets (call sign COFFEE flight) mistakenly thought they were being engaged by hostile forces from the ground, this led to one of the fighter jets dropping a Mark 82 500lb laser guided bomb on the Canadian forces, resulting in the "friendly-fire" deaths of four soldiers (Baril, Matte, Hodgson, & Levesque, 2002). One of the conclusions of the American investigation cited "the lack of situational awareness exhibited by COFFEE flight follows from poor planning and preparation combined with problems with attention, misperceptions, and fatigue" (Dumais & Sargeant, 2002).

The relevance of these examples is to provide context to the fluid, dangerous, and unpredictable life and death conditions for which team situational awareness becomes paramount and for which the present study seeks to simulate. When examining such cases for improvements, the obvious response is to analyze team processes, such as communication and planning. However, improvements can also be gained by ensuring that the initial personnel selection only selects those that possess the cognitive abilities to perform in this specific environment. Specifically, as a result of the dynamic and, at times, novel situations members of extreme teams are placed in, attention and working memory capacity (WMC), especially in novice personnel, may likely play a significant role in the attainment of individual and TSA. Indeed, as Endsley (2000) explains:

> The way in which attention is employed in a complex environment with multiple competing cues is essential in determining which aspects of the situation will be processed to form situation awareness. Once taken in, information must be integrated with other information, compared to goal states and projected into the future – all heavily demanding on working memory (p. 8-9).

As one will see below in the section on WM theories, individuals vary with respect to their WMC and their attentional control, therefore, understanding these differences and how they may impact TSA and team performance becomes important.

Therefore, the primary purpose of this study is to examine the relationship between individual attention/WMC, TSA and team performance, something that has not been investigated before. To accomplish this, a literature review of the relevant constructs and their relationships will be explored. The first half of the review will examine theories on attention/working memory (WM), individual SA, and TSA. The second half of the paper will discuss studies that have explored the relationships between TSA and team performance, WMC and team performance, and WMC and individual SA. This will lead us to a formulation of relevant research questions and their hypothesis.

Attention/Working Memory

A General Information Processing Model

The human information processing (IP) system is capable of astounding feats of intellectual prowess, yet, it is inherently limited. Before focusing on the limiting IP components of attention and WM, it is important to appreciate the purpose of IP models and to be able to understand how attention and WM fit within it. As Massaro and Cowan (1993) explain, the basic premise of IP models is to be able to trace information flow from the stimulus in the environment through the human IP system all the way to the individual's output (or response). Atkinson and Shiffrin's (1968) model is one of the earlier and most influential (Baddeley, 2007). Their model is comprised of three stores: a sensory register, which registers external stimuli, with decay occurring within several hundred milliseconds; a short-term store (STS), a unitary component that holds information and processes it with decay occurring around 15-30 seconds; and a long-term store which, unlike the two previous stores, information is relatively permanent with the exception that it may be modified or become temporarily irretrievable (Atkinson & Shiffrin, 1968). Over the years, researchers have expanded and modified this original IP model. Industrial organizational (IO) psychologists Arthur, Doverspike, and Bell (2004), who are particularly interested in the testable individual differences that can be derived from IP theories, outline a more recent general model (see Figure 1). The main purpose of showing this model is to highlight how attention and WM fit within a more complex IP architecture. It begins with the input of stimulus information from the environment through the sensory system into a short-term sensory store. This information is then detected by a perceptual mechanism which discriminates the importance of the information and organizes it. A decision and response selection mechanism then compares it to information in memory and makes a decision. Working memory, which sometimes is still referred to and confused with short-term memory, is responsible for holding the information obtained from the environment and simultaneously retrieving information from long-term memory (LTM). The storage capacity of LTM is thought to be limitless and holds all of our knowledge (e.g. in schemas and episodic memories). Once a decision has been made, a mechanism then controls the selection and response execution. Finally, attentional resources overarch all of these mechanisms and responses.

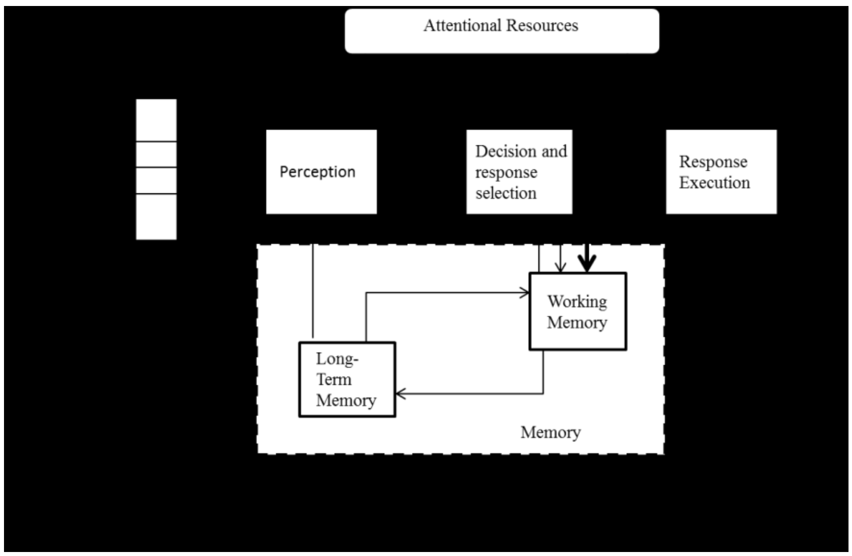


Figure 1. Arthur, Doverspike, and Bell's (2004) general information processing model.

It is important to note that while attention/WM may be just one aspect of the model, they are quite important. Indeed, as Cowen (2005) explains, limited attentional resources and WMC are the reason people may feel overwhelmed by new information, whether it is because there is too much, too confusing, or too complex. This capacity issue would quite obviously be of great concern for extreme work teams and will therefore be a central focus of the attention/WM review. Overall, there is no agreed upon model of attention/WM, however, there is little disagreement that it is of limited capacity and that individuals differ in this capacity. The following sections will discuss several of the major theories.

Baddeley's Working Memory Model

Until Baddeley and Hitch's (1974) seminal work on WM, short term store (STS) or short term memory (STM) were the terms most often used to describe the unitary storage system that held and processed information obtained from the environment and from LTM (see Atkinson & Shiffrin, 1968). Baddeley and Hitch (1974) advanced the theory and proposed a multi-component WM model which allowed for information to be held separate from the component processing it. Key components of their model include a central executive responsible for the processing of information and two "slave systems" responsible for the temporary storage of information called the articulatory loop and the visuo-spatial sketchpad (VSSP). Throughout the years, their model has been extensively studied and referenced while undergoing minor modifications. Figure 2 depicts Baddeley's (2012) most current model.

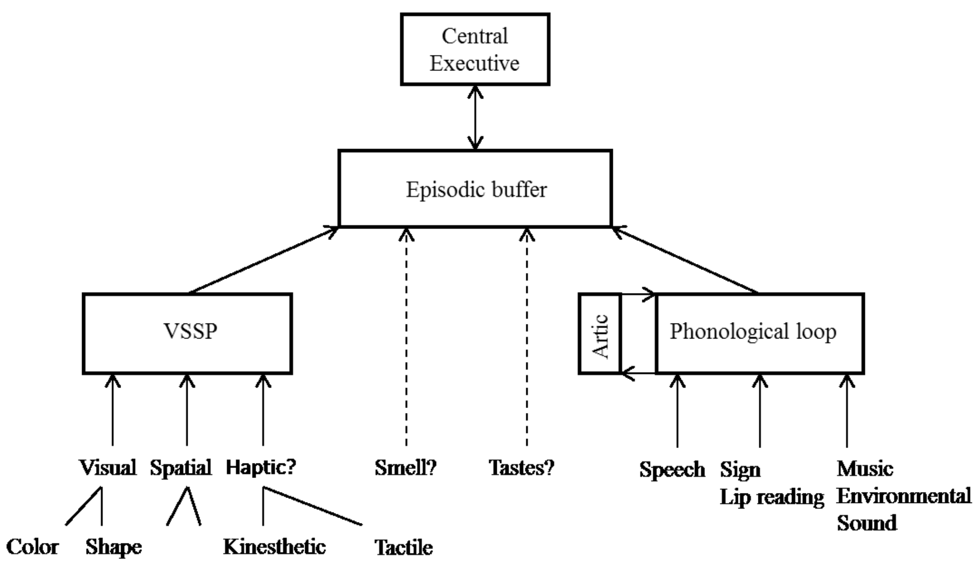


Figure 2. Baddeley's (2012) most current model of WM.

The current model is comprised of the central executive, the VSSP, the phonological loop (formerly known as the articulatory loop), and the episodic buffer, all of which having limited capacity to process information. Baddeley's approach over the years has been to tackle WM from the bottom up. Focusing first on the phonological loop, Baddeley (1986, 2007, 2012) explains that it is comprised of a phonological store which temporarily holds verbal material and visually presented material that has been verbally recoded, as well as, an articulatory control process that keeps that information active through vocal and sub-vocal rehearsal. According to Baddeley, Thomson, and Buchanan (1975), the phonological loop has a capacity equal to the number of words that an individual can say in two seconds.

The VSSP is also a temporary storage and rehearsal system that is "capable of integrating visual and spatial information, whether acquired from vision, touch, language or LTM, into a unitary visuo-spatial representation" (Baddeley, 2007, p. 64). The VSSP can be divided into a spatial component, responsible for dynamic information such as movement and direction, a visual component, responsible for storage of static information such as objects and their features, and a temporal component, which stores sequential actions like those required to learn new motor skills (Baddeley, 2007; Dehn, 2008). The capacity limitations for these components appear to be around four, whether it is the number of sequences an individual can mimic in the Corsi's Block Tapping Test, or the number of objects held in memory (Baddeley, 2007).

While the two slave systems discussed above temporarily store the information being used, the executive control attends to the information and processes it. Baddeley (1986) originally incorporated Norman and Shallice's (1980) supervisory attention system (SAS) into his model to explain the central executive. According to their model, while much of the time the SAS is minimally engaged due to the automaticity of our behavior in the form of schemas (e.g. overlearned behavior such as driving a car), it does become heavily engaged for planning and decision making, when our schemas come into trouble, novel or poorly learned sequences, or dangerous or difficult tasks (Baddeley, 1986). While eventually introducing his own functions of the central executive (see Baddeley, 1996), his most recent list includes: being responsible for the division of attention between two tasks, the ability to focus on one task while disregarding other stimuli, switching between tasks and, finally, interfacing with LTM (Baddeley, 2007, 2012).

As a result of several unexplainable phenomena over the years, Baddeley (2000) added a fourth component to his model called the episodic buffer. The episodic buffer is "a limited-capacity temporary storage system that is capable of integrating information from a variety of sources. It is assumed to be controlled by the central executive, which is capable of retrieving information from the store in the form of conscious awareness, of reflecting on that information and, where necessary, manipulating and modifying it" (Baddeley, 2000, p. 421). The information stored in the episodic buffer comes in preformed chunks that have already been binded in the LTM, as in the case of sentences, or binded at perception, as in the case of objects and their features in the visual system (Baddeley, 2012). With respect to a capacity limitation of both the central executive and the episodic buffer, Baddeley (2007, 2012) refers to Cowan's (2005) limitation, on average, of four chunks. With multiple components each with their own limited capacity, one can begin to get a sense for how easily an extreme work environment may overwhelm the capacity of anyone of these components, leaving the individual at a loss for information and, ultimately, at a loss for SA. While Baddeley's WM model has been the "the industry standard and for reason" (Cowan, 2005, p.21), others have proposed solid alternatives.

Cowan's Embedded-Processes Model

Unlike Baddeley's model which suggested multiple components, Cowan (1988, 1995) proposed a model in which the focus of attention and a short-term store are embedded within LTM. Cowan (1988) identified the main components of his model as being: (a) a sensory storage, (b) long-term memory, (c) short-term store (activated portions of the long-term memory), and (d) a central executive that controls the focus of attention. In this model, sensory information from the environment is held for a brief time in the sensory store. This sensory information activates schemas within LTM, however, if this sensory information is not novel or special in any way (i.e. the individual has been habituated to the information), it will remain within the activated, but unconscious, short-term memory store. Conversely, if the information is novel or special, it will activate the focus of attention automatically at which time the central executive will become engaged, controlling the focus of attention outward toward the stimuli or inward toward information stored in the LTM, see figure 3.

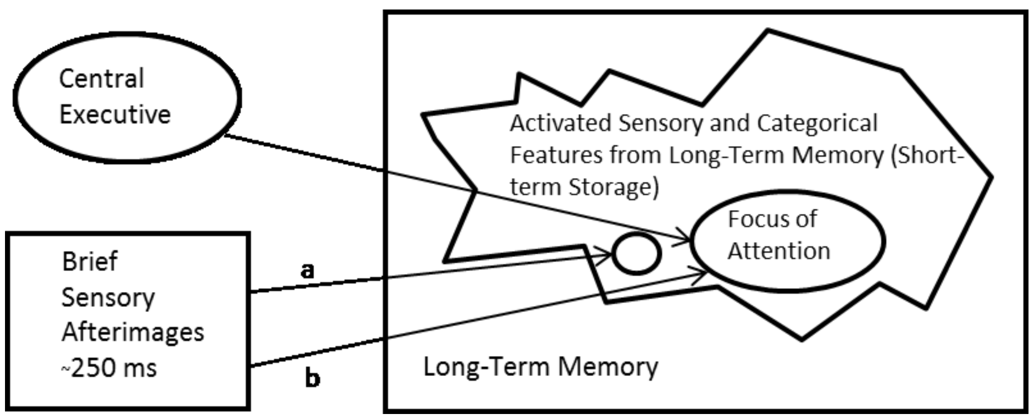


Figure 3. Cowan's (1988, 1995, 2005) Embedded Processes Model.

Another and perhaps one of the most important differences between Baddeley's and Cowan's models, is that the only capacity limitation in the Embedded Process Model is the number of chunks that a healthy adult can hold in the focus of attention. Cowan (2001) concluded that this number, on average, was four, however, individuals varied. To understand what constitutes a chunk, Cowan (2001) defines it as "a collection of concepts that have strong associations to one another and much weaker associations to other chunks concurrently in use" (p.89). While Cowan (2005) uses Miller's (1956) example of being able to recall nine letters, IBM, CIA, and FBI, because they were chunked into three familiar acronyms, he also explains that chunking can occur not just with preexisting associations but can also occur on the spot by converting seven digits into a smaller number of chunks as is commonly done with telephone numbers. In addition to capacity differences, individuals may also differ in their ability to adjust their focus of attention, zooming in or out to the right level depending on complexity, while at the same time deactivating irrelevant chunks, which Cowan called effortful processing (Cowan, 2005, 2010). Whether completing a novel task or a complex one, Cowan (1999) explains that LTM must be placed in a highly activated state in the focus of attention. Information coming into the focus of attention from the environment must be combined and integrated with existing LTM information, often in new ways, by the central executive. Knowing that individuals have a limited focus of attention, it can be easy to see how an extreme environment might bombard the senses with too much information, leaving much of it outside the focus of an individual's attention and therefore not processed. Should critical pieces of information not be perceived and processed, SA begins to degrade, making the situation potentially dangerous for everyone. Obviously, the bigger someone's WMC, the better. While Baddeley (1996) explained that he consciously did not pursue trying to understand individual differences per se, Cowan's model certainly explored it and the next model focused on it.

Engle et al.'s Individual Differences Model

The individual differences model is structured very similarly to Cowan's model with various levels of activation within LTM. Specifically, Engle, Kane, and Tuholski's (1999) explain that memory traces within LTM become activated to varying degrees, either to the point they can be considered as entering into short-term memory, or if activated to a high enough degree, they enter controlled attention. The fundamental difference between this model and Cowan's model, which emphasizes how many chunks can be in the focus of attention, is that

Engle et al.'s model is about "the capacity for controlled sustained attention" (Engle et al., 1999, p. 104), not about a number or time (Engle et al., 1999; Heitz, Unsworth, & Engle, 2005). This controlled sustained attention is "a voluntary, effortful cognitive act that serves to maintain information through activation of relevant brain circuitry, inhibit the irrelevant and distracting information that impinges on us at any one time, and suppress prepotent response tendencies that are task irrelevant" (Heitz et al., 2005, p. 63). This attention may take the form of vigilance (maintaining attention), divided attention (dual task), and selective attention (disregard multiple sources of information and focus on the required stimulus) (Heitz et al., 2005). To better understand the nature of individual differences, experiments were conducted by grouping together participants who scored either high or low on traditional measures of WMC. While the score itself was of little concern, the idea was to determine how these two groups differed with respect to controlling their attention. As an example, Kane, Conway, Bleckley and Engle (2001) required participants to fixate on a particular point on a computer screen, then either attend to or disregard a flash cue. As predicted, there were no differences between the groups with high and low WMC when having to attend to the flash (since it's an automatic behavior), there were, however, differences between the groups when they had to disregard the flash. In an extreme environment then, it would appear from this model that those that can maintain focus on the important tasks at hand and disregard distractors (perhaps including one's own emotions) will allow for the perception and processing of information that leads to good SA. Those that cannot focus would likely be dangerously susceptible to large losses of SA.

Measuring Working Memory Capacity

Far less controversial than the model of working memory is how to measure it. The first widely accepted and dominant measure of WMC is Daneman and Carpenter's (1980) reading span (RSpan) task which taps both the storage and processing components of WM (Heitz et al., 2005; Baddeley, 1996). Daneman and Carpenter (1980) required participants to read out loud sentences typed onto index cards, one sentence per card, each unrelated and ending with a different word. Each set was presented three times and consisted of either two, three, four, five or six sentences each that were separated by a blank card. The blank card signaled the participant to recall the last word in each sentence in the correct order. As the sets continued to increase, the level (the maximum number of words) at which the participant recalled correctly two of the three sets was taken as their WMC. The result of their study showed a quantitative difference between high and low span individuals which they attributed to the readers chunking efficiency, that is, efficient readers chunk the material better leaving more WMC. Their explanation implied that their task is domain specific to reading. An alternative to the efficiency explanation was put forth by Turner and Engle (1980) who simply proposed that some people have greater WMC than others whether it is a reading or non-reading task, in other words, a complex span task that measures WMC is domain general. To demonstrate this, Turner and Engle (1980) theorized that reading comprehension can be predicted by any complex span task. They used four types, one in which participants read out a sentence and verified if the sentence made sense, then had to remember the last word (sentence word). The second was the same except they had to remember a digit at the end (sentence digit). The third involved performing a simple arithmetic operation and then remember a word at the end (operation word), while the fourth was the same except after the simple arithmetic, a digit had to be remembered (operation digit). For all tasks, it was the total number of words or digits remembered in order that reflected WMC. Results showed that all four complex span tasks were significantly correlated with reading comprehension. Knowing that verbal and quantitative skills significantly correlate with each other and therefore may be the reason the operation span (OSpan) tasks were correlated with reading comprehension, they partialed out the participant's quantitative abilities by using their QSAT scores. Results still showed a significant partial correlation between sentence word and operation word span tasks and reading comprehension. Ultimately, this demonstrated that WMC scores are not based on a specific strategy or efficiency of a reader but reflects a more domain general capacity. Therefore, regardless of the complex span task, they are designed to resemble the demands placed on the WM during complex cognitive tasks (Miyake & Shah, 1999). Complex span tasks such as the RSpan and OSpan are widely used to this day and are available in automated versions (see Redick et al., 2012).

While complex span tasks are the norm in measuring WM's storage and processing function, what exactly it is measuring depends on the model. Baddely (2007) explains that complex span tasks are influenced by both the central executive and the phonological loop. Therefore, using the RSpan as an example, the measure takes into account the limited capacity of the phonological loop, which would be responsible for the short term storage and rehearsal of the to be remembered word and sentence, while the limited attentional capacity of the central executive would be responsible for processing the sentence and maintaining focus on the dual task. On the other hand, as Cowan (2005) would likely explain, the RSpan would reflect a participant's maximal capacity (i.e. their focus of attention is "zoomed out") to keep both the word to remember and sentence to be processed activated in memory while disregarding any interference. Finally, Engle et al. (1999) describe that differences in complex span tasks such as the RSpan and OSpan, are reflections of the participant's capacity to maintain attention. Therefore, participants that score high can maintain focus on the dual task of the RSpan while those that score low may become easily distracted (either from thoughts or stimulus in the environment) and unable to focus on the task at hand.

Summary of Working Memory

While the structure and emphases between the three WM models presented above may differ, they all agree that WM is of limited capacity and that individuals differ in this capacity. Further, it is clear that novel or complex tasks require WM to process incoming information, hold it, incorporate information stored in LTM and, once processed, store the new information into LTM. Thus, in a dynamic, complex environment such as the ones faced by extreme work teams, WM may play a fundamental role in individual and TSA. Finally, while attention and WM are generally studied as different constructs, it is clear from the literature above that attention is a part of WM. For this reason and for ease of reading, the term WM will include attention unless a differentiation is required. As one will see below, the importance of WM in achieving SA is crucial. However, the constructs of SA and especially TSA are still not well understood. The following is a brief review of the challenging concepts of SA and TSA.

Situational Awareness and Team Cognition

Individual SA

Exploring TSA would not be very successful without an understanding of individual SA from which TSA emerges. Endsley's (1995) individual SA model will be used as the theoretical backdrop of this study for several reasons. First, it is the most widely cited model (Breton & Rousseau, 2003; Shanmugaratnam & Parush, 2009), second, it focuses on human information processing, e.g. WM, a key construct of the present study (Endsley et al., 2000; Shanmugaratnam & Parush, 2009), and finally, the environment, namely the military, from which Endlsey has conducted many of her experiments, is precisely the sort of extreme environment from which this study seeks to simulate. It should be noted, however, that despite the popularity of Endsley's model, the construct remains profoundly contentious (Salmon et al., 2008). As Stanton, Salmon, Walker, and Jenkins (2010) point out, there are different schools of thought, including the psychology discipline, which focuses on SA as an individual phenomena, design engineers, who look to displays as containing all the required SA information, and finally, ergonomics that say that SA emerges from the interaction between the individual and the objects in the environment. This study is taking the psychological approach, specifically Endsley's, not only for the reasons stated above but also because there is no interaction with artifacts that help maintain SA (e.g. displays) within this environment. That is, the firefighters within this study do not use any technological aids to assist them in maintaining SA.

Individual SA has been defined by Endsley (1988) as "the perception of the elements in the environment within a volume of time and space, comprehension of their meaning, and the projection of their status in the near future" (p. 97). Endlsey (1995) followed this definition with a three level model of SA. The first level begins with an individual's "perception of elements in the environment" (p. 36). As an example, Endsley (1995) uses a military tactical commander who must perceive accurate data on relevant enemy and friendly force's locations, type, number, and capabilities. The second level is the "comprehension of the current situation" (p.37). She explains that this comes about from synthesizing the information from the fragmented elements in level one into a coherent understanding of the current, dynamic, environment within the context of one's own goals. To continue with the military tactical example, after perceiving a certain number and type of enemy vehicles at a particular location, a soldier may comprehend that this is a reconnaissance element from a much larger enemy fighting force. Finally, the third level is the "projection of future status" (p.37). From understanding the situation in level two, military tactical commanders can now, for example, project that within a few hours, a large enemy force will be in their location and that this necessitates the request for additional support from artillery, air, or ground troops. As Endlsey (1995) explains, there are numerous task, system, and individual factors which affect the process of obtaining SA. Within the individual factors, human-information processing mechanisms, that is, perception, working memory, and LTM, all play a role. Indeed, “in dynamic environments, the development of situational awareness and the decision process are restricted by limited attention and working memory capacity for novices and those in novel situations” (Endlsey & Jones, 1997, p. 17). Understanding that attention and WM play a fundamental role in attending to, or perceiving, stimuli in the environment, holding that information and processing it with the help of stored knowledge within LTM, it is easy to see the importance of WM in obtaining individual SA. In fact, WM is so tied to Endlsey’s model of individual SA that Salmon et al. (2008) question whether they can even be considered two distinct constructs. Now take that individual and place them in a team setting where, as one will see in the next section, additional demands are placed on the individual’s WM as they must juggle the demands of sharing and receiving the right information at the right time in the right quantity. It now becomes possible to see the potential relationship between WMC and TSA.

Team Situational Awareness

Before moving on to TSA, it would be helpful to contextualize it within the larger team cognition literature and to operationalize several relevant definitions. Perhaps the most basic, what exactly is a team? Salas, Dickinson, Converse, and Tannenbaum (1992) define a team as “a distinguishable set of two or more people who interact dynamically, interdependently, and adaptively toward a common goal/object/mission, who have each been assigned specific roles or functions to perform, and who have a limited life span of membership” (p.4). From this definition, it is clear that the accomplishment of the mission is predicated not only by the way in which the team interacts but also by who forms the team. What then is team cognition? The

Handbook of Applied Cognition defines team cognition as simply the “cognitive activity that occurs at a team level” (Cooke, Gorman, & Winner, 2007, p.240). To expand on the definition, team cognition includes: team decision-making, TSA, team perception (vigilance), and team knowledge, e.g. mental models (Cooke, Salas, Cannon-Bowers, & Stout, 2000). TSA is, of course, the team construct at the heart of this study.

As discussed above, there is still no wide spread agreement on the model of individual SA. Unfortunately, TSA is even more complex than individual SA (Salmon et al., 2008). A good starting point in an attempt to understand this topic is once again Endsley’s (1995) seminal paper on SA. In that paper, she expanded her model on individual SA to include TSA. She explains that TSA can be thought of as “the degree to which every team member possesses the SA required for his or her responsibilities” (Endsley, 1995, p. 39). While individual SA allows each team member to achieve their sub-goals, these sub-goals lead to an overall team goal (Endsley, 1995; Endsley & Jones, 1997). It is in pursuit of this team goal that TSA emerges. That is, some of the SA required by an individual in order to perform their tasks is the same amongst some or all of the other team members, this overlap is referred to a shared SA (Endsley & Jones, 1997), see Figure 4 below. Shared SA that is essential to all team members is referred to by Endsley and Jones (1997) as shared SA requirements. Endlsey and Jones (1997) are careful to point out that there are several possible states of shared SA. That is, two team members may possess the same SA but are both incorrect, they may possess the same SA that is correct, and they may also possess different understandings of the situation with one member being correct and the other not. Ultimately, “shared SA constitutes one important component of team SA, along with the level of SA possessed by individual team members” (Endsley et al., 2000, p.55).

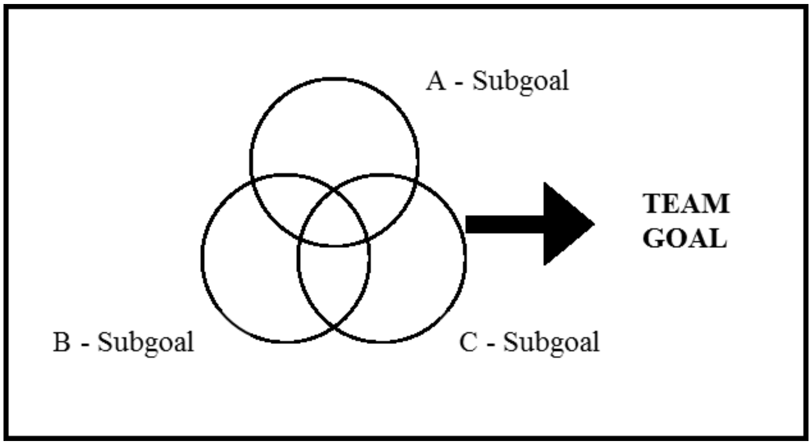


Figure 4. Team Situation Awareness (from Endsley & Jones, 1997). The circles represent the SA individuals require to complete their sub-goals. Where the circles overlap, it represents the portion of SA that must be shared amongst the team members. Together, team members pursue a common goal.

To solidify the concept of TSA, an example will be used from a firefighting context. In a dyad of firefighters entering a burning house, many of the SA requirements must be shared. For example, not only should the location of the fires be shared, but also the locations of other firefighters within the structure need to be known by everyone. Not knowing this could result in steam burns when water droplets are applied to a room that unknowingly contains other firefighters (Shane Causton, Ottawa Firefighter, personal communications, October 2012). In fact, not only does the location of other fighters need to be known by everyone in the structure but it is also vital for the firefighter whose job it is to track the location of all of the firefighters with the use of a white board set up near one of the fire engines. Not all SA, however, needs to be shared. For example, every firefighter must be aware of their individual emergency egress route at all times. This route is highly dynamic as it changes as the firefighter moves throughout the structure, and should a significant event occur (e.g. back draft), each firefighter must get out as fast as possible (Shane Causton, Ottawa Firefighter, personal communications, November 2012).

Like individual SA, several models compete with Endsley's (1995) model of TSA. One that has been recently proposed is Stanton et al.'s (2006) distributed situation awareness (DSA). Like the ergonomics approach in individual SA, DSA takes the approach that knowledge is distributed throughout a whole system, including human and non-human agents. Through their interaction, information is shared amongst all agents in the system which, in turn, serves to modify the information. Interestingly, they use Endsley's (1995) model on SA to explain that as a system, some agents are there to perceive information, others comprehend, while others project the information for future tasks. At first appearance, this appears to be quite at odds with Endsley's (1995) model. At the heart of this TSA debate is whether TSA "only exists within the individual" (Endsley et al., 2000, p. 53), which aggregates individual SA to measure TSA, or whether TSA is a phenomena in and of itself, owned by the entire system (Stanton et al., 2006) and, therefore, should be measured in a holistic manner. While it would appear from this debate that the two approaches to TSA are very far apart, it can be argued that they are in fact not much different. To argue that Endlsey's (1995) approach is individualistic, down plays those aspects of the theory that specifically integrates agent interactions. Specifically, Endsley and Jones (1997) discuss TSA devices (non-human agents or artifacts in Stanton et al.'s, 2006, terms), TSA mechanisms (e.g. shared mental models), and TSA process (e.g. planning and communication). The information solicited at the individual level is the product of a collective effort that results from the interaction between all of these factors and therefore can be considered more than the sum of SA that an individual could obtain on their own. The main difference between the two approaches appear to be whether SA devices (or non-human agents) can be considered a tool to assist humans in achieving SA or whether they can also have SA, which makes the debate metaphysical. For the purposes of this study, Endsley's (1995) model of TSA will be used primarily because it is the naturel extension of her popular individual SA model and because her individual SA model provides the theoretical underpinnings for five of the seven TSA models in a relatively recent review (Salmon et al., 2008). As the model continues to be debated so does the more practical issue of how to measure it.

Measuring Team Situational Awareness

Traditionally, the measurement of TSA has been obtained through the aggregation of individual SA, such as averaging individual team member's Situation Awareness Global Assessment Technique (SAGAT) scores (Cooke et al., 2007). This direct method assumes to tap directly into TSA as a state of knowledge at a particular time rather than processes used to achieve it (Endsley et al., 2001). More contemporary approaches attempt to measure TSA in a more holistic manner such as collective answers to queries (Cooke et al., 2004), documenting how teams adapt to unexpected roadblocks and realign themselves with team goals, i.e. Coordinated Awareness of Situations by Teams (CAST) (Gorman, Cooke, & Winner, 2006), and team communication as an indicator of TSA (Parush et al., 2011). For the purposes of this study, the traditional SAGAT method will be used for the following four reasons: (a) since the team consists of two firefighters with very similar tasks and SA requirements, aggregating their SAGAT scores appears reasonable in light of the large overlap of shared SA; (b) since this study is in a lab setting using a simulator that can be frozen, it lends itself to the SAGAT method; (c) it will be assumed that the benefits of team interaction are captured with the SAGAT score; and (d) employing the SAGAT method is much simpler and less resource intensive than attempting to use the CAST method or analyze communication. While allowing the teams to collaborate on

SAGAT queries is an option in this study, it is felt that this would not be an accurate reflection of TSA in this circumstance. In essence, by allowing the teams to collaborate on answers, it will cue the communicative processes that *should* have taken place during the simulation rather than accurately reflect the TSA at the point of freezing the simulator.

After having reviewed working memory, individual SA and TSA, the remaining sections will now examine previous studies with respect to the relationships between these constructs, as well as, the importance of TSA to team performance.

Examining the Construct Relationships

Team Situational Awareness and Team Performance

As can be seen from the 9/11 and friendly fire examples in the introduction, loss of TSA can have a devastating effect on team performance. The positive relationship between TSA and team performance has been studied both at the qualitative and quantitative levels. Qualitatively, Sonnenwald and Pierce (2000) examined "information behavior" in a synthetic military command and control environment. Collecting and coding data from a variety of sources including documents, observations and interviews, the researchers, using qualitative methods, discovered that interwoven situational awareness, among other themes, help determine success. They refer to interwoven situational awareness as a term used to describe the SA at the individual, intragroup and intergroup levels. In another qualitative study, Rafferty, Stanton, and Walker (2010) examined errors in teamwork with respect to fratricide based on a literature review and case study. Coding and analyzing the data from the literature, they found that team mental models, coordination, communication, cooperation, and TSA were all major factors contributing to fratricide.

There have also been several quantitative studies which have demonstrated that TSA can predict team performance. According to Cooke, Salas, Kiekel, and Bell (2004a), “team SA (situational awareness) is a consistently good predictor of team performance” (p. 28). Indeed, multiple experiments using a unmanned aerial vehicle (UAV) simulated environment, in which teams of three must navigate a UAV and capture aerial photographs, have consistently shown that TSA predicts team performance (Cooke & Kiekel, 2001; Cooke et al., 2004b). Interested in the influence of individual differences, including WMC, on team performance, Cooke et al. (2004b) also ran a secondary analysis to examine WMC and team performance that will be discussed in the next section. While there appears to be a positive relationship between TSA and team performance, there is no guarantee of it (Rafferty et al., 2010). As noted by Parush and Ma (2012) who found that while the presence of a team display facilitated both team performance and TSA, particularly during communication breakdown, no direct relationship was found between the two. They hypothesized that while team performance is associated with TSA, their relationship may be indirect as other factors may be at work. Indeed, in a recent study by Rustandjaja (2013) who examined TSA in a similar firefighting simulation as the present study, found no consistent link between TSA and team performance.

Working Memory Capacity and Team Performance

To the best of the author’s knowledge, Cook et al.’s (2004b) study mentioned above has been the only one that has examined WMC of the team members and the team’s performance. In their UAV studies, their primary research goal was to examine team performance, processes, and knowledge differences between distributed versus co-located teams. However, as a secondary analysis, they also examined individual differences, including WMC, on team performance. These teams consisted of three members; the air vehicle operator (AVO), who controls airspeed, heading, altitude, and monitors UAV systems, the payload operator (PLO), who adjusts camera settings, takes photos, and monitors the camera equipment, and the Data Exploitation, Mission Planning, and Communications Operator (DEMPC), who oversees the mission and determines flight paths. The results found that WMC scores associated with a specific role did predict team performance. Specifically, the AVO WMC predicted team performance as did the DEMPC for mission 5, the mission with a high workload. As Cooke et al. (2004b) explained, this finding makes sense as AVO and DEMPC must plan and navigate, requiring WM in order to remember where they are and where the team is going, whereas, the PLO job is less taxing on WM. In the present experiment, both participants will be completing the same firefighting task which is believed to heavily tax WM. It should be noted that Cooke et al. (2004) tried to predict team performance based on using the individual maximum score on a team as well as the range between the highest and lowest on the team but did not find significant results. While a team's score based on an individual's maximum score has been found to correlate with team performance in studies involving general mental ability, the strongest correlation has been found using the team's mean score (Devine & Philips, 2001) which was not used in Cooke et al.'s (2004) study. Overall, there appears to be mixed and uncertain results between WMC and team performance. Therefore, this study will further explore this relationship in order to better understand it, including team composition factors.

As a final examination of the constructs relevant to this study, the relationship between WMC and individual SA will be reviewed.

Working Memory and Individual Situational Awareness

While the relationship between WM and TSA has not been explored and are the primary constructs of concern in this study, the relationship between WM and individual SA has been studied in the past. By looking at this, it will provide evidence to make a hypothesis for the link between WMC and TSA.

As a reminder from the section on individual SA above, Endlsey and Jones (1997) theorized that in dynamic environments, “the development of situation awareness and the decision process are restricted by limited attention and working memory capacity for novices and those in novel situations” (p.17). From this, the hypothesis would be that those with high WMC would have better SA than those with lower WMC in novices and/or those in novel situations. When looking at specific experiments to support such a proposition, the first experiment that examined this relationship proved to be a little discouraging at first glance. Endsley and Bolstad’s (1994) work on measuring individual differences in SA among pilots revealed, in fact, no relationship between WM and SA. While surprising, a closer look at their memory measure shows that they used a test that only measured STM rather than WM. That is, they measured the short-term storage aspect of WM but not the processing aspect. Therefore, this was not a study measuring WM that is of interest here. Next, Carretta, Perry, and Ree (1996) conducted an experiment in which they wanted to determine if pilot SA could be predicted by several measures including cognitive ability, psychomotor ability, and personality. Results showed that, while holding flying experience constant, the partial correlations for verbal working memory, spatial working memory, and divided attention (among others), all significantly predicted SA. Further studies have supported the positive relationship between WMC and SA. Doane (2003) found that not only does WMC help identify individuals with adequate levels of SA but she also found that those novice pilots with high WMC were less likely to make certain error types. Furthermore, WMC has been shown to be more predictive of SA in novices than in more experienced individuals (Sohn & Doane, 2003, 2004; Gonzalez & Wimisberg, 2007). Such findings, are of course, exactly what one would expect knowing that WM becomes heavily engaged in novices or even more experienced individuals confronted with novel situations. Otherwise, if they had sufficient experience in the situation, they would rely on pre-formed schemas, or mental models residing within LTM that reduce the cognitive load on attention and WM (Endsley et al., 2000). Along similar lines, a recent and relevant study by Cass and Parush (in press) investigated the use of CogScreen – AE to predict situation awareness in Canadian Forces pilots. CogScreen – AE is a computer administered neuropsychological test used to detect changes or deficits in attention, WM, visual-perceptual functions and a variety of other cognitive abilities (Cass and Parush, in press; CogScreen website). In addition, Cass and Parush (in press) also wanted to see if inexperienced pilots who obtained high levels of SA would also score high on the CogScreen test, since they considered SA an ability, irrespective of training and experience. After splitting the pilots into high and low SA performance groups, results showed that the high SA group also scored high on the CogScreen – SA sub-tests symbol digit coding (SDC) and dual task tracking (DTT), which among other constructs, tap attention and WM. Furthermore, because one third of the inexperienced pilots scored high on SA performance and the SDC and STT sub-tests, they surmised that individual differences in attention, working memory and the other cognitive abilities measured by SDC and STT could be potentially used in personnel selection. It seems clear from all of these studies that WMC appears to have a positive relationship with SA, however, it remains to be seen in this study as to whether this can also be extended to TSA.

Summary

The next section is a summary of the literature presented above. Following this, a description of the urban firefighting domain for which this study takes place is given, why this study should be conducted, and the specific research questions it seeks to answer.

- WM is an information-processing construct that temporarily holds and processes information perceived from the environment or retrieved from LTM (Baddeley, 1987, 1997). It becomes heavily engaged in novel or complex situations (Endsley, 2000) and is responsible for the feeling of being overwhelmed by information that is too much or too complex (Cowen, 2005).
- WM is of limited capacity and individuals differ with respect to this capacity. Where and how the capacity occurs depends on the theory. In Baddeley's WM model, each sub-component has its own limiting capacity. In Cowan's model, how much can be in the focus of attention is the only limiting capacity, whereas Engle et al.'s model focuses on an individual's ability to control their attention. Complex span tasks, such as Daneman and Carpenter's (1980) reading span tasks are the norm for measuring the storage and processing ability of an individual's WM.
- SA at the individual level is comprised of three levels, perception of stimulus information from the environment, comprehension of the information, and projection of what this information means for the future. The attainment of SA relies heavily on WM for novices or for individuals in novel situations (Endsley, 2000). TSA emerges from individual SA and includes the sharing of overlapping SA, called SA requirements, to the right teammates, at the right time, in the right quantity (Endsley, 1995).

- Research has shown TSA helps predict team performance (Cooke, Salas, Kiekel, & Bell, 2004).
- Research on WMC and team performance is very limited. One study was found that linked the WMC of an individual and that person's role on a team to higher levels of team performance (Cooke et al., 2004).
- Research has shown that WMC can help predict SA at the individual level (e.g. Carretta, Perry, & Ree, 1996).

The Urban Firefighter Domain

Before moving to the goals and objectives of this study, it would be helpful to first put into context the domain of urban firefighting for which the participants will be immersed in. Urban firefighting, while not on the scale of the September 11 example in the introduction, is inherently dangerous, marked by limited or partial information, unforeseen hazards, and a fire that is very dynamic. Based on information obtained from two Ottawa firefighters (Shane Causton, October, 2012; Mark Messier, November, 2012), the following is a brief account of what an urban fire incident entails. A typical fire incident usually involves the firefighters getting notified at the fire station, at which time, as much information about the fire is passed to them, to include location, how dispatch was notified (911 call, witness) and whether it's a commercial or residential fire. Along the way to the incident, dispatch continues to update any information they have. Concurrently, the firefighters are assessing the situation along the route and as they arrive on scene. Their assessment can include the time of day (i.e. if at night they will likely head for the bedrooms), the wind direction, if smoke and flame are visible, the kind of construction, and what information any bystanders may have (i.e. if people are inside), to name a few examples. Based on this assessment, they will obtain the required tools, such as forced entry tools and hoses, and make their way into the structure. Generally, teams enter in three; two firefighters and an officer. They start with a primary search looking for and evacuating any victims found, with "attacking" the fire being secondary. Depending on the size of the fire, as more firefighters arrive there are multiple teams providing multiple roles, such as: the ventilation team, responsible for effectively removing smoke, a backup crew that will replace the crew inside, a rapid intervention team (RIT) that is responsible for going in and getting out firefighters that are in trouble, and ladder crews, who make sure there are second floor egress routes. Throughout, teamwork and communication leading to increased TSA is crucial. Within the structure, firefighters must continually update each other and the officer as to where they are, what they are doing, and notifying the others for any signs of victims, fire or immediate danger to themselves (i.e. back draft). Complicate this by having their field of view restricted by the respirator, smoke, and perspiration, as well as a limited oxygen supply for which they must continually be cognizant of and one can see the vital importance of teamwork, communication, and TSA. Ultimately, such a dynamic and dangerous environment as urban firefighting epitomizes extreme teams and makes it an ideal domain for reflecting extreme teams in general. Within this specific urban firefighting domain, a recent study by Rustandjaja (2013) examined the effects of building complexity and the amount of information given to the participants before commencing the firefighting simulation on TSA. In addition, Rustandjaja also looked at workload shift in which a fire obstruction was introduced within the scenario. Surprisingly, results showed that team performance was actually better in the building which was initially thought to be more complex as calculated by its inter connection density (ICD), or more simply, the number of opportunities to make a decision to go in a different direction. A closer examination revealed that the building which was thought to be less complex may have actually been more complex in that it contained smaller rooms and was more “maze like”. Additional findings revealed that less information given to the participants before the simulation started was related to lower SA, while performance before the workload shift was less than afterwards. Finally, as mentioned in the TSA and team performance section above, no significant link was made between TSA and team performance.

Goals and Objectives

This study has both theoretical and practical motivations. On the theoretical side, this study seeks to break new ground by examining the relationship between WMC of team members and TSA. Concurrently, it will also expand the current understanding of the effect of WMC on a team’s performance, something that has received minimal attention in the literature. Furthermore, this study is expected to provide further evidence of the positive relationship between TSA and team performance. On the macro-level, Wildman et al.’s (2012) literature review on team knowledge identified dynamic team knowledge as the number one research need. They explain that “more than 87% of the team knowledge research to date has examined mental models, transactive memory, and strategic consensus, with only 13% of the research examining situation awareness and other constructs” (Wildman et al., 2012, p. 102). This study is expected to fulfill this research need.

On the practical side, there are several reasons for running this experiment. If a relationship can be found between WMC and TSA, as well as WMC and team performance, organizations can use this information as a factor when considering team composition, especially in any environment that requires TSA. Furthermore, despite the fact that information-processing tests have been shown to predict job performance on complex tasks like flying an aircraft, monitoring and maintenance tasks, and air traffic control, overall, its use in personnel selection has been limited (Arthur et al., 2004). Since individuals vary with respect to information-processing ability, the results from this study can also support the use of WMC for personnel selection purposes.

A Caveat

It is understood that in any of these extreme team environments, training, which forms mental models, play a significant role. While this training provides the best chance of survival and success in these environments, information flow can still easily overwhelm even the most experienced members. Indeed, the enemy, whether it is an enemy soldier, a house fire, or a criminal, all have a say in how the situation unfolds. Therefore, this study is not undermining the importance of training but rather stresses the critical importance of WM in novices with minimal training or even experienced members in novel or highly complex situations.

Research Questions

1. Do extreme differences between team member's WMC affect their TSA or team performance?
2. What is the relationship between the WMC of team members and TSA?
3. What is the relationship between the WMC of team members and team performance?
4. What is the relationship between TSA and team performance?
5. What affect will the difficulty level of the scenario have on the relationship between WMC of team members and TSA?
6. What affect will the difficulty level of the scenario have on the relationship between WMC of team members and TSA?
7. What affect will the difficulty level of the scenario have on the relationship between TSA and team performance?

Hypothesis

1. Extreme differences between team member's WMC will not affect TSA or team performance.
2. Teams high in WMC will demonstrate better TSA than teams low in WMC.
3. Teams high in WMC will demonstrate better performance than teams low in WMC.
4. Teams high in TSA will demonstrate better performance than teams low in TSA.
5. Teams high in WMC will be less affected by the difficulty of a scenario on their TSA than teams low in WMC.
6. Teams high in WMC will be less affected by the difficulty of a scenario on their performance than teams low in WMC.
7. The more difficult the scenario, the greater the detrimental effect will be on both TSA and team performance.

Methods

Participants

A total of 40 teams, consisting of two participants each (80 participants total), were recruited using the Carleton Universities SONA system. After data cleaning, six participants, and subsequently their team, were removed leaving 34 teams (68 participants) for the analysis. Age of the remaining participants ranged from 17 to 33 with a mean age of 20.12. Furthermore, 37 were female and 31 were male. Of the 34 teams, only one team had teammates that were well acquainted, two teams had teammates that knew each other but were not well acquainted, with the remainder of the teams having teammates that had never met before. Participants were all undergraduate students taking the first year psychology or neuroscience class where they each received a 2% credit towards their final mark for their participation.

Environmental Tasks and Design

Experimental Tasks

This experiment took place within the firefighting domain; however, the intent is for the results to reflect any extreme team environment that is dynamic, dangerous, and taxes WM. The main reason the firefighting domain was chosen in particular was due to the availability of a relatively cost effective simulator that is both realistic and could be manipulated to fit the experiment's requirements. The simulators' task environment involved a fire truck arriving on the scene of a dwelling fire, at which time two firefighters of equal rank controlled by the participants, got out of the fire truck and set about accomplishing their tasks. The main tasks of these firefighters were to locate and rescue victims, extinguish fires, and to turn off the utilities (gas, water, and electricity).

Design

This study is a one-way within-participants design. The independent variable, scenario difficulty, was manipulated in order to create more extreme and complex situations to better test the role of WMC. The order of teams through the scenarios was counterbalanced to control for order effects. The scenarios differed in the size of the house, the number of victims, the number of fires and the degree to which smoke hindered visibility. Specifically, the three factor levels were:

1. The "easy" scenario: A single story range house with no garage or basement, with two victims to rescue and four fires to put out, smoke was not a visibility issue.

2. The "medium" scenario: A two story house with a garage but no basement. There were four victims and five fires and, at times, smoke became a visibility issue.

3. The “hard” scenario: A two story house with a garage and basement. There were six fires and seven victims to rescue. Smoke was a greater visibility issue than the other scenarios.

Scenario configurations and screen captures can be found in Appendix A.

Setting and Apparatus

The experiment took place in a lab located in the Human Computer Interaction (HCI) building, room 3113. The lab had no windows and was therefore illuminated through artificial light alone. Both the WMC test and the firefighting simulator were loaded onto two Gateway DX4850 computers with Intel (R) Core processors and 8.6 GB RAM. For the WMC test, computers were placed on two different desks facing away from each other in order to reduce distractions to the participants. Once the WMC tests were complete, the computer set up was reconfigured so that they were beside each other for the firefighting simulation, see Figure 5.

Figure 5. Configuration of the two computers for the firefighting simulation.

Participants used a standard Xbox controller to manipulate their avatar firefighters. The monitors were angled away from each other in order to prevent participants from easily seeing one another's screen; however, no dividers between participants were used due to concerns that this would have had a negative effect on the communication between teammates.

The computer software used in this experiment was called Flam-Sim, an off-the-shelf product designed to train professional firefighters in a high fidelity 3-D first-person simulator. Flame-Sim was installed on both computers which were connected through a local area network (LAN). Either computer could be used to host multi-user sessions while the other computer merely had to join the session. While the primary purpose of the software is for training purposes, the high quality graphics, the ability to create scenarios, and the functionality of automatically capturing performance and timeline data in an after action report (AAR), makes it a valuable tool for experimental purposes. Specifically, Flame-Sim had teams maneuver throughout a very realistic and dynamic simulated environment, utilizing equipment such as respirators, hoses and entry tools, to open doors, fight fires and extract victims while having to deal with the realities of reduced visibility caused by smoke. See Figure 6 for a screen capture of the simulated environment.

Figure 6. This screen capture highlights the realistic first person view of Flam-Sim. A participant's teammate can be seen attacking a fire in the kitchen.

Furthermore, Flame-Sim allowed the experimenter to edit scenarios by being able to select: the dwelling type from 12 choices (e.g. farm house, apartment complex, bungalow, etc…); the construction type, which determines if and how the fires spread through the structure; the fire locations and ignition start times; whether doors or windows were open or closed, which influenced smoke distribution; and the victim locations and health status (e.g. healthy, injured, incapacitated); see Figure 7. Once the scenarios were created, they were then saved and uploaded for each team to ensure uniformity.

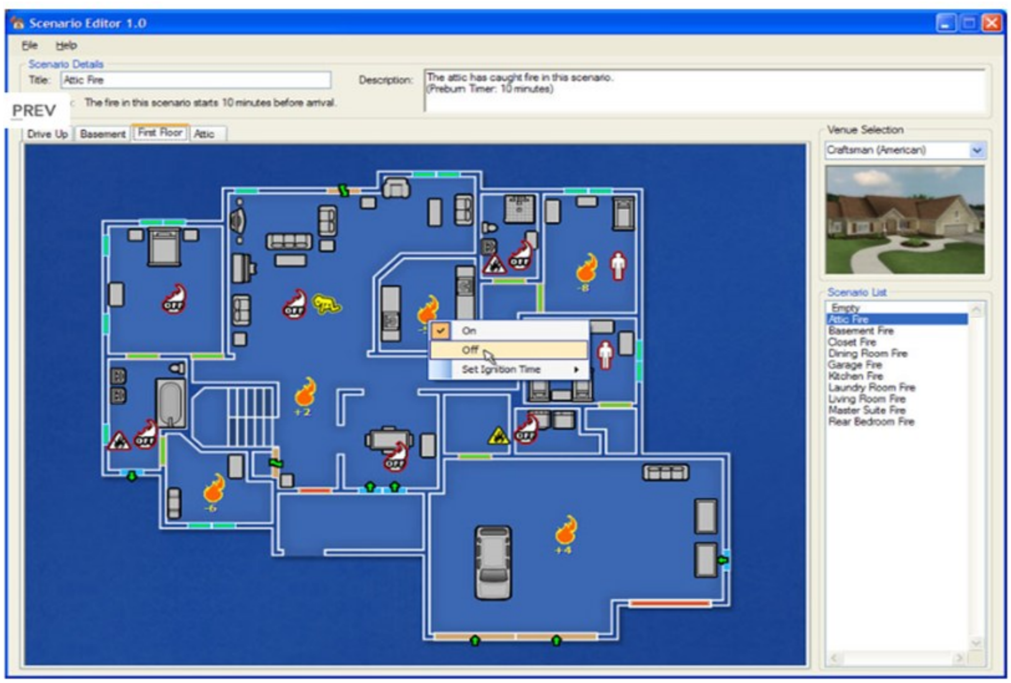


Figure 7. The scenario editor allowed the experimenter to manipulate several features including; the type of structure and construction type, when and where fires ignite, where the victims are and their health status, and whether the windows are doors are open or closed. (Obtained from the www.Flame-Sim.com website)

Once each scenario was complete, the Flame-Sim software automatically captured valuable data in an after action report (AAR). This report provided information on virtually all of the firefighters' actions to include when and where they searched, when and by whom a victim was rescued, when and by whom a fire was attacked and extinguished, as well as, the location of the fires. Finally, a Sony Camcorder was positioned behind the participants to capture each scenario for later play back in order to score the TSA measure.

Procedure

After being recruited through the Carleton Universities' SONA system, participants arrived at the arranged date and time. Upon arrival, participants were greeted and given Carleton University's standardized consent form which briefly outlined the purpose of the study, the environment and tasks, as well the risks and benefits of the study, see Appendix B. After answering any questions, a demographic questionnaire was administered that has been used in previous studies within the lab (see Chunyun, 2012) but modified to suit the present experiment, see Appendix C. This questionnaire captured basic demographic information such as sex, age, first language, familiarity with their teammate for the study, as well as information regarding the participant's video game experience. It took no longer than 10 minutes to read and complete both the consent form and demographic questionnaire, at which time it was collected by the experimenter. The experimenter then asked if there were any questions and reminded them that if they did not feel comfortable for any reasons they could withdraw from the experiment at any time without penalty. Once the preliminary administration was completed testing commenced.

The first measure that was obtained was each participant's WMC. This was accomplished by running the participants through the automated complex span tasks located on the same two computers used for the simulation. As both participants ran through the WMC at the same time, these computers were separated and placed on different desks facing away from each other in order to minimize distraction. Completion time varied between participants, taking on average between 20-30 minutes. Once compete and before commencing the firefighting training program, participants were given a five minute break which also allowed the experimenter to reconfigure the computers so that they were beside each other.

Before the training session began, participants were handed training instructions (Appendix D) and a handout with the Xbox controller functions (Appendix E). Once given a chance to read them, they were told to press start which brought them to Flame-Sims' built-in training environment. Situated in a virtual warehouse, the experimenter walked each team through the training instructions step-by-step in order to ensure that each team was trained in the exact same manner. Training consisted of learning and practicing basic controller functions to include walking, running, and crawling. Participants were also shown how to retrieve and use three basic tools from the fire truck necessary for the accomplishment of the mission. They included: (a) a forced entry tool, needed to smash open locked doors; (b) the fire hose, necessary to extinguish fires; and (c) the gas wrench, necessary to shut off the gas to the house. Once the formal step-by-step training was complete, teams were given five minutes to further practice what they had learnt. During this time, each team also practiced pausing the simulation and turning of the monitor in preparation for the TSA queries. By the end of the training session, participants had gained the necessary understanding of the basic controller functions without having to actually run through a scenario session. In addition to the training session, Flame-Sim also cued the participants throughout each session as to what controller button to push when confronted with a door, a victim (e.g. "push X to pick up victim), or a host of other actions. Once training was complete, participants were given another five minute break at which time the first scenario was loaded on the computers by the experimenter.

Participants ran through three scenarios varying in difficulty level in a counterbalanced order. In order to obtain a measure of TSA, Endsley's (1988) situation awareness global assessment tool (SAGAT) was used. Participants were told to "freeze" the simulator twice during each scenario by pressing the start button at the 3 and 6 minute mark. Once the simulator was paused, participants turned off the monitor and were then given a sheet containing random SA queries taken from the query bank found in Appendix G. Randomization was achieved using the Excel spreadsheet randomization function. Each scenario was considered complete either when the team had successfully rescued each victim, put out each fire, turned off the three utilities and returned to the fire truck or when eight minutes had elapsed, whichever came first. After each scenario, participants were given a five minute break while the next scenario was being loaded. Once all of the scenarios were complete, a post session questionnaire was given to the participants (Appendix H). This questionnaire obtained information regarding the difficulty the participants experienced at the task at hand, as well as way-finding difficulty (difficulties experienced throughout the scenarios understanding where they had been, where they were, and where they had to go). The questionnaire took no more than five minutes to complete. Upon completion of the experiment, the participants were given the standardized Carleton University debriefing form (Appendix I) to read and take with them explaining the background and purpose of the study in more detail, where they could seek more information, contact numbers if they were feeling distressed and the contact numbers of the researchers should they have future questions.

Measures

Working Memory Capacity

Recently, complex span tasks have been automated (i.e. computerized), validated and, compared to the traditional measures, are quick to administer, mouse driven and automatically scored (Redick et al., 2012). As it appears from the discussion in the introduction on WMC measures, any validated and reliable complex span task is domain general and therefore could have been used here. The automated RSpan task was selected; however, the OSpan or any other complex span task could have been used. The automated RSpan, based on Daneman and Carpenter's (1980) test and discussed in Redick et al. (2012), was purchased and downloaded from http://www.millisecond.com/. The script itself was created by Engle (2005) and got the participants to read a sentence and, unlike the original RSpan task, decide whether the sentence made sense or not by clicking the mouse on true or false. After the decision, a letter was presented for which the participant had to remember. Like the RSpan, the number of sentences and letters to be remembered varied from three to seven. After all of the sentences and letters were presented, the participant had to then correctly select the letters in order of presentation. Two RSpan scores were obtained; one was based on the absolute total number of letters recalled from correct sets. That is, if the participant recalled five of five letters in the correct order on one set, they would be given five points for that set, conversely, if they got four of five letters correct they get zero points. Points were then summed across correct sets. Alternatively, the partial RSpan score was given for every correct letter in sequence irrespective of whether a letter was missed. In other words, participants got part marks for partially correct sets. For the purpose of this study, the partial RSpan score was used as research has indicated that the psychometric properties of partial-scoring is better since the absolute RSpan score discards useful information on an individual's performance (Redick et al., 2012). A team mean WMC score was obtained by averaging the teammate's WMC score.

Team Situational Awareness

As discussed, Endsley's (1988) SAGAT was used. As Endsley (2000) points out, one of the most important issues is developing the queries. A list of SA queries were created which aimed to assess the SA requirements, that is, the dynamic information needs associated with the accomplishment of major goals or sub-goals (Endsley, 2000). These SA requirements were derived from two separate hierarchical task analyses (hTA). One hTA was completed for this study with the help of Shane Causton (October, 2012) a firefighter from the Ottawa fire Services (see Appendix J). This hTA served to augment another hTA already completed within the firefighter domain by Nadya Rustandjaja, another Carleton student, in 2011. In addition, to help structure the questions in accordance with the three SA levels and to differentiate between individual/generic queries and those aimed at teams, Parush's (personal communication, November, 2012) SA generator template was used (see Appendix K). Taken from the SA generator, an SA query bank was created for ease of randomization (Appendix G). Two queries were taken from the individual/generic queries at each of the three levels (i.e. perception, comprehension, and projection) and two were taken from the team queries at each of the three levels. The lines depicted in the queries reflect information such as location or quantity that were filled in by the experimenter based on the scenario. A total of 4 (queries) x 3 (levels) x 2 (Freezes) x 3 (scenarios) for a total of 72 queries per participant were asked. Any queries for a scenario that would elicit an identical response to a previous query (i.e. "the front door was locked") were discarded and a random replacement was drawn. Each query was scored as being correct or incorrect based on participants answering true, false, or unknown for level 1 and level 2 TSA queries, and true or false for Level 3 TSA. Unknown was not an option for Level 3 TSA since it is a future projection of events that have not happened, therefore, if unknown was an option, this could be selected and the participants would always be right. Participant answers were then evaluated against the video footage, after action report (AAR) and what was pre-determined by the experimenter in order to determine its true accuracy. For example, an SA query may ask, "a victim was found in the upstairs bathroom". If a victim was found in the upstairs bathroom, their answer should be true. However, it may be likely that due to the dynamic nature of the simulator, the participants may not have seen this victim yet. Therefore, unknown would be the correct answer. Incorrect scores were also given not only to incorrect answers but also to correct answers that have been guessed based on the review of the video footage and AAR and knowing whether the participant had been exposed to the stimulus or not. For example, if the participant accurately responded true to a victim being in an upstairs bathroom but the AAR and video footage show that the room had not been searched, this would be classified as incorrect. This method is loosely based on signal detection theory (SDT) in which a participant either accurately responds to a stimulus that is present (hit), accurately responds that the stimulus is not present (correct rejection), inaccurately responds that the stimulus is present (false alarm) or inaccurately responds that the stimulus not present (miss). Finally, in the same manner as Salmon et al. (2009), SAGAT scores were calculated by summing across all correct responses. These summed individual SAGAT scores were then summed again for the team's TSA score.

Team Performance

The primary team performance score was obtained for each scenario by aggregating the total number of victims saved, the total number of fires extinguished and the total number of utilities turned off. This score was then divided by the total number of possible victims saved, fires put out and utilities turned off, in order to get a total proportional, or relative, performance score that could be compared across scenarios.

In addition to the team performance measure above, two team performance measures based on way-finding were also obtained. The first measure was simply based on the total number of rooms searched divided by the total number of rooms, which gave a relative search performance score. The second measure was obtained by determining the optimum search sequence and comparing it to the search pattern of the teams automatically collected in the AAR. According to Ottawa firefighter Shane Causton (personnel communication, October, 2012), unless it is night time, in which firefighters go straight to the bedrooms, they general perform a left or right hand search technique running along the walls room to room. The optimal searches during these day time scenarios was the sequential search of rooms going in a clockwise or counter clockwise fashion, e.g. A-B-C-D-E-F or vice versa, for each floor (see Figure 8 for an example).

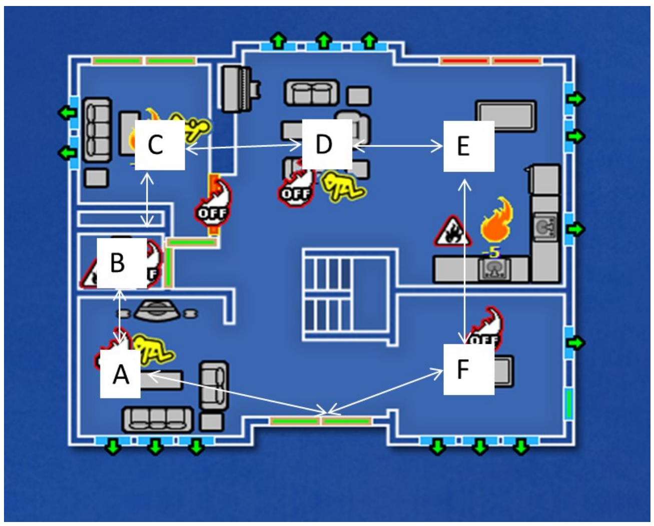


Figure 8. An example of an optimal search pattern.

An optimum score was obtained based on the number of pairings, for the example, a total of six pairings is present (AB, BC, CD, DE, EF). The team's actual sequence was then compared, for example, a team's sequence may have been C-A-B-E-F-D. In this example, the team only has two optimal pairings, AB and EF, therefore they score a two. A total score across floors was obtained for each scenario and then divided by the optimum score, giving a relative optimal search performance score.

Video Game Experience

The video game covariate measure was obtained from the participant's questionnaire. Three questions were asked pertaining to the participant's experience with computer games and video game console (i.e. Xbox). Specifically, "how many hours a week do you play video games on a computer" and "how many hours a week do you play video games on a console (i.e. Xbox, Wii, etc…), and "how many hours a week do you play games on your phone". These scores were combined to measure video game experience.

Demographic and Post-Test Questionnaire

In addition to the video game experience, a demographic questionnaire was used to obtain information pertaining to age, sex, first language and how well the participants were acquainted with each other.

Results

The results outlined below have been organized based on the three relationships at the core if this study; working memory capacity and team situational awareness, working memory capacity and team performance, and team situational awareness and team performance. Data were initially checked for valid RSpan scores, outliers, and normality. First, the data from the RSpan task were examined for participants who failed to achieve the 85% accuracy score in the processing component of the task. Due to possible trade-offs between the processing and storage tasks (i.e. the participant focuses only on remembering the letters to recall rather than reading the sentences as well), Conway et al. (2005) explained that the processing component of the task needs to be close to perfect, typically 85%, otherwise the entire data set for the participant is to be discarded. Indeed, as Engle (personal communication, May, 2013) further explained, having the processing cut-off score quite high is necessary for the test to be a valid and reliable measure.

Of the 80 participants, four participants did not achieve this cut-off and were excluded from the analysis, as was the team they belonged too. Next, the data were examined for outliers. From the remaining 76 participants, participant #8 was an outlier on their RSpan score; they obtained a raw score of 5 out of a possible 75, which was 3.45 SD below the mean. This score, in all likelihood, indicated that the participant put little effort into the task and was thus excluded, along with their team, from the analysis. Finally, participant #80 had several outliers including their score for their Level 2 TSA, 3.04 SD below the mean, their overall TSA score, 3.09 SD below the mean and that at age 62, she was 7.47 SD above the mean. Since the sample in this study is meant to reflect the cognitive abilities of potential members of extreme teams, a 62 year old did not match this profile, thus, this participant and their team was also excluded. With a total of six participants excluded from the analysis, and subsequently their team, 34 teams remained. In addition to identifying outliers, all variables used in the analysis below were checked for normality, none of which violated this assumption. To note, each pairwise comparison presented below was Bonferroni adjusted.

Working Memory Capacity and Team Situational Awareness

The following sub-section examines WMC and TSA, specifically; it first provides an analysis of team composition, followed by an explanation on how the sample was eventually divided into sub-groups based on the team's mean WMC scores. It then provides the results for the main analysis, which includes: examining the relationships between WMC and TSA, WMC and team specific queries, and finally, WMC and when teammates were both correct. One will recall that TSA is comprised of all SA questions including both individual level (e.g. "I have extinguished 2 fires") and team level queries (e.g. "as a team we have searched 4 rooms"). The team specific queries score is comprised of just those SA questions which pertain to one's teammate and the team in general, and the "both correct" score is the aggregate of when both teammates get an SA query correct, thus indicating if they had correct overlapping SA.

Team Composition with Respect to Working Memory Capacity. In order to assess whether within-team WMC differences had any relation to TSA or team performance indices, the following analysis was performed. First, a WMC difference score was obtained for each team by subtracting the lowest WMC score on the team from the highest. Based on this WMC difference score, the median was used to divide the sample into two groups, those teams with large within-team WMC differences and those with small within-team WMC differences. Second, the mean WMC score for each team was computed and, once again, the median was used to divide the sample into two groups, those with a high team mean WMC scores and those with a low team mean WMC scores. Third, to be sure that the distribution of the WMC difference scores within the sample would not affect results (i.e. the difference scores were not associated with the mean WMC score of the teams), a 2 x 2 Chi-Square analysis was performed using the frequencies in which teams fell into the following categories: (a) teams with high team mean WMC scores and large WMC differences; (b) teams with high team mean WMC scores and small WMC differences; (c) teams with low team mean WMC scores and large WMC differences; and (d) teams with low team mean WMC scores and small WMC differences (see Table 1). Results were not significant, $p > .05$, thus indicating that the WMC differences scores were evenly distributed amongst the teams high and low in WMC.

Table 1

Frequency of WMC Difference Scores within Team Mean WMC Scores

	WMC Difference Score		
Team Mean WMC Score	Small	Large	Total
Low	8	8	16
High	9	9	18
Total	17	17	34

Lastly, knowing the WMC difference scores were not associated with the mean WMC scores of the teams in the sample, the following mixed ANOVAs were performed in order to determine whether differences between the team member's WMC score, irrespective of whether that team had a high or low average WMC score, were associated with the team's TSA score or any of the team performance indices.

1. 2 (WMC Diff) x 3 (Scenarios) x 3 (TSA Levels) mixed ANOVA, with TSA as the dependent variable.
2. 2 (WMC Diff) x 3 (Scenarios) x 3 (TSA Levels) mixed ANOVA, with Team Queries as the dependent variable.
3. 2 (WMC Diff) x 3 (Scenarios) x 3 (TSA Levels) mixed ANOVA, when both teammates are correct as the dependent variable.
4. 2 (WMC Diff) x 3 (Scenarios) x 3 (Performance Type) mixed ANOVA, with Relative Performance as the dependent variable.
5. 2 (WMC Diff) x 3 (Scenarios) mixed ANOVA, with Relative Search Performance as the dependent variable.
6. 2 (WMC Diff) x 3 (Scenarios) mixed ANOVA, with Relative Optimal Search Performance as the dependent variable.

In all six analyses, there were no significant results involving WMC differences, $p > .05$.

In addition to the difference score analysis performed above, an analysis was performed to compare groups in which both team members were above the samples individual WMC median score ($n = 7$), both below the median ($n = 7$), and when one was well above the median and one well below ($n = 7$). Once again, using the 6 analysis outlined above, with the between group variable now WMC (high-high, low-low, and high-low), results did not show a significant difference between the WMC high-low group and the other groups, $p > .05$. See Table 2 for the between-group TSA results and Table 3 for the between-group performance results.

Table 2

Descriptives of WMC (high-high, high-low, low-low) and TSA

	TSA		Team Queries		Both Correct	
WMC	*M (SD)*	95% CI	*M (SD)*	95% CI	*M (SD)*	95% CI
Low-Low	9.52 (.46)	[8.55, 10.50]	4.62 (.25)	[4.10, 5.14]	3.05 (.27)	[2.48, 3.62]
High-Low	9.59 (.46)	[8.62, 10.56]	4.43 (.25)	[3.91, 4.95]	3.25 (.27)	[2.68, 3.83]
High-High	10.91 (.46)	[9.91, 11.88]	5.30 (.25)	[4.78, 5.82]	4.11 (.27)	[3.54, 4.68]

Note. CI = confidence interval.

Table 3

Descriptives of WMC (high-high, high-low, low-low) and Team Performance

	Relative Performance		Relative Search Performance		Relative Optimal Search Performance	
WMC	*M (SD)*	95% CI	*M (SD)*	95% CI	*M (SD)*	95% CI
Low-Low	.668 (.05)	[.558, .777]	.813 (.04)	[.772, .905]	.354 (.05)	[.259, .449]
High-Low	.671 (.05)	[.562, .781]	.753 (.04)	[.661, .844]	.356 (.05)	[.261, .451]
High-High	.716 (.05)	[.606, .825]	.775 (.04)	[.683, .866]	.317 (.05)	[.222, .412]

Note. CI = confidence interval.

While caution must be used in this last analysis due to the relatively small group sizes and resulting small power, it does, along with the WMC differences analysis, indicate that within-team WMC differences do not appear to have an effect. Knowing this, the analysis below focuses on the division of the sample into sub-groups based on the team's average WMC.

Dividing the Overall Sample into Sub-Groups Based on Working Memory Capacity. While the division of the sample into two WMC sub-groups above was necessary in order to achieve enough of a cell count for the Chi-Square (greater than 5) analysis, the remainder of the analysis used a division based on three WMC sub-groups, those teams with a low WMC, medium WMC, and high WMC, see Table 4. The decision to divide the sample into three sub-groups based on the teams' average WMC was ultimately made as it was thought to be less arbitrary than using the median split and still utilized the whole sample rather than splitting it into quartiles (n = 9 in each group) or when comparing teams with both teammates above the median against teams with both below the median (n = 7 in each group). All of these analyses were still performed and can be found in Appendix L; however, they all showed similar results.

Table 4

WMC Sub-Group (Thirds) Descriptives

WMC Sub-Group	n	Mean	Median
Low	12	43.83	45.50
Median	11	52.63	53.00
High	11	58.59	57.00

Team Situational Awareness Queries. In order to assess the relationship between WMC and TSA, a 3 (WMC) x 3 (Scenarios) x 3 (TSA Levels) mixed ANOVA with TSA as the dependent variable, was performed. Results revealed a significant main effect between the WMC groups, $F(2, 31) = 7.679$, $p = .002$, $\eta_p^2 = .331$. Pairwise comparisons further identified that it was the high WMC group that had a significantly higher TSA ($M = 10.919$, $SE = .319$) than the low WMC group ($M = 9.213$, $SE = .305$), p = .002. Teams in the medium WMC group ($M = 10.273$, $SE = .319$) did not show a significant difference between the high or low WMC groups.

A main effect of scenario, $F(2, 62) = 5.897$, $p = .005$, $\eta_p^2 = .160$, was also significant, with the easier Ranch scenario having a significantly higher TSA score ($M = 10.553$, $SE = .279$) than the hardest scenario, California ($M = 9.184$, $SE = .246$), p = .008. The medium difficulty scenario, Colonial ($M = 10.315$, $SE = .242$), was not significantly different than the other two scenarios.

The analysis also revealed a main effect of TSA levels, $F(2,62) = 18.693$, $p = .000$, $\eta_p^2 = .376$. A pairwise comparisons showed that level 1 TSA, perception, was significantly higher than level 2 TSA, comprehension, and level 3 TSA, future projection, both p = .000. There was not a significant difference between level 2 TSA and levels 3.

Finally, a significant interaction between scenarios and TSA levels was found, $F(4,124) = 2.587$, $p = .040$, $\eta_p^2 = .077$. Figure 9 shows that a significant decline in level 3 TSA, projection of future status, occurs in the California scenario. In other words, in the hardest scenario, understanding what will happen becomes much more difficult than the easier scenarios.

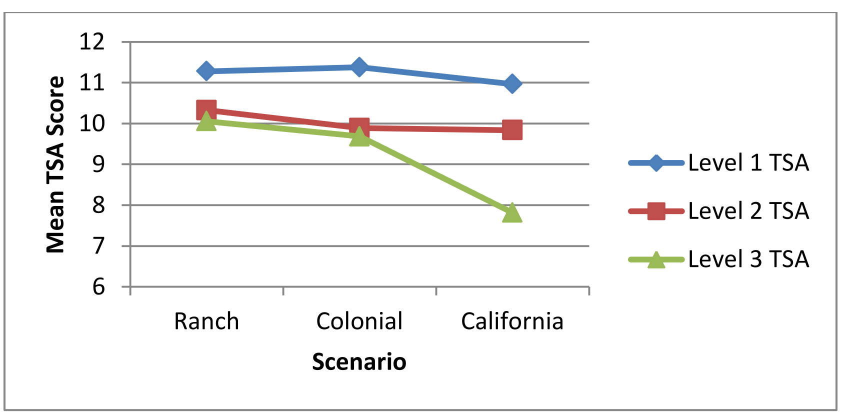


Figure 9. Total TSA score as a function of Scenario and TSA Level.

Team Specific Queries. In order to evaluate the relationship between WMC and team specific SA queries, a 3 (WMC) x 3 (Scenarios) x 3 (TSA Levels) mixed ANOVA with team queries as the dependent variable, was performed. Results revealed a main effect between the WMC groups, $F(2,31) = 8.312$, $p = .001$, $\eta_p^2 = .349$. Team's high in WMC had a significantly higher score on the team questions ($M = 5.364$, $SE = .176$) than teams low in WMC ($M = 4.380$, $SE = .168$), p = .001. The medium WMC group ($M = 4.747$, $SE = .176$) was not significantly different than either the high or the low WMC group.

A main effect of TSA levels, $F(2,62) = 11.419$, $p = .000$, $\eta_p^2 = .269$, was also found with team specific queries. Level 1 TSA ($M = 5.330$, $SE = .156$) was once again significantly higher than Levels 2 ($M = 4.864$, $SE = .173$) and Level 3 TSA ($M = 4.297$, $SE = .150$), both p = .000.

Finally, a significant interaction between scenarios and TSA levels was discovered, $F(4,124) = 2.705$, $p = .033$, $\eta_p^2 = .08$. Once again, a significant decline occurs for level 3 TSA in the California scenario, indicating that projecting team specific information into future status becomes much more difficult in the hardest scenario, see Figure 10.

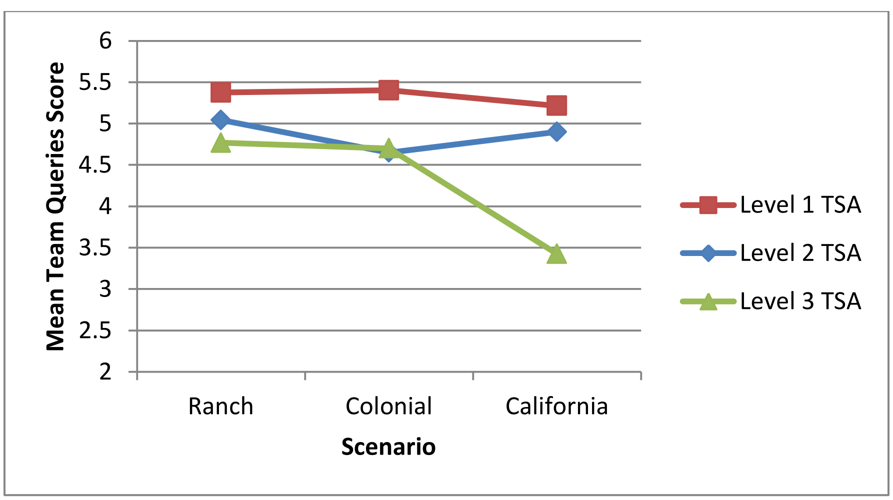


Figure 10. Total team queries score as a function of scenario and TSA level.

Both Correct. Using a score based on when both participants got the same TSA queries correct, a 3 (WMC) x 3 (Scenarios) x 3 (TSA Levels) mixed ANOVA was performed in order to assess the relationship between WMC and when both team members were correct. Once again, results revealed a significant main effect between WMC groups, $F(2,31) = 8.990$, $p = .001$, $\eta_p^2 = .367$. Pairwise comparisons showed that the low WMC group had significantly fewer queries in which both team members were correct ($M = 2.852$, $SE = .200$), than the medium WMC group ($M = 3.707$, $SE = .209$), p = .017, or the high WMC group, ($M = 4.030$, $SE = .209$), p = .001.

A main effect of scenario was observed, $F(2,62) = 7.535$, $p = .001$, $\eta_p^2 = .196$, with a pairwise comparison indicating a significant difference between the Ranch Scenario ($M = 3.896$, $SE = .193$) and the California Scenario ($M = 3.038$, $SE = .157$), p = .003. Finally, a main effect of TSA levels was revealed, $F(2,62) = 18.565$, $p = .000$, $\eta_p^2 = .375$, with significant differences between Level 1 TSA ($M = 4.294$, $SE = .175$) and Levels 2 ($M = 3.340$, $SE = .184$), and Level 3 TSA ($M = 2.955$, $SE = .170$), both p = .000.

Working Memory Capacity and Team Performance

The following analysis uses three different team performance indices; relative performance, relative search performance, and relative optimal search performance. One will recall that a teams' relative performance score is the total number of victims saved, fires extinguished, and utilities turned off by the team, divided by the total number of victims, fires, and utilities present. Furthermore, relative search performance is the number of rooms searched divided by the total number of rooms. Finally, the relative optimal search performance score was obtained by first calculating the optimal search score of each floor based on a clockwise or counter clockwise search pattern (e.g. rooms searched in the order A-B-C-D-E-F), where each sequential pairings received a point. For the example above, there is a total of five possible pairings (AB, BC, CD, DE, EF), thus, the total optimal search score for the floor is five, see Figure 8 above for an example. Next, each team's actual sequence was obtained based on the after action report, for example, a team's sequence may have been C-A-B-E-F-D. In this example, the team only has two pairings, AB and EF, therefore they scored a two. A teams' relative optimal search performance score was then obtained by adding the teams' score across floors and then dividing it by the total optimal search score. Once again, each pairwise comparison presented below has been Bonferroni adjusted. It should be noted that the covariate, video game experience, often correlated with the performance indices of the Ranch and Colonial conditions but did not correlate with the California dependent variables, indicating that the homogeneity of regression slopes assumption for ANCOVAs was violated. As a result, video game experience was not used in the analyses.

Working Memory Capacity and Relative Performance. In order to determine if a relationship exists between WMC and relative performance, a 3 (WMC) x 3 (Scenarios) x 3 (Performance Type) mixed ANOVA, with relative performance (of each type) as the dependent variable, was performed. It should be noted that the homogeneity of variance assumption was violated for the Ranch victim, Colonial victim, and Colonial fire performance types, not surprising considering the near perfect performance on these measures by most teams. No corrections were made to address this violation due to the robust nature of the F test to this violation when group sizes are proximately equal (Kutner, Nachtsheim, Neter, & Li, 2005). Results showed a significant main effect of Scenario, $F(2,62) = 33.059$, $p = .000$, $\eta_p^2 = .516$, with teams performing significantly worse in the California scenario ($M = .563$, $SE = .025$), than the Ranch scenario ($M = .771$, $SE = .024$) or the Colonial scenario ($M = .754$, $SE = .035$), both pairwise p = .000.

A main effect of Performance Type was also revealed, $F(2,62) = 103.641$, $p = .000$, $\eta_p^2 = .77$. The proportion of victims saved ($M = .91$, $SE = .013$) was much higher than the proportion of utilities turned off ($M = .668$, $SE = .036$), and proportion of fires extinguished ($M = .496$, $SE = .031$). All performance types were significantly different from each other, p =.000.

Furthermore, a significant interaction between scenarios and the type of performance was found, $F(4,124) = 3.006$, $p = .021$, $\eta_p^2 = .088$. Figure 11, shows that extinguishing fires in the easiest scenario was actually more difficult than in the medium difficulty scenario.

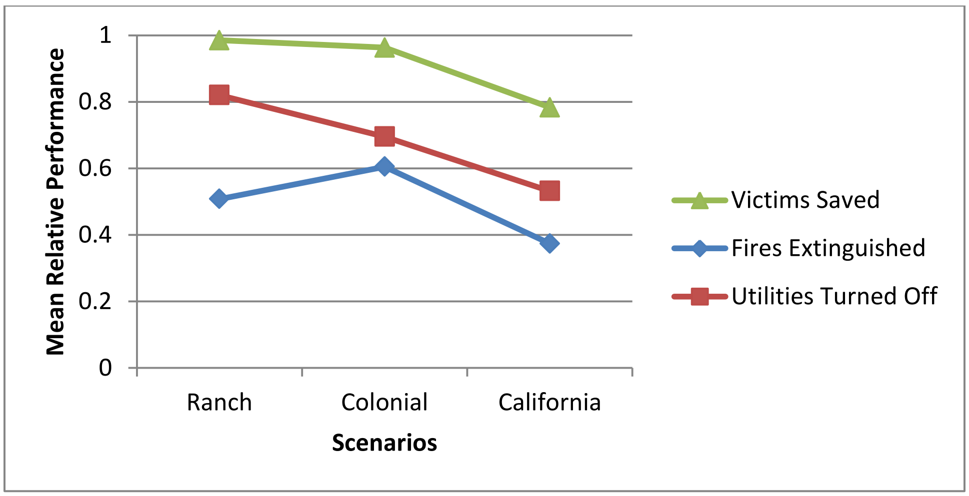


Figure 11. Proportion of each performance type as a function of scenario.

There were no significant differences between the WMC groups, low WMC ($M = .679$, $SE = .039$), medium WMC ($M = .694$, $SE = .040$), and high WMC ($M = .710$, $SE = .040$). One can note however, that the means are trending as one would expect.

Working Memory Capacity and Relative Search Performance. The relationship between WMC and relative search performance was examined using a 3 (WMC) x 3 (Scenarios) mixed ANOVA, with relative search performance as the dependent variable. A significant main effect was found for scenarios, $F(2,62) = 15.829$, $p = .000$, $\eta_p^2 = .999$. Significantly fewer rooms were searched in the California scenario ($M = .753$, $SE = .024$) compared to either the Ranch scenario (*M* = .859, *SE* = .031) or the Colonial scenario (*M* = .890, *SE* = .019), both comparisons were p = .000. There were no significant differences between WMC groups, p > .05.

Working Memory Capacity and Relative Optimal Search Performance. In order to determine if a relationship exists between WMC and the relative optimal search performance, a 3 (WMC) x 3 (Scenarios) mixed ANOVAs was performed using the relative optimal search performance score as the dependent variable. A significant main effect was found for scenarios, $F(2,62) = 18.063$, $p = .000$, $\eta_p^2 = 1.00$. Pairwise comparisons showed that teams demonstrated significantly poorer optimal search performance in the California scenario (*M* = .220, *SE* = .02), than in the Ranch (*M* = .386, *SE* = .034), or Colonial scenario (*M* = .426, *SE* = .031). Once again, there were no significant differences amongst the WMC groups, p > .05.

Team Situational Awareness and Team Performance

Team Composition with respect to Team Situational Awareness. In order to assess whether within-team TSA differences had any impact on team performance, a within-team TSA difference score was obtained for each scenario by taking the team member with the lowest TSA score and subtracting it from the team member with the highest score. This was similar to the analysis examining within-team WMC differences above; however, unlike the WMC difference analysis, three simple regressions for each scenario were performed based on the three performance indices (relative performance, relative search performance, and relative optimal search performance). The teams' TSA difference score was used as the independent variable with a performance indices used as the dependent variable for each analysis. Results showed no significant results, p > .05, thus indicating that large or small differences between team member's TSA did not impact any of the team performance indices. Knowing this, the remainder of the analysis below focuses simply on whether increased TSA resulted in increased performance.

Team Situational Awareness and Relative Performance. For each scenario, three separate simple regressions were performed to see if the TSA score, team queries score, or when teammates were both correct score, were related to the team's relative performance score. Results showed that none of the regressions were significant, $p > .05$. A further analysis was performed by first aggregating each team's TSA score across scenarios for a total TSA score, as well as their relative performance score for a total relative performance score. A multiple regression was then performed utilizing the team's WMC score and total TSA score as independent variables, and total relative performance as the dependent variable. Diagnostics revealed that Team 16's DFBeta of .664 for TSA did exceed the cut-off of .515. When examining the raw data this team had the highest TSA, however, there was no reason to remove them from the analysis. Ultimately, the results were not significant, $p > .05$.

Team Situational Awareness and Relative Search Performance. Once again, for each scenario, the TSA score, team specific queries score, and when teammates were both correct score were each analyzed separately in a simple regression to see if they were related to relative search performance. For the Ranch scenario, it was revealed that the team specific queries score was significantly related to the relative search performance, $B = .018$, $SE = .008$, $\beta = .385$, $t(32) = 2.358$, $p = .025$, CI [.002, .034], $sr^2 = .148$. In other words, the greater the team's understanding of team related queries (e.g. their teammate's location, what they have searched as a team, etc…), the greater the number of rooms searched.

Team Situational Awareness and Relative Optimal Search Performance. Three simple regressions were also performed for each scenario to determine if there was a relationship between a team's TSA score, team queries score, or when teammate's were both correct score, and their relative optimal search performance. Results showed that Ranch TSA was significantly related to the Ranch scenarios' relative optimal search performance, $B = .016$, $SE = .005$, $\beta = .467$, $t(32) = 2.990$, $p = .005$, CI [.005, .027], $sr^2 = .218$. That is, the higher the team's TSA, the more efficient they were in navigating the house. Furthermore, team specific queries and when both teammates were correct, were also significantly related to the relative optimal search performance, $B = .034$, $SE = .008$, $\beta = .444$, $t(32) = 2.805$, $p = .008$, CI [.006, .040], $sr^2 = .197$, and $B = .026$, $SE = .009$, $\beta = .473$, $t(32) = 3.038$, $p = .005$, CI [.009, .044], $sr^2 = .224$, respectively.

Discussion

Using undergraduate participants on a high fidelity firefighting simulator, the primary objectives of this study were to examine the relationships between working memory capacity (WMC) and team situational awareness (TSA), WMC and team performance, and TSA and team performance. It was hypothesized that teams high in WMC would have higher TSA and higher team performance levels; similarly, the higher the TSA levels the higher the team performance. Furthermore, those teams higher in WMC would be less affected by the difficulty of the scenario. The subsequent subsections will discuss the main findings, followed by their implications, the limitations of the study, and finally, suggested future research.

Working Memory Capacity and Team Situational Awareness

Teams with a higher mean WMC had higher TSA, regardless of the team's composition, difficulty of the scenario, or type of TSA query (total TSA or team specific queries). Furthermore, the higher the team's mean WMC, the greater the number of both teammate's getting the TSA query correct, thus giving an indication that they had correct overlapping SA.

The overall finding supports the primary hypothesis, in an extreme team environment, the WMC of each team member matters, the more capacity within the team, the greater the overall TSA. The following sub-sections will discuss this result in greater detail.

Team Composition with respect to Working Memory Capacity. Results showed that teams with large or small differences between the teammates' WMC scores were not related to their TSA levels or team performance scores. Additionally, an analysis comparing large team differences (those teams in which one teammate had a WMC score well above the median and one had a score well below the median) also did not differ significantly in their TSA levels or team performance scores compared to teams where both teammates were below the median or both above the median. While the absence of a relationship between WMC difference scores and TSA advances current knowledge, the lack of relationship between WMC differences and team performance is in line with an earlier study by Cook et al.'s (2004b). They also found that Range, the maximum WMC minus the minimum WMC score, did not predict team performance in a UAV task. While Cook et al. (2004b) did not offer an interpretation on this particular finding, it would have been reasonable to expect that those teams with extreme WMC differences to have a disproportionality lower TSA and team performance score than teams with the same team mean WMC score but little difference in their individual WMC score. One could have expected this as a result of the higher WMC team member feeling the requirement to compensate for a team member that has lower WMC. Instead, as will be discussed in further detail below, it appears that TSA is associated with the overall average of the team's WMC.

Working Memory Capacity and Team Situational Awareness. The hypothesis that teams with a higher mean WMC score have higher TSA was supported in this study. In an attempt to understand this result, one must be reminded of the role of working memory (WM) in information-processing and its relationship with SA. WM temporarily holds and processes information perceived from the environment or retrieved from LTM (Baddeley, 1987, 1997). While there are several models of WM which differ in their components and structure (see Baddeley & Hitch, 1974; Baddeley, 2012; Cowan, 1988, 1995; Engle et al., 1999) they all agree that WM is of limited capacity and that individuals differ with respect to this capacity. Furthermore, "the development of situation awareness and the decision process are restricted by limited attention and working memory capacity for novices and those in novel situations" (Endlsey & Jones, 1997, p.17). Since the present study takes place in a simulated extreme team environment, with dynamic and numerous stimuli (e.g. sirens, flames, shattering glass, etc…), multiple team goals (saving victims, putting out fires, and turning off utilities), and the additional demands of exchanging Shared SA Requirements (see Endsley & Jones, 1997), it appears that the heavy demands placed on team member's WM allow those team's with a greater capacity to hold and process more information resulting in higher TSA. In order to discuss more precisely why this occurs, one must do so in relation to the various WM models. For Baddeley and Hitch's (1974) working memory model, as well as, Cowan's (1988, 1995) embedded process model, teams with higher WMC simply can perceive and process more of the stimulus present, with those team's lower in WMC becoming more overwhelmed by the sheer volume of stimulus. Engle et al.'s (1999) individual differences model would explain the differences between team's higher and lower in WMC as reflecting the ability of the team to focus on the pertinent information in the environment while disregarding unimportant distractions.

In light of the results which found no relationship between WMC differences and TSA but a strong relationship between the team's mean WMC score and TSA, it simply appears that the more WMC that can be brought to the team, the better.

Scenarios and Team Situational Awareness Levels. Across all teams, as expected, TSA was higher in the easiest scenario compared to the hardest scenario. The importance of such a finding demonstrates that the intended manipulation of the scenario difficulty levels were achieved as it was harder for teams to achieve TSA when the environment was more challenging. Likewise, also as expected, level 1 TSA (perception) was higher than level 2 TSA (comprehension) and level 3 TSA (projection of future status). Once again, this provides evidence that the SA queries produced for this study did achieve their intended effect since level 2 and 3 TSA should be more difficult. This is because not only is the team member relying on what they perceived in Level 1 but they also must retrieve information from the LTM to make sense of it and understand what it means for the future. Furthermore, the interaction between scenarios and TSA levels showed that in the hardest scenario, level 3 TSA, future projection, was much lower than in the other scenarios, indicating that trying to project what should happen in the most difficult scenario becomes that much harder. Therefore, taken together, all of these findings show that the more difficult the scenario, the greater the detrimental effect will be on TSA. Not only do these findings help indicate that the scenario manipulation and the TSA queries achieved their intended effect but it also shows that the lack of significant results involving an interaction with the team's mean WMC score goes against what one would expect. It was hypothesized that teams higher in WMC will be less affected by the difficulty of a scenario on their TSA than teams lower in WMC. Instead, teams higher in WMC appear to be equally affected by the difficulty level than teams lower in WMC. Likewise, teams high and low in their mean WMC score were equally affected by the more challenging level 2 and level 3 TSA queries.

To try to explain such results requires some speculation. With respect to TSA and scenario difficulty, perhaps the sheer amount of stimuli present in each of the three scenarios meant that the WMC of each team was completely consumed throughout, with the higher capacity teams continually having a proportionate advantage in each scenario. With respect to TSA levels, while teams higher in WMC had higher TSA scores across levels, the proportional decrease of scores between level 1 TSA and level 2 and 3 TSA, regardless of WMC, could reflect the same limited knowledge contained in LTM by all participants. In other words, the higher level 2 and 3 TSA scores exhibited by teams high in WMC, could simply be the follow on to perceiving more of the stimulus in the level 1 TSA. Ultimately, these results may also reflect that while having higher WMC is a benefit, it is only so up until the capacity of when WM is reached, after which, it ceases to be advantageous.

Working Memory Capacity and Team Specific Queries. Where the TSA measure above reflected "the degree to which every team member possesses the SA required for his or her responsibilities" (Endsley, 1995, p. 39) which included all SA queries directed both at the individual (e.g. where you are and what you are doing) and team (e.g. where your teammate is and what they are doing), WMC also had a relationship to the team specific queries. That is, teams with a higher mean WMC had higher team specific queries scores than teams with low WMC. This indicated that teams with greater WMC were better able to perceive, comprehend and project information present in the environment pertaining to one's teammate and the team as a whole. Furthermore, one can speculate that in order for this result to occur, teams higher in mean WMC were better able to attend to and process pertinent information given to them by their teammate (i.e. in the form of communication) or by what they saw. For example, if a teammate finds and rescues an incapacitated victim from a 2nd floor bathroom, this information needs to be exchanged with the other teammate in order to pass on the information required to get the team query correct. This places greater demands on the WM and is, perhaps, a more concentrated measure of TSA. Overall, this result was particularly interesting because it makes certain the importance of WMC as it pertains exclusively to the team in an extreme team environment.

Scenarios and Team Situational Awareness Levels. Similar to the TSA sub-section above, a main effect was found for TSA levels, as well as an interaction between scenarios and TSA levels. No interaction was found involving the WMC group or a main effect of scenarios. Once again, this reflects the increase in difficulty moving from just perceiving information related to one's teammate and the team as a whole to comprehending it and then projecting future status. The interaction also shows the particular difficulty in projecting future status of the teammate and the team as a whole, particularly in the most difficult scenario. While these results are as expected, having no main effect with respect to scenario was unexpected. Examining the results; however, does show that the means are trending in the expected direction, Ranch (easiest) scenario ($M = 5.061$, $SE = .197$), Colonial (medium) scenario ($M = 4.917$, $SE = .146$), and California (hardest) scenario ($M = 4.513$, $.162$), with the difference between Ranch and California scenario approaching significance, $p = .096$, after making the Bonfferoni correction.

Working Memory Capacity and Both Correct. In an attempt to get a sense of whether the teammates had overlapping correct SA, a score was obtained which reflected how many of the same SA questions the teammates both got correct. As Endlsey and Jones (1997) pointed out, two team members may possess the same SA that is incorrect, the same SA that is correct, and they may also have different understandings of the situation with one member being correct and the other not. While the TSA measure above reflects the team as a whole, there is the possibility that both teammates had high TSA scores but answered different questions correctly indicating that while they had good overall TSA, they weren't necessarily extracting the same correct SA. Finding a relationship between team's higher in their mean WMC and both teammates getting the same queries correct provides evidence that not only do team's with higher mean WMC have better TSA, but that those scores reflect the same and accurate knowledge of their environment, thus indicating that they have correct overlapping SA. Why this occurs precisely, requires some speculation. Perhaps teams with higher mean WMC have more WM resources to commit to communication and coordination behaviours with their teammates' while still being able to focus on their individual avatars. Furthermore, teams with higher mean WMC may also have quickly developed similar and accurate mental models of the simulator functions and environment during the training session. This is turn could have reduced the load on their WM and allowed them to attend to more of the same stimulus in the environment which they both felt were important. This line of speculation does lead to future possible studies to better understand the relationship between a team's mean WMC and their communication/coordination behaviours, as well as, the impact of team mental models.

Scenarios and Team Situational Awareness Levels. When teammates were both correct, a main effect was found for scenarios and for TSA level, however, no interaction was present between the two or involving the WMC groups. This result is as expected, the hardest scenario made it more difficult for the teammates to achieve correct overlapping correct SA compared to the easier scenario; likewise, teammates were more often both correct on level 1 TSA queries than level 2 or 3 TSA queries. The explanation for this is fairly straight forward, the easier something is, the greater chance both teammates will get it right.

Working Memory Capacity and Team Performance

There was no relationship found between WMC and any of the three team performance indices: relative team performance, relative search performance, and relative optimal search performance. While this is contrary to the hypothesis, only one study in the literature could be found that explored WMC and team performance, that is, Cook et al.'s (2004b) study in which higher WMC associated with a specific role within the team predicted team performance. While the present study attempted to expand what little there is in the literature on WMC and team performance, the lack of significant results found here may be attributable to one or more factors. Perhaps one of the most plausible explanations is that the relationship between the team's mean WMC and relative team performance is confounded by too many variables. Indeed, a review of team effectiveness within the framework of Ilgen et al.'s (2005) input-mediator-output-input (IMOI) model, show numerous factors which can influence a team's ultimate performance or output. Individual inputs can include; knowledge, skills, and abilities (KSAs), personality, and attitudes/values; team inputs can include: team training, technology, and leadership (Mathieu, Maynard, Rapp, & Gilson, 2008). In turn, these influence mediators, such as team processes: e.g. mission analysis, planning, monitoring, and coordinating; and, emergent states: e.g. cohesion, and team situational awareness (Burke et al., 2006; Mathieu, Maynard, Rapp, & Gilson, 2008). While any of these factors could be confounding the relationship between WMC and relative team performance, evidence for TSA as a mediator between WMC and the two way-finding performances was found within the present study. As one will see below, TSA and the two way-finding performance measures did have a relationship in the Ranch (easiest) scenario. Perhaps then, WMC results in higher TSA which then results in higher way-finding performances. Although mediation analysis would provide evidence for this speculation, the small sample size (n = 34) makes this prohibitive; however, it should be taken into consideration for future studies.

Another input factor that could be confounding the results is video game experience. While video game experience was measured by the demographic questionnaire, it may not have been measured with precision. The measure asked the participant for the amount of time spent per week playing video games but did not give a time based reference. That is, was it a typical week over the past year or a typical week over the past five years? This would make a difference as the demands placed at university may have reduced the amount of playing time compared to high school. One participant could be thinking of all of the time spent playing in high school and put a high number on the questionnaire where, conversely, someone may have just been thinking about the past year and put a low number despite being an avid gamer in high school. Anecdotal observations from the experimenter and comments made from the participants themselves indicated that those who play video games had an easier time achieving the performance goals and navigating their way through the scenarios. Therefore, a lack of relationship between WMC and team performance could be the result of video game experience being measured imprecisely and therefore, could not be controlled for in the analysis.

Scenarios and Performance Type

As indicated above, no relationship was found between WMC and the three team performance indices: relative team performance, relative search performance, and relative optimal search performance. However, for each of these three analyses, performance was worse in the hardest scenario compared to the medium or easiest scenario. Once again, this serves to show that the scenario difficulty manipulation was achieved.

With respect to the WMC and relative team performance analysis, significant differences were found between the types of performances. Across scenarios, 91% of the victims were saved, as opposed to 66.8% of the utilities turned off, and only 49.6% of the fires extinguished. While better performance was observed for victims over fires, it is in contrast to Rustandjaja's (2013) study who found that performance involving fires were better than victims. Such differences in the two studies are easily explained due to the specific nature of the tasks in each study. In Rustandjaja's (2013) study, participants were simply required to indicate if fires or victims were found by jumping up and down twice, the simulator was not advanced enough to allow victims to be saved or fires to be extinguished. As one would expect, finding fires would be easier due to the size, dynamic nature of the fire, and our prepotent response to look towards flashes. Finding victims, especially ones that are not moving, would no doubt be more difficult. The present study, however, used a much more realistic firefighting simulator which required participants to not only find victims but also carry them outside to safety. Likewise, participants had to find fires and extinguish them using a fire hose. Furthermore, due to the realistic nature of the fires' behaviour, fires were especially challenging to extinguish as they needed to be adequately doused with water; otherwise, they could flare back up even though it had looked like the fire was out. Therefore, while upon first glance it looks as though the present results are in contrast to Rustandjaja (2013), in reality, what the two simulators enabled were quite different and both studies' results make sense.

Finally, in the analysis, an interaction between performance types and scenario was also found, showing that the performance measures decreased the harder the scenario was, except the proportion of fires extinguished. Instead, there were a higher proportion of fires extinguished in the Colonial (medium difficulty) scenario compared to the Ranch (easiest) scenario. While this is in contrast to what one would expect, it became quite evident from observations made by the experimenter, which was confirmed by the simulators' after action report, that the kitchen fire in the Ranch (easiest) scenario was unexpectedly very challenging to extinguish. Of the 34 teams, only four teams were successful in extinguishing it. In order to try and speculate why this occurred, it is important to first note that the fires in the simulator were not one big fire but rather a room full of objects on fire, all of which needed to be put out in order for the simulator to register that the room was extinguished. It appeared that the challenging aspect of the kitchen fire may have been a small fire on top of the fridge. Not only was this fire in an unexpected place, but it was also above eye level and the flames blended closely with the surroundings, causing it to go unnoticed.

Team Situational Awareness and Team Performance

Throughout the literature, the relationship between TSA and team performance appears to be mixed and uncertain; the present study continues with this trend. While no relationship was found between TSA and relative team performance, which was the accomplishment of the primary firefighting tasks, there was a relationship found between TSA and two different Team way-finding performances. Before explaining these results, the analysis of TSA difference scores will be discussed.

Team Composition with respect to Team Situational Awareness. No relationship was found between TSA difference scores (taking the lowest TSA score in the team and subtracting it from the highest) and team performance scores, including the two different way-finding performance scores. In other words, having large or small differences between the teammates' understanding of the situation was not related to team performance. Such a finding suggests that the team member with higher TSA is not trying to compensate for their teammate with lower

TSA. That is, they are not spending enough time trying to help their teammate's understanding of the situation as to detrimentally affect team performance.

Team Situational Awareness and Way-finding Performance. A relationship was found between a team's TSA score and two different way-finding scores in the Ranch (easiest) scenario. It is important to note that the cause and effect relationship between TSA and way-finding performance is unclear. Does higher TSA result in better way-finding, or does better way-finding ability result in higher TSA? The following discussion is meant to examine the relationship of TSA and way-finding performance in general terms before discussing the specifics of the results (i.e. why this relationship was only found in the easiest scenario). In order to try and first understand whether better way-finding performance increases TSA, reviewing the literature on spatial ability is a good starting point. Spatial ability is defined as "the ability to generate, retain, retrieve, and transform well-structured visual images" (Lohman, 1993, p.3). There are three dimensions of spatial ability that have been identified, they include: spatial visualization, "the ability to mentally rotate, manipulate, and twist two- and three-dimensional stimulus objects" (McGee, 1979, p.896); spatial orientation, "the comprehension of the arrangement of elements within a visual stimulus pattern, the aptitude to remain unconfused by the changing orientations in which a spatial configuration may be presented, and the ability to determine spatial orientation with respect to one's own body" (McGee, 1979, p. 897); and spatial relations, which, among others things, include way-finding in real-world environments (Self & Golledge, 1994, p.236). Understanding the connection between way-finding and TSA in the present study follows quite simply, "knowing where one is in space is very helpful in obtaining and maintaining situation awareness" (Shanmugaratnam & Parush, 2009). Conversely, causation could go the other way; greater TSA could allow for more effective way-finding. That is, those teams that possessed higher TSA better understood what rooms had been searched and by whom, thereby making it easier to ensure all of the rooms had been searched and in an efficient manner. For those teams with lower TSA, it resulted in a much more disorganized and less effective search.

Way-finding Performance and Scenarios

It is important to note that the relationship found between TSA and way-finding only occurred in the Ranch (easiest) scenario. While one would expect such a relationship in all of the scenarios, perhaps the harder two scenarios simply overwhelmed every team's ability to effectively and efficiently search. That is, teams higher in TSA simply could not leverage that information like they could in the easier scenario to ensure they have searched every room or searched the house in an optimal fashion. No doubt having multiple floors decreased the likelihood of inexperienced participants being able to maintain the SA required to coordinate the search. Additionally, while relative optimal search performance was related to TSA, team specific queries, and when teammates were both correct, only team queries was significantly related to relative search performance in the Ranch scenario. This result may reflect the level of difficulty inherent in either way-finding task. Relative search performance represents the effectiveness of the search which appears to tap into just team specific knowledge (e.g. what rooms have been searched by the team), while relative optimal search performance reflects not only the basic effectiveness but also the efficiency. Therefore, it makes sense that both teammates need to have correct overlapping SA (as represented in the both correct measure) as they coordinate an efficient rooms search; likewise, in addition to team knowledge, a better understanding of oneself relative to the teammate is required, hence why this measure also correlates with total TSA.

Team Situational Awareness and Relative Team Performance. The lack of findings between these constructs is not entirely surprising. Rafferty et al. (2010) explained that while there appears to be a positive relationship between TSA and team performance, it doesn't guarantee it. Indeed, recent studies conducted within the same lab also found no relationship between TSA and team performance (Ma, 2012; Rustandjaja, 2013). To try and understand why there was no relationship found in the present study, it became plausible that the team's way-finding performance may be acting as a mediating variable. The impetus for this speculation first came from examining the literature in which TSA and team performance held a relationship. Many of these studies had a way-finding, or spatial, component linked with performance. For example, Cooke and Kiekel (2001) and Cooke et al.'s (2004b) UAV studies had a score associated with the rate of waypoints acquired which constituted part of the overall team performance score and in Sulistyawati, Wickens, and Yoon's (2009) study, the main goal was to sweep the enemy planes from an assigned navigation route. A *post-facto* analysis examining the relationship between the two way-finding performance scores and relative team performance provided strong evidence for this assertion. Relative search performance was significantly correlated with relative team performance in all three scenarios; Ranch ($r = .64$, $p = .000$), Colonial ($r = .60$, $p = .000$), and California ($r = .53$, $p = .001$). Relative optimal search performance was also significantly correlated with Ranch's relative team performance ($r = .396$, $p = .02$) and approached significance with the Colonial ($r = .279$, $p = .11$) and California ($r = .313$, $p = .07$) scenarios. Simply speaking, way-finding performance correlated with team performance and may be mediating the relationship between TSA and Relative Team Performance.

In the process of examining the literature another possible explanation for the lack of findings between TSA and relative team performance arouse. It is hard to miss that the studies which found a relationship are most often connected to expert teams performing multiple complex tasks. For instance, in the two examples at the beginning of the paper, the loss, or lack, of TSA was one of the causes attributed to the deaths of firefighters at the World Trade Centre during 9/11 (Garrity, 2007; Lipton & Glanz, 2002) as well as the deaths of Canadian soldiers at Tarnak Farms, Afghanistan (Dumais & Sargeant, 2002). Likewise, Rafferty, Stanton, and Walker (2010) also concluded that TSA was a major factor in the deaths of British soldiers in Iraq after analyzing a fratricide incident in 2003. Studies that were conducted in a laboratory setting continue with this trend. Sonnenwald and Pierce (2000) gathered data from experienced personnel in a simulated military C2 environment, Entin and Entin's (2000) study used mid-career naval officers, and Sulistyawati, Wickens, and Yoon's (2009) study used experienced pilots. With the exception of Cooke & Kiekel (2001) and Cooke et al.'s (2004b) UAV studies involving undergraduate participants, many of the links found between TSA and team performance in the literature involved expert teams performing complicated tasks. One possible explanation for this can be found in Endsley's (1990) study in which only those participants with the technical and operational capabilities could take advantage of SA and leverage it into performance (Endsley, 1995). Therefore, it is plausible that experts may be more apt to leverage TSA into team performance especially when more complex tasks are performed.

Practical Implications

Wildman et al.'s (2012) literature review on team knowledge identified dynamic team knowledge as the number one research need; this study fulfilled its goal to help meet this need.

In the process, it also achieved its main practical goal, to give support for using WMC as a selection tool.

Support for Working Memory Capacity as a Selection Tool. The use of information processing tests for personnel selection has been limited (Arthur et al., 2004). This study advanced the understanding of the importance of WMC in TSA; the higher the team's average WMC, the higher their TSA. Therefore, for teams in which TSA is considered important, WMC should be considered for personnel selection.

Use of Flame-Sim for Research. Although Flame-Sim is designed to train professional fire-fighters, it demonstrated its utility in a research environment. The ability to create and save dynamic, and realistic scenarios, coupled with the after action report that logged key data, took away many of the concerns associated with experimenter bias. With any software, bugs can be expected, Flame-Sim was no exception. However, the stability of program (it did not crash during any of the trials), as well as, the willingness of Flame-Sim's team to correct any bugs found, make this a current and future platform for which to study extreme teams.

Limitations of the Study

Undergraduate Students as Participants. For the present study, it was desirable to have young undergraduate students, as they, generally speaking, reflect the age and the inexperience from which police, the military and firefighters often recruit from. On the other hand, extreme team members are given a substantial amount of training before being placed in harm's way. With training, comes the formation of team mental models which allow for the integration and comprehension of information that help project future events without taxing the working memory (Endsley et al, 2000). Therefore, while the present study appears to have adequately taxed the WM, it is unclear what the impact would be on more experienced firefighters in the same simulated environment.

Sample Size. While the sample size was adequate to detect several relationships within the study, a much larger sample size could have allowed for more complex mediation and causal modeling analysis.

Video Game Experience. The present study was unable to control for video game experience. Upon close examination of the demographic questionnaire, it appears that one issue may have been the nature of the video game experience questions. The questionnaire asked the participants to provide the amount of time spent per week playing video games on a console (e.g. Xbox, computer, and phone. These scores were then added together to provide a total video game experience measure. The issue lies with the ambiguity of the questions as it did not give a time based reference. For example, some participants may have taken the questions to mean time spent recently; others may have taken it to mean time spent since being introduced to video games. These two interpretations would have an impact; those that took the question to mean recently, would likely have seen a decrease in video game playing since attending university compared to a participant that took, for example, high school into consideration. Therefore, future demographic questionnaires should be more specific.

Team Situational Awareness Queries. Originally, the utilities performance type was not a part of the overall team performance measure. However, after several pilot studies, this measure was added not only to increase the amount of tasks needed to be performed but also because teams were attracted to the revolving green icons in Flame-Sim that pointed to the three utilities. Despite not originally being a task, teams would still feel the need to shut them off. Unable to remove this stimulus from Flame-Sim, they were added as a performance type. Due to an oversight, additional SA questions pertaining specifically to utilities were not added to the SA query bank. The effect of this would have been to somewhat dilute the TSA measure.

Future Studies

Experienced Firefighters. While it was advantageous to use inexperienced undergraduates due to the reasons mentioned above, future studies should also include real firefighters with varying degrees of experience, if possible. While studies have shown that the relationship between WMC and SA declines with experience at the individual level (Sohn & Doane, 2003, 2004; Gonzalez & Wimisberg, 2007), Gonzalez and Wimisberg (2007) also note that the dynamics of this relationship over time is not well understood. Therefore, while future studies involving WM should continue to incorporate inexperienced participants, in an ideal situation, having two groups of real firefighters, one which has been trained but lack experience as well as one that is trained and very experienced, should tease out the relationship between WMC and TSA over time. By incorporating real firefighters, it may also be found that WMC continues to hold a relationship with TSA. Since WM is thought to be heavily engaged in not only novices but novel situations (Endlsey & Jones, 1997), as well as dangerous or difficult tasks (Norman & Shallice, 1980; Baddeley, 1986), it is unclear whether WMC will continue to hold a relationship with TSA or not in an extreme team environment.

Working Memory Capacity and Team Effectiveness Factors. A relationship between teams' mean WMC and team performance was not found. However, this may be attributable to several individual and team factors including: personality, communication, team mental models, planning and coordination, to name a few. Future research, using a much larger sample size, could further explore a teams' mean WMC and these factors.

Spatial Ability. Since the present study found a relationship between way-finding performance and TSA, as well as way-finding and team performance, it does present an intriguing question as to the role of spatial ability. For example, since WM is but one of five possible factors of individual differences in spatial ability (Shanmugaratnam & Parush, 2010), perhaps, it is spatial ability that is driving both TSA, and team performance, including way-finding performance, in an extreme team environment. While the relationship between spatial ability and individual performance has been studied extensively, future studies could better examine the role of spatial ability of team member's in an extreme team environment. Similar to the present studies' examination of WMC and its relationship with TSA and team performance, the impact of spatial ability on these and other team effectiveness factors could be explored. Unlike the present study which did not measure spatial ability, valid and reliable measures that tap into all three dimensions of spatial ability (for a comprehensive list see Shanmugaratnam & Parush, 2010) should be utilized.

Conclusion

This study advanced the literature by finding a positive relationship between the teams' mean WMC and TSA. Likewise, the higher the teams' mean WMC, the higher they scored on their team specific SA queries and the more they had correct overlapping SA. It also appears that large differences between team member's WMC do not have an association with TSA; therefore, for extreme teams that rely on TSA, it simply appears that the more WMC that can be brought to the team, the better. Furthermore, higher TSA was also found to be associated with higher way-finding performance, which in turn, was also found to be related to higher team performance. In the domain of personnel selection, information-processes tests, such as complex span tasks, are underutilized; therefore, this study supports the use of WM tests in personnel selection.

References

Arthur, W. Jr., Doverspike, D., & Bell, S, T. (2004). Information-processing tests. In Thomas, J.C. (Eds.), *Comprehensive Handbook of Psychological Assessment Vol. 4, Industrial and Organizational Assessment* (pp. 56-74). Hoboken, New Jersey: John Wiley & Sons, Inc.

Atkinson, R.C. & Shiffrin, R.M. (1968). Human memory: A proposed system and its control processes. In Spence, K.W., & Spence, J.T. (Eds.)., *The Psychology of Learning and Motivation: Advances in Research and Theory* (pp. 89-195). New York, NY: Academic Press.

Baddeley, A.D., & Hitch, G. (1974). Working Memory. In Bower, G.A., *The Psychology of Learning and Motivation: Advances in Research and Theory* (pp. 47-89). New York: Academic.

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

Baddeley, A.D. (1996). Exploring the central executive. *The Quarterly Journal of Experimental Psychology, 49A(1)*, 5-28.

Baddeley, A.D. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Sciences, 4(11)*, 417-423.

Baddeley, A.D. (2007). *Working Memory, Thought, and Action*. Oxford, UK: Oxford University Press.

Baddeley, A.D. (2012). Working memory: Theories, models and controversies. *Annual Review of* Psychology, 63, 1-29.

Baril, M., Matte, G., Hodgson, M., & Levesque, D. (2002). Tarnak Farm Board of Inquiry – Final Report. Retrieved from http://publications.gc.ca/collections/Collections/D2-138-2002E.pdf

Bell, S.T. (2007). Deep-level composition variables as predictors of team performance: A meta-analysis. *Journal of Applied Psychology, 92*(3), 595-615.

Breton, R., & Rousseau, R.(2003). *Situation awareness: a review of the concepts and its measurement.* (Technical Report TR-2001-220). DRDC Valcatier.

Burke, C. S., Stagl, K. C., Salas, E., Pierce, L., Kendall, D., & Darwin, C. (2006). Understanding Team Adaptation : A Conceptual Analysis and Model, *91*(6), 1189-1207.

Cass, E., & Parush, A. (in press). Can situation awareness be predicted? Investigating relationships between Cogscreen – AE and pilot situation awareness. *International Journal of Aviation Psychology*.

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*(1), 21-41.

CogScreen. (2012). *Overview*. Retrieved from http://www.cogscreen.com/Default.aspx

Cohen, S., Eimicke, W., & Horan, J. (2002). Catastrophe and the public serve: A case study of the government response to the destruction of the world trade center. *Public Administration Review*, 62, 24-32.

Conway, A.R.A, Kane, M.J., Bunting, M.F., Hambrick, D.Z., Wilhelm, D., & Engle, R.W. (2005). Working memory span tasks: A methodological review and user's guide. *Psychonomic Bulletin and Review*, *12*(5), 769-786.

Cooke, N.J., Shope, S.M., & Kiekel, P.A. (2001). Shared-knowledge and team performance: A cognitive engineering approach to measurement (Technical Report: F49620-98-1-0287). AFOSR.

Cooke, N.J., Gorman, J.C., & Winner, J.L. (2007). Team Cognition. In F.T. Durso (Eds.), *Handbook of Applied Cognition* (pp. 239-270). West Sussex, England: John Wiley & Sons, Ltd.

Cooke, N., Salas, E., Kiekel, P., & Bell, B. (2004a). Advances in measuring team cognition. In E. F. Salas, Stephen M. (Eds.), *Team cognition: Understanding the factors that drive process and performance.* (pp. 83-106): Washington, DC, US: American Psychological Association.

Cooke, N.J., DeJoode, J.A., Pedersen, H.K., Gorman, J.C., O'Connor, O., and Kiekel, P.A. (2004b). *The role of individual and team cognition in uninhabited air vehicle command-and-control*. (Technical Report No. F49620-03-1-2004).Arlington, VA: AFOSR

Cooke, N.J. & Kiekel, P.A. (2001). Measuring team knowledge during skill acquisition of a complex task. *International Journal of Cognitive Ergonomics, 5*(3), 297-315.

Cowan, N. (1988). Evolving conceptions of memory storage, selective attention, and their mutual constraints within the human information-processing system. *Psychological Bulletin, 104(2)*, 163-191.

Cowan, N. (1995). *Attention and Memory: An Integrated Framework*. New York, New York. Oxford University Press.

Cowan, N. (1999). An embedded process model of working memory. In Miyake, A., & Shah, P., *Models of Working Memory: Mechanisms of Active Maintenance and Executive Control* (pp. 62-101). Cambridge, UK: Cambridge University Press.

Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences, 24*, 87-185.

Cowan, N. (2005). *Working memory capacity*. New York, New York. Psychology Press.

Cowan, N. (2010). The magical mystery four: How is working memory capacity limited, and why? *Current Directions in Psychological Science*, 19(1), 51-57.

Daneman, M., & Carpenter, P.A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior, 19*, 450-466.

Devine, D.J., & Philips, J.L. (2001). Do smarter teams do better: A meta-analysis of cognitive ability and team performance. *Small Group Research, 32*(5), 507-532.

Doane, S.M. (2003). *Relating memory processes to aviation flight situation awareness abilities*. (Technical Report 20030929 038). Office of Naval Research, US.

Dumais, M.J., & Sargeant, S.T. (2002). Summary of Facts: Tarnak Farms Friendly Fire Incident near Kandahar, Afghanistan, 17 April 2002. Coalition Investigation Board Report: Tarnak Farms, Afghanistan

Endsley, M.R., Holder, L.D., Leibrecht, B.C., Garland, D.J., Wampler, R.L., & Mathews, M.D. (2000). *Modeling and measuring situation awareness in the infantry operational environment.* U.S. Army Research Institute for the Behavioural and Social Sciences. Infantry Forces Research Unit.

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

Endsley, M.R., & Boldstad, C.A. (1994). Individual Differences in pilot situation awareness. *The international Journal of Aviation Psychology, 4*, 241-264.

Endsley, M.R. & Jones, W.M. (1997). *Situation awareness, information dominance and information warfare.* (Technical Report 97-01). US Air Force Armstrong Laboratory.

Endsley, M.R. (1988). Design and evaluation for situation awareness enhancement. *Proceedings Of the Human Factors Society 32nd Annual Meeting* (pp. 97-101). Anaheim, CA: Human Factors Society.

Endsley, M.R. (1987). SAGAT: A methodology for the measurement of situation awareness (NORDOC 87-83). Hawthorne, CA: Northrop.

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

Engle, R.W., Kane, M.J., & Tuholski, S.W. (1999). Individual differences in working memory capacity and what they tell us about controlled attention, general fluid intelligence, and functions of the prefrontal cortex. In Miyake, A., & Shah, P., *Models of Working Memory: Mechanisms of Active Maintenance and Executive Control* (pp. 102-134). Cambridge, UK: Cambridge University Press.

Entin, E. B. & Entin, E. E. (2000). Assessing team situation awareness in simulated military missions. *Proceedings of the Human Factors and Ergonomics Society 44th Annual Meeting*, San Diego, CA.

Flam-Sim (2012). *Flam-Sim User Manual*. Retrieved: http://www.flame-sim.com/upload/presentationmaterials/15/presentationmaterials-1270669247.pdf

Garrity, M. (2007). *Investigating team collaboration of the fire department of new york using transcripts from September 11, 2001* (Unpublished master's thesis). Naval Postgraduate School: California.

Gonzalez, C. & Wimisberg, J. (2007). Situation awareness in dynamic decision making: effects of practice and working memory. *Journal of Cognitive Engineering and Decision Making, 1*(1), 56-74.

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

Granlund, R., Dahlback, N., Granlund, H., and Johansson, B. (2010). *The effect of a geographical information system on communication in a professional emergency response organization.* Paper presented at the seventh international information systems for crisis response and management (ISCRAM) conference, Seattle, USA.

Heitz, R.P., Unsworth, N., Engle, R.W (2005). Working memory capacity, attention control, and fluid intelligence. In Wilhelm, O., & Engle, R.W., *Handbook of Understanding and Measuring Intelligence* (pp. 61-78). Thousand Oaks, California: Sage Publications.

Ingle, D.R., Hollenbeck, J.R., Johnson, M., & Jundt, D. (2005). Teams in organizations: From input-process-output models to IMOI models. *Annual Review of Psychology, 56*, 517-543.

Johansson, B, Trnka, J. and Granlund, R. (2007). *The Effect of Geographical Information Systems on a Collaborative Command and Control Task*. Paper presented at the fourth international information systems for crisis response and management (ISCRAM) conference, Delft, Netherlands.

Jones, H., & Hinds, P. (2002, November). *Extreme work teams: Using SWAT teams as a model for coordinating distributed robots*. Paper presented at the Conference on Computer Supported Cooperative Work (CSCW), New Orleans, LA.

Kutmer, M.H., Nachtsheim, C.J., Neter, J., &Li, W. (2005). *Applied Linear Statistical Models* (5th Ed.). New York, NY: McGraw-Hill/Irwin.

Lipton, E., & Glanz, J. (2002, January 23). Ground zero firefighting: 9/11 inspires call to review response plan for crisis. *New York Times*, B4.

Lohman, D. F. (1993). Spatial ability and G. *Paper presented at the first Spearman seminar.* Plymouth, UK.

Massaro, D.W., & Cowan, N. (1993). Information processing models: Microscopes of the mind. *Annual Review of Psychology, 44*, 383-425.

Ma, C. (2012). Situation awareness, communication, and teamwork: Re-examining the role of a team display in wildland firefighting command and control. (Unpublished master thesis). Carleton University, Ottawa, Canada.

Mathieu, J., Maynard, M.T., Rapp, T., & Gilson, L. (2008). Team effectiveness 1997-2007: A review of recent advancements and a glimpse into the future. *Journal of Management, 34*(3), 410-476.

McGee, M. G. (1979). Human Spatial Abilities: Psychometric Studies and Enviromental, Genetic, Hormonal, and Neurological Influences. *Psychological bulletin*, *86*(September), 889-918.

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

Miyake, A., & Shah, P. (1999). *Models of Working Memory: Mechanisms of Active Maintenance and Executive Control*. Cambridge, UK: Cambridge University Press.

Rustandjaja, N. (2013). Situation awareness in firefighting: The impact of sudden obstructions, building complexity, and information availability. (Unpublished master thesis). Carleton University, Ottawa, Canada.

Norman, D.A., & Shallice, T. (1980). Attention to action: willed and automatic control of behavior. In Davidson, R.J., Schwartz, G.E., & Shapiro, D. (Eds.), *Consciousness and Sel-Regulation. Advances in Research and Theory* (pp. 1-18). New York, NY: Plenum.

Parush, A., Kramer, C., Foster-Hunt, T., Momtahan, K., Hunter, A., & Sohmer, B. (2011). Communication and team situation awareness in the OR: Implications for augmentative information display. *Journal of Biomedical Informatics, 44*, 477-485.

Parush, A., & Ma, C. (2012). Team displays work, particularly with communication breakdown: performance and situation awareness in a simulated forest fire. Proceedings of the Human Factors and Ergonomics Society 56th Annual Meeting, 383-387.

Rafferty, L.A., Stanton, N.A., & Walker, G.H. (2010). The famous five factors in teamwork: a case study of fratricide. *Ergonomics, 53*(10), 1187-1204.

Redick, T.S., Broadway, J.M., Meier, M.E., Kuriakose, P.S., Unsworth, N., Kane, M.J. and Engle, R.W. (2012). Measuring working memory capacity with automated complex span tasks. *European Journal of Psychological Assessment, 28*(3), 164-171.

Roth, E. M., Multer, J., & Raslear, T. (2006). Shared situation awareness as a contributor to high reliability performance in railroad operations. *Organization Studies*, 27(7), 967-987.

Rustandjaja, N. (2012). Team Situation Awareness in Firefighting: The Impact of Workload History, Building Complexity and Information Availability. (Unpublished master thesis). Carleton University, Ottawa, Canada.

Salas, E. Dickinson, T.L., Converse, S.A. & Tannenbaum, S.I. (1992). Toward an understanding of team performance and training. In R.W. Swezey & E. Salas (Eds.), *Teams: Their Training and Performance* (pp. 3-29). Norwood, NJ: Ablex.

Salmon, P.M., Stanton, N.A., Walker, G.H., Baber, C., Jenkins, D.P., McMaster, R., & Young, M.S. (2008). What is really going on? Review of situation awareness models for individuals and teams. *Theoretical Issues in Ergonomics Science, 9*(4), 297-323.

Salmon, P. M., Stanton, N. A., Walker, G. H., Jenkins, D., Ladva, D., Rafferty, L., & Young, M. (2009). International Journal of Industrial Ergonomics Measuring Situation Awareness in complex systems : Comparison of measures study. *International Journal of Industrial Ergonomics*, *39*(3), 490-500.

Self, C. M., & Golledge, R. G. (1994). Sex-related Differences in Spatial Ability: What Every Geography Educator Should Know. *Journal of Geography*, *93*(5), 234-243.

Shanmugaratnam, S., & Parush, A. (2009). *Situation awareness: a review of definitions, measurements and predictors, and implications for creating a task-based selection battery.* (Technical Report DGMPRA CR 2009-005). Defence R&D Canada.

Sohn, Y.W. & Doane, S.M. (2003). Roles of working memory capacity and long-term working memory skill in complex task performance. *Memory & Cognition, 31*(3), 458-466.

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

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

Sulistyawati, K., Wickens, C.D., & Yoon, P. C. (2009) Exploring the concept of team situation awareness in a simulated air combat environment. *Journal of Cognitive Engineering and Decision Making, 3*(4), 309-330.

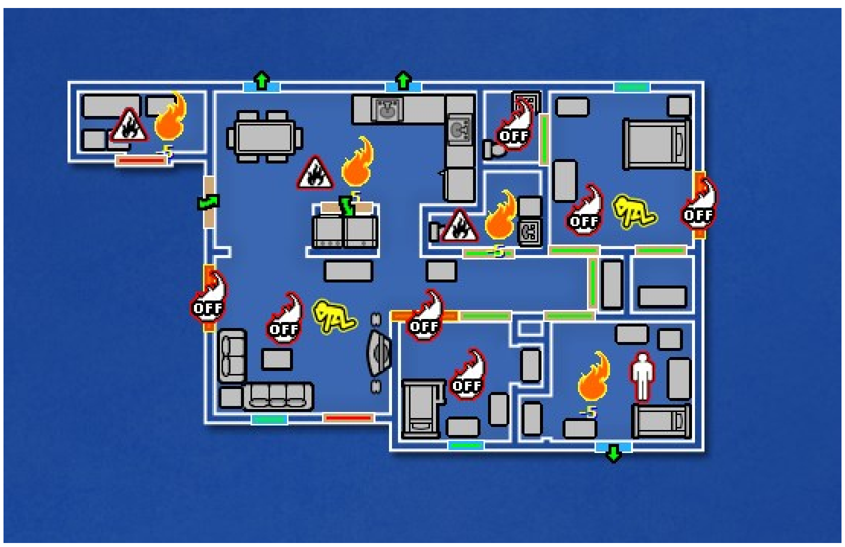
Turner, M.L. & Engle, R.W. (1989). Is working memory task dependent? *Journal of Memory and Language, 28*, 127-154.

Wildman, J.L., Thayer, A.L., Pavlas, D., Salas, Stewart, J.E., & Howse, W.R. (2012). Team knowledge research: Emerging trends and critical needs. *Human Factors*, *54*(1), 84-111.

Appendix A

The Scenarios

The Ranch House – Easy Scenario

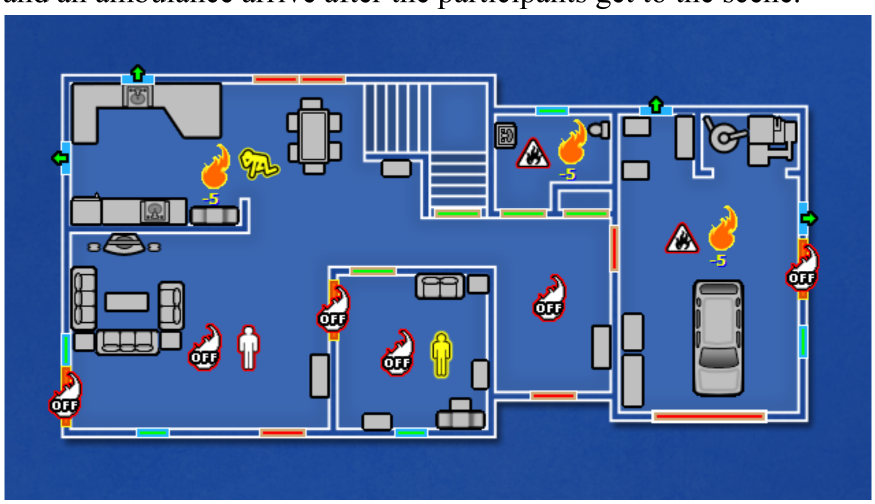


- The ranch house consists of one floor, an attic, and no basement.
- There are four fires, denoted as orange/yellow flame icons. Non-activated fire icons are white with the word off.
- There are two victims, denoted as yellow human figures either incapacitated (laying down icon) or injured (crawling icon). Non-activated victims are white human figures standing upright.
- There are three utilities: gas, water, and electricity that must be turned off.
- Smoke levels can be manipulated by opening (green arrows) or closing windows. Smoke is not an issue for visibility in this scenario.

- An officer and an ambulance arrive after the participants get to the scene.

Colonial Revival – The Medium Scenario

- The medium difficulty scenario takes place in a house that has two floors, a garage, an attic but no basement.
- There are a total of five fires
- There are four victims.
- There are three utilities: gas, water, and electricity that must be turned off.
- Smoke becomes a minor visibility issue for one room in this scenario.
- An officer and an ambulance arrive after the participants get to the scene.

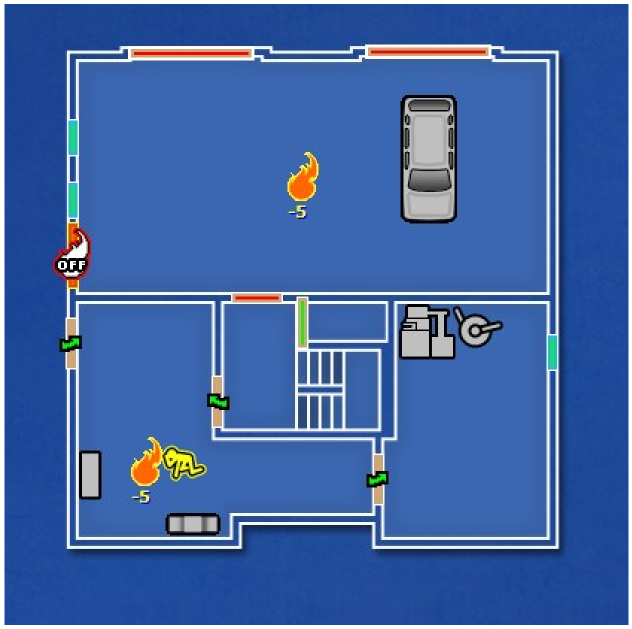


The first floor has three fires and two victims.

The second floor has two fires and two victims.

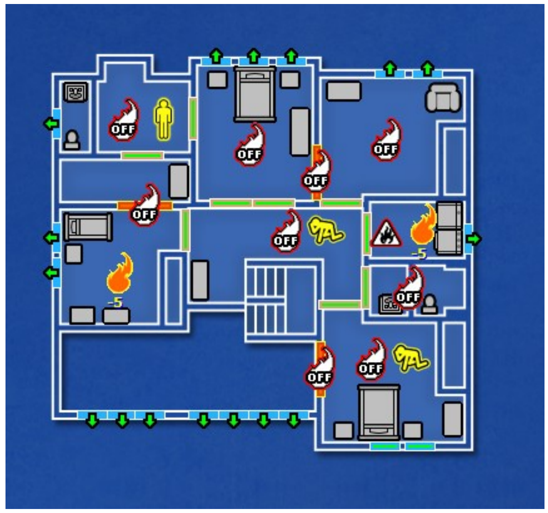
Bungalow (California) - The Difficult Scenario:

- This house has a basement/garage, first floor, second floor, an attic.
- A total of five fires and seven victims are present.
- There are three utilities: gas, water, and electricity that must be turned off.
- Smoke on the second floor restricts visibility in this scenario.
- An officer and an ambulance arrive after the participants get to the scene.



The basement/garage has two fires and one victim.

The first floor has two fires and three victims.



The second floor has two fires, three victims and smoke which obscures visibility.

Appendix B

Informed Consent

The purpose of an informed consent is to ensure that you understand the purpose of the study and the nature of your involvement. The informed consent must provide sufficient information such that you have the opportunity to determine whether you wish to participate in the study.

Present study: Examining the Relationship between Working Memory Capacity, Team Situational Awareness and Team Performance in Extreme Work Teams

Research personnel. The following people are involved in this study, and may be contacted at any time if you have questions or concerns:

Garrett Morawiec (email: gmorawie@connect.carleton.ca, phone:[redacted]
Dr. Avi Parush (e-mail: avi_parush@carleton.ca, phone: 613-520-2600, ext. 6026)

Concerns. Should you have any ethical concerns about this research, please contact Dr. Chris Davis (Associate Chair, Psychology Ethics Board, chris_davis@carleton.ca, 613-520-2600, ext.2251). For other concerns, please contact Dr. Anne Bowker (Chair, Department of Psychology, psychchair@carleton.ca, 613-520-2600, ext. 8218).

Purpose. The purpose of this study is to investigate the relationship between Working Memory Capacity, Team Situational Awareness and Team Performance in Extreme Work Teams. We are interested in understanding whether the working memory capacity of team members corresponds to teams that have higher team situational awareness and team performance.

Task requirements. There are several task requirements contained within the study. After signing this questionnaire we will ask you to fill out a demographic questionnaire (e.g. your past video game experience). Next, you will be asked to complete a working memory test administered on a computer. Once complete, the primary task can begin. After a short training session, we will ask you to participate in three firefighter simulation scenarios in which you and your teammate will be required to enter a burning structure and rescue victims and extinguish fires. Finally, we will ask you to fill out a post-test questionnaire (e.g. how difficult you found the firefighting task). The study will take approximately two hours to complete including multiple breaks.

Benefits/compensation. You will receive a 2% increase in your final grade in PSYC 1001, 1002, 2001, 2002 or NEUR 2001, 2002 for participating in this study.

Potential risk/discomfort. Risks associated with this experiment are similar to any video game environment and therefore would be considered minimal risk. Due to the high fidelity and immersion into the simulated environment, there is a chance that you may feel "cyber sickness" which includes symptoms such as dizziness and nausea. Overall, this risk is very minimal and will be mitigated by taking a break after each scenario; however, if you begin to experience any of these symptoms immediately stop the simulator and inform the experimenter, you will not be penalized. The debriefing form at the end of the study provides contact information for local support services that you may contact if you need or want help.

Anonymity/Confidentiality. The data collected in this study are strictly confidential and anonymous. All data are coded such that your name is not associated with the responses you provide. Any identifying information associated with your code will be confined to a single page that will be separated from your questionnaire, and kept in a separate, secured file by the research investigators, who will keep this information confidential. The informed consent and all other identifying information will be secured in a locked cabinet located within a code protected lab. This information will then be destroyed after three years. The anonymously coded data will be kept and used by the primary investigator and the faculty sponsor for research and teaching purposes.

Right to withdraw. Your participation in this study is entirely voluntary. At any point during the study, you have the right to not complete certain questions, or to withdraw without penalty. If you withdraw, you have the right to request that your data be deleted. This study has received clearance by the Carleton University Psychology Research Ethics Board
(Reference #12-193).

Signatures
I have read the above form and understand the conditions of my participation. My participation in this study is voluntary, and I understand that if at any time I wish to leave the experiment, I may do so without having to give an explanation and with no penalty whatsoever. Furthermore, I am also aware that the data gathered in this study are confidential and anonymous with respect to my personal identity. My signature indicates that I agree to participate in this study.

Participant's Name: ________________________ Participant's Signature: _________________

Researcher’s Name: _______________________ Researcher’s Signature: _________________

Date:

Appendix C

Demographic Questionnaire

Team#:

Participant#:

Background Information
What is your sex?
☐ Male
☐ Female

What is your age? __________

What is your first language?
☐ English
☐ French
☐ Other (Please specify) ____________________________

If English is not your first language, how do you rate your fluency?
☐ Poor
☐ Average
☐ Superior

You are working with a teammate in this study, how familiar are you with him/her?

☐ I have never met him/her before
☐ I have met him/her before, but we are not well acquainted
☐ I am well acquainted with my teammate

How many hours a week do you play video games on a computer?

How many hours a week do you play video games on a console (e.g., wii, xbox, playstation)?

How many hours a week do you play video games on a smart phone or tablet (e.g. iPhone, iPad)?

Appendix D

Training Instructions

In a moment, you will be situated in a training warehouse, when the training scenario commences, there will be several tools in front of you and a fire truck and an avatar fireman behind you. To maximize the training, these are the following movements and equipment functions that must be practiced, we will do this together:

Step 1: Using the Look/Turn button, scan the room.

Step 2: Using the Move/Strafe button, move forward and backward, side to side.

Step 3: Combine Step 1 and 2 to move and look around at the same time.

Step 4: Using the stances button, get on your hands and knees. Once on the ground perform step 3. Now get up using the same button.

To note, to move faster at any time, press the sprint button (the speed will be dictated by your stance and load).

You can now perform the basic movement functions. Next, the following functions are required to obtain and utilize the key pieces of equipment.

Step 1: Go to the fire truck, press B to obtain a tool. Find an entry tools. Now, move to the other side of the warehouse where you will encounter a door requiring “forced entry”. Open door with forced entry tool. Return to fire truck.

Step 2: Notice at the bottom right hand corner three red boxes. Each box signifies a different tool that can be carried at once by your avatar. Press the Y toggle button to free up an empty red box. Now ensure you are beside the fire truck and press B to obtain a hose.

Step 3: Use the hose by pressing A. Turn off the hose by pressing A.

Step 4: Press X to drop the hose, press X again to pick up the hose. Note, this is also the same method to pick up victims and drop them off.

Step 5: Finally, the last tool you will need to know how to get is the gas shut off wrench, obtain this now by going to the fire truck and finding it under miscellaneous.

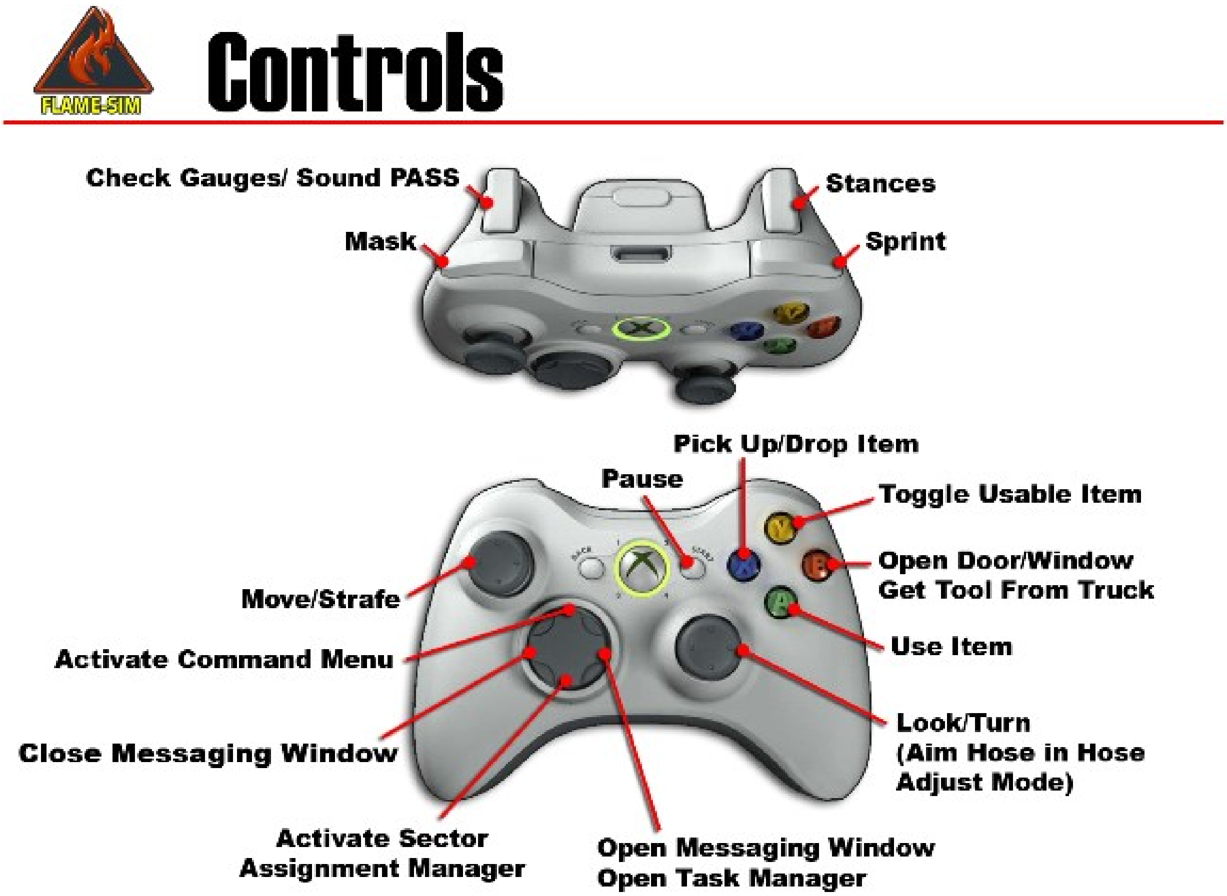
Do not dawn the mask in training or the scenarios. Furthermore, take note of the human figure in the bottom left hand corner of your screen, disregard this icon throughout.

Lastly, during the scenario, you will be given the command “freeze”. When this happens, the firefighter on the right is to press the “pause” button. Practice this now.

You have now learnt how to use the equipment function buttons. Now take 5 minutes to practice both the movement and equipment functions.

Appendix E

Flame-Sim Controller Functions



Appendix F

Participant Instructions

You are about to begin three sessions in which you and your teammate will be required to rescue victims, extinguish fires, and ensure the safety of yourself and fellow firefighters by ensuring the utilities are shut-off. Your first priority is for the safety and well-being of the victims, extinguishing the fire is secondary and, lastly, it is also important to shut off the gas, electricity, and water connected to the house.

When the simulator begins, there will be radio chatter displayed on the screen regarding the location and type of fire. Disregard this information and listen to the information provided to you by the investigator who will act as dispatch. Once the fire engine arrives on scene, your avatar will automatically get out of the fire truck. At this point, you will be able to see your teammate's avatar and you can commence your mission. The mission will be considered complete once you, as a team, have rescued all of the victims, put out all of the fires, and shut off the gas, electricity, and water connected to the house or the time limit of 8 minutes (run time) has expired.

If you see a victim standing up, he/she is considered "uninjured", if he or she is on their hands and knees, they are considered "injured", if they are laying on the ground, they are considered "incapacitated". Once a victim is found, you need to take them outside of the structure and place them on the grass.

While it is possible to stick your head up into any attic that may be found and look around, it is not possible to get off the ladder and complete a physical search of the attics. Therefore, try your best to look around from the ladder but do not spend too much time doing this.

If, between you and your teammate, you believe that you have saved all of the victims and put out all of the fires, return to the fire truck.

Freezes during the Scenario

When given the command "freeze", the participant on the right computer will press the "start/pause" button on the controller which will freeze the computer, at which time you are to turn around. A list of questions will be given to you, the majority of them will have either a true, false, or unknown option. For an answer to be considered true, all aspects of the question must be true. For example:

The victim I found in the upstairs bathroom was incapacitated. T / F / U

In this example, you must have found the victim and the victim has to be lying on the floor. If your teammate found a victim lying on the floor, it would be false. If you do not know the answer because you or your teammate have not been exposed to the event (no way of knowing), circle unknown. If there is only a T or a F option to a question, try your best and circle one. YOU CANNOT DISCUSS THE ANSWERS WITH YOUR TEAMMATE.

Some helpful advice

1. In order to best achieve the mission, teamwork (communication, collaboration, and coordination) is important.

2. Also, according to professional firefighters, the optimum search pattern is to follow the walls and rooms clockwise (left hand search technique) or counter clockwise (right hand search technique) to make sure you have searched each room in an organized manner.

3. When you encounter a fire, there will likely be several items on fire within a room. In order for the fire to be considered extinguished, you must extinguish all fires within a room. Therefore, make sure you douse the fire with lots of water, once you believe it is out, take a second and look around the room to make sure.

Do you have any questions?

Appendix G

SA Query Bank

Note:
1. Lines indicate a piece of information (quantity, location etc…) that will be added by the experimenter based on the scenario.

2. Unknown will not be an option for level three queries as this answer effectively allows the participant to avoid the question.

Query ID	TYPE	**FACTS/PERECPTION**
1.1.1	Individ/Generic	There is _____ stories to the house
1.1.2	Individ/Generic	There is a fire hydrant present
1.1.3	Individ/Generic	The address of this house is _________
1.1.4	Individ/Generic	This appears to be a new construction
1.1.5	Individ/Generic	Your Engine number is _______
1.1.6	Individ/Generic	As the engine approached the house, smoke was visible
1.1.7	Individ/Generic	As the engine approached the house, fire was visible
1.1.8	Individ/Generic	There is an officer present on the scene.
1.1.9	Individ/Generic	There is an ambulance present on the scene
1.1.10	Individ/Generic	There was a bystander in front of the house
1.1.11	Individ/Generic	The bystander spoke with me
1.1.12	Individ/Generic	The front door was locked
1.1.13	Individ/Generic	There are more than ________ rooms to search
1.1.14	Individ/Generic	I have searched ______ rooms
1.1.15	Individ/Generic	A victim was found in the ______(loc)
1.1.16	Individ/Generic	I found the victim in ________(room) _________(not moving on the floor, on hands and knees, standing up)
1.1.17	Individ/Generic	I have saved _________ victims
1.1.18	Individ/Generic	I can hear the crackle sound of fire
1.1.19	Individ/Generic	I found a fire(s) in the _______(loc)
1.1.20	Individ/Generic	I found _______(Qty) rooms on fire
1.1.21	Individ/Generic	I have extinguished ______ fires
1.1.22	Individ/Generic	I am located in _________(loc)
1.1.23	Individ/Generic	I am currently ________(task)
1.2.1	Team	We have searched the whole first floor
1.2.2	Team	We have searched the whole second floor
1.2.3	Team	We have searched the attic
1.2.4	Team	We have searched the whole basement
1.2.5	Team	We have searched the whole house
1.2.6	Team	My teammate has searched ________ rooms

1.2.7	Team	As a team we have searched ________ Rooms
1.2.8	Team	My teammate found a victim______(loc)
1.2.9	Team	My teammate found the victim in ________(room) _________(not moving on the floor, on hands and knees, standing up)
1.2.10	Team	My teammate has saved ______ victims
1.2.11	Team	Together we have saved ______ victims
1.2.12	Team	My teammate found fire(s) in the _______(loc)
1.2.13	Team	My teammate found ______(Qty) rooms on fire.
1.2.14	Team	There are _______(Qty) rooms on fire in this house.
1.2.15	Team	My teammate has extinguished _________ fires.
1.2.16	Team	As a team we have extinguished _________ fires.
1.2.17	Team	My teammate saw a bystander in front of the house
1.2.18	Team	The bystander spoke to my teammate
1.2.19	Team	My teammate is in the same room
1.2.20	Team	My teammate is on the same floor
1.2.21	Team	My teammates task is currently to___________
		MEANING/COMPREHENSION
2.1.1	Individ/Generic	This house is too big for the two if us to search
2.1.2	Individ/Generic	An entry tool was/is required
2.1.3	Individ/Generic	The victim I found _______(loc) is ok.
2.1.4	Individ/Generic	The Victim I found_______(loc) is injured.
2.1.5	Individ/Generic	The victim I found _______(loc) is incapacitated
2.1.6	Individ/Generic	There are still ________ fires to extinguish.
2.1.7	Individ/Generic	I should bring the victim to the ambulance
2.1.8	Individ/Generic	This is primarily a rescue mission
2.1.9	Individ/Generic	This is primarily a fire attack (extinguish fires) mission.
2.1.10	Individ/Generic	I am currently doing what I need to achieve the primary objectives
2.2.1	Team	The victim my teammate found ________is ok.
2.2.2	Team	The victim my teammate found________is injured.
2.2.3	Team	The victim my teammate found ________is incapacitated
2.2.4	Team	My teammate is too far away
2.1.5	Team	My teammate is currently doing what he needs to do to achieve the primary objectives
		FUTURE/PROJECTION
3.1.1.	Individ/Generic	Should an emergency occur, you would be able to get out of the house immediately (less than 10 sec) through a nearby window or door.
3.1.2	Individ/Generic	The victim I found _______(loc) is priority for first aid
3.1.3	Individ/Generic	I will have enough time to extinguish _________ more fires.

3.1.4	Individ/Generic	I will be ______(loc) in one minute
3.1.5	Individ/Generic	Additional firefighters should be requested
3.1.6	Individ/Generic	I will be doing _______ (task) in 1 minute
3.2.1	Team	The victim my teammate found _______(loc) is priority for first aid
3.2.2	Team	Should an emergency occur, your teammate would be able to get out of the house immediately (less than 10 sec) through a nearby window or door
3.1.3	Team	We will have enough time to extinguish all of the fires.
3.1.4	Team	Additional firefighters should be requested
3.1.5	Team	My teammate will be ______ (loc) in one minute
3.1.6	Team	My teammate be doing _______(task) in 1 minute

Appendix H

Post-Scenario Questionnaire

Difficulty Level

1. On a scale from 1 – 10, I found the single story ranch house to be:

Not Difficult									Extremely Difficult
1	2	3	4	5	6	7	8	9	10

2. On a scale from 1 – 10, I found the two story all brick house to be:

Not Difficult									Extremely Difficult
1	2	3	4	5	6	7	8	9	10

3. On a scale from 1 – 10, I found the two story plus basement house to be:

Not Difficult									Extremely Difficult
1	2	3	4	5	6	7	8	9	10

Spatial Difficulty

7. On a scale from 1 – 10, I found trying to remember where I had been, where I was, and where I was going in the single story ranch house to be:

Not Difficult									Extremely Difficult
1	2	3	4	5	6	7	8	9	10

8. On a scale from 1 – 10, I found trying to remember where I had been, where I was, and where I was going in the two story all brick house to be:

Not Difficult									Extremely Difficult
1	2	3	4	5	6	7	8	9	10

9. On a scale from 1 – 10, I found trying to remember where I had been, where I was, and where I was going in the two story plus basement house to be:

Not Difficult									Extremely Difficult
1	2	3	4	5	6	7	8	9	10

Appendix I

Debriefing Form

What are we trying to learn in this research?

Contained within the larger human information processing architecture, working memory is responsible for both holding information coming in from the environment and from long-term memory while simultaneously processing it (Baddeley, 1987, 1997). Attention, which is controlled by the central executive, a component of working memory, is responsible for being able to divide attention between two tasks, being able to focus on one task while disregarding distractors, switching between tasks, and retrieving information from long-term memory. Working memory becomes heavily engaged in novices or those in novel or complex situations. While there are several different models to explain working memory, they all agree that it is of limited capacity and that individuals differ with respect to this capacity (Baddeley, 1987, 1997; Cowan; 1988, 1995).

Situational awareness (SA) has been defined by Endsley (1988) as "the perception of the elements in the environment within a volume of time and space, comprehension of their meaning, and the projection of their status in the near future" (p. 97). Obtaining situational awareness is limited by one's attention and working memory (Endlsey and Jones, 1997). Indeed, research has shown that working memory capacity can predict individual situational awareness (Carretta, Perry and Ree, 1996). Team situational awareness can be thought of "the degree to which every team member possess the SA required for his or her responsibilities" (Endsley, 1995, p. 39). In team research literature, team situational awareness has been shown to predict team performance (Cooke, Salas, Kiekel, and Bell, 2004).

What has not been studied is the relationship between the working memory capacity of individual members of a team and that team's situational awareness. Likewise, team member's working memory capacity and the team's performance has received little attention. Therefore, the primary purpose of this study is to explore these relationships.

Why is this important to scientists or the general public?

There are two primary research goals. One, to advance our understanding of the relationship between working memory capacity and team situational awareness, as well as, working memory capacity and team performance. Second, since individuals differ with respect to working memory capacity, it is hoped that results from this study can be used by organizations to help select the right individuals to work in environments that require high levels of team situational awareness.

What are our hypotheses and predictions?

We predict that teams higher in working memory capacity will demonstrate higher levels of team situational awareness and team performance and that these effects will be more pronounced in more complex environments. We also predict that teams higher in team situational awareness will have higher levels of team performance.

Where can I learn more?

In addition to contacting the principal investigator for more information, the following references provide the basis for this topic area:

Baddeley, A.D. (2012). Working memory: Theories, models and controversies. *Annual Review of* Psychology, 63, 1-29.

Cowan, N. (2005). *Working memory capacity*. New York, New York. Psychology Press.

Engle, R.W., Kane, M.J., & Tuholski, S.W. (1999). Individual differences in working memory capacity and what they tell us about controlled attention, general fluid intelligence, and functions of the prefrontal cortex. In Miyake, A., & Shah, P., *Models of Working Memory: Mechanisms of Active Maintenance and Executive Control* (pp. 102-134). Cambridge, UK: Cambridge University Press.

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

If you would like to learn more about Flame Sim, the firefighter simulator used in this study, visit this site: http://www.flame-sim.com/

Is there anything I can do if I found this experiment to be emotionally upsetting?

Yes. If you feel any distress or anxiety after participating in this study, please feel free to contact the Carleton University Health and Counseling Services at: 613-520-6674, or the Distress Centre of Ottawa and Region at 613-238-3311 (http://www.dcottawa.on.ca).

What if I have questions later?

If you have any remaining concerns, questions, or comments about the experiment, please feel free to contact Garrett Morawiec (Principal Investigator), at: gmorawie@connect.carleton.ca [redacted] or Dr. Avi Parush (Faculty Sponsor) at: avi_parush@carleton.ca, 613-520-2600, ext. 6026. Should you have any ethical concerns about this research, please contact Dr. Chris Davis (Associate Chair, Psychology Ethics Board, chris_davis@carleton.ca, 613-520-2600, ext.2251). For other concerns, please contact Dr. Anne Bowker (Chair, Department of Psychology, psychchair@carleton.ca, 613-520-2600, ext. 8218).

Thank you for participating in this research!

Appendix J

Hierarchical Task Analysis

INCIDENT MANGEMENT

1 Primary Assessment of the Situation

1.1 Gain any information from printout in the fire hall (e.g. location of fire, residential or commercial, 911 or alarm)

1.2. Note time of day

1.3 Communicate with dispatch

1.3 Arrive on scene and look for fire hydrant first

1.4 Observe for fire

1.4.1 If seen, fire well started, consider risk of entry

1.5 Observe for smoke

1.5.1 What's the colour?

1.5.2 What's it doing? Light waft (not well seated) or heavy plumes (well seated).

1.6 Observe for victims

1.7 Consider structure type (e.g. new suburban house vs old downtown house)

1.8 Circle the structure observing

1.9 Talk with bystanders, determine if there is a possibility of someone in side. If so, search for victim is primary task. If not (e.g. family has accounted for everyone), skip to fighting the fire as primary task.

2 Search for Victims

2.1 Decide on tools to bring (e.g. forced entry and hose)

2.1.1 Prep entry check

2.1.1.1 Seal on mask (positive pressure)

2.1.1.2 Test line, purge air

2.2 Bring hose even if search and rescue primary task

2.3 Gain entry through front door if possible, easier to gain bearings

2.4 Victim locations known, go direct

2.5 Victims locations not known (daytime), room to room search

2.5.1 Search technique: left hand search follow walls

2.5.2 Children in the house, detail search under beds in closets as this is where children will go when scared

2.5.3 Just adults in house, less detail, usually found near exits

2.5 Victims locations not known (night time), straight to bedroom

2.6 Monitor oxygen levels

2.6.1 Observe lights on respirator

2.6.2 Know how long you have been so you know you how long you have to get out

2.7 Maintain SA on emergency egress routes in case of significant event (e.g. back draft)

3 Ventilation

3.1 Decide: ventilation can fuel fire with oxygen vs benefit of visibility

3.2 Break Windows for visibility

3.3 Vent enter search: Don't know if victims inside room

3.3.1 Crew chainsaws roof above fire seed

3.3.1 Search team waits until smoke pulled out of vent then moves in

3.3.1 No victims, fight fire

4 Rescue victims

3.1 Communicate with command

3.2 Immediately start dragging out, no first aid

5 Fight Fire

3.1 Driver becomes pump operator

3.2 Entire crew goes after fire

3.3 Locate fire

3.3.1 Feel for heat, too much heat, do not enter

3.3.2 Visual inspection for fire

3.6 Determine fire characteristics and strategy

3.6 Attack Fire

3.6.1 Very small fire: water direct on material on fire

3.6.2 Typical fire or room engulfed: deprive fire with steam by shooting short bursts at ceiling then close door, repeat

3.6.3 Keep together with partner, maintain SA on all fire fighters locations to avoid steam burns

3.6.3. Avoid tunnel vision, watch arcs

2.5 Monitor oxygen levels

2.5.1 Observe lights on respirator

2.5.2 Know how long you have been so you know you how long you have to get out

2.6 Maintain SA on emergency egress routes in case of significant event (e.g. back draft)

3.7 Burst open walls to search for fire in walls and insulation

3.8 Use thermal imager to detect remaining hot spots

6 Overhaul (post firefighting)

4.1 Fire watcher

4.1.1 Leave pump on scene in case of unexpected flare ups

4.2 Leave line out and ready

4.3 Rotate crews as time elapses

Appendix K

SA QUERIES GENERATOR

Note:

1. Lines indicate a piece of information (quantity, location etc…) that will be added by the experimenter based on the scenario.

2. Unknown will not be an option for level three queries as this answer effectively allows the participant to avoid the question.

					SA Questions (TRUE/FALSE/UNKNOWN)		
					Individual	**Generic**	**Team**
	Element	**SA Level**	**Information & Knowledge**				
Situational Elements	Structure	Facts	Number of stories			There is _____ stories to the house	1.We have searched the whole first floor 2.We have searched the whole second floor
			Attic				We have searched the attic
			Basement				We have searched the whole basement
			Quantity of rooms		I have searched ______ rooms	There are more than ________ rooms to search	1.My teammate has searched ________ rooms 2.As a team we have searched ________ Rooms 3.We have searched the whole house
			Fire hydrant present			There is a fire hydrant present	
			Address			The address of this house is _________	
			New or old construction			This appears to be a new construction	
			Location of Windows and doors			1. There more than ______windows and doors 2. The front door was locked	
		Meaning	Number of stories/ Quantity of rooms				This house is too big for the two if us to search
			Location of windows and doors			1. There are multiple egress routes should an emergency occur 2. A entry tool was required	

		Future	Quantity of rooms: to search				1.We will have enough time to search all rooms. 2. We will need to ask for back up.
			Location of windows and doors		Should an emergency occur, you would be able to get out of the house immediately (less than 10 sec) through a nearby window or door.		Should an emergency occur, your teammate would be able to get out of the house immediately (less than 10 sec) through a nearby window or door
	Victims	Facts	Location			A victim was found in the ______(loc)	My teammate found a victim______(loc)
			Health of victims: how they were found		I found the victim in ________(room) _________(not moving on the floor, on hands and knees, standing up)		My teammate found the victim in ________(room) _________(not moving on the floor, on hands and knees, standing up)
			Quantity		I have saved _________ victims		My teammate has saved ______ victims Together we have saved ______ victims
		Meaning	Health of the victim		1. The victim I found _______(loc) is ok. 2. The Victim I found_______(loc) is injured. 3. The victim I found _______(loc) is incapacitated		1. The victim my teammate found ________is ok. 2. The Victim my teammate found________is injured. 3. The victim my teammate found ________is incapacitated
		Future	Priority		The victim I found _______(loc) is priority for first aid		The victim my teammate found _______(loc) is priority for first aid
	Fire/Smoke	Facts	Smoke			As the engine approached the house, smoke was visible	
			Senses		I can hear the crackle sound of fire	As the engine approached the house, fire was visible	
			Location		I found a fire(s) in the _______(loc)		My teammate found fire(s) in the _______(loc)
			Quantity		I found _______(Qty) rooms on fire		1.My teammate found ______(Qty) rooms on fire. 2. There are _______(Qty) rooms on fire in this house.
			Fires extinguished		I have extinguished ______ fires		1. My teammate has extinguished _________ fires. 2. As a team we have extinguished _________ fires.
		Meaning	Quantity				There are too many fires for the two of us to put out
			Fires extinguished			There are still ________ fires to extinguish.	
		Future	Quantity/ Fires extinguished		I will have enough time to extinguish		We will have enough time to

					_________ more fires.		extinguish all of the fires.
	Assets	Facts	Engine Number			Your Engine number is _______	
			Officer			There is an officer present on the scene.	
			Ambulance			There is an ambulance present on the scene	
		Meaning	Ambulance		I should bring the victim to the ambulance		
		Future	Officer			Additional firefighters should be requested	
	Bystander	Facts	Present		There was a bystander in front of the house		My teammate saw a bystander in front of the house
			Who did the bystander talk too		The bystander spoke with me		The bystander spoke to my teammate
			Information Provided			The bystander said __________ (there is someone inside, no one inside, doesn't know if someone is inside)	
		Meaning	Information Provided			1. This is primarily a rescue mission 2. This is primarily a fire attack (extinguish fires) mission.	
		Future					
	Teammate	Fact	Location		I am located in _________(loc)		My teammate is in the same room
							My teammate is on the same floor
			Task		I am currently ________(task)		My teammates task is currently to___________
		Meaning	Location				My teammate is too far away
			Task		I am currently doing what I need to achieve the primary objectives		My teammate is currently doing what he needs to do to achieve the primary objectives
		Future	Location		I will be ______(loc) in one minute		My teammate will be ______ (loc) in one minute
			Task		I will be doing _______ (task) in 1 minute		My teammate be doing _______(task) in 1 minute

Appendix L

Additional Analysis

A team's WMC score was divided in several different ways during the analysis of the data. While the primary analysis was presented in the results section, this appendix displays the results of when the WMC variable was split at the median, into quartiles and when both teammates were above the median and both below the median. The following are the results between these WMC categorizations and TSA. WMC and Team Performance results based on the various categorizations of the WMC groups are not displayed here since none of the analyses achieved significance.

Median Split

Utilizing the SPSS rank cases function, the sample was split in half. As several teams fell on the same median score, the split was not even, the WMC high group had 18 teams while the WMC low group had 16 teams. The median split found the following results:

Total TSA. A significant Scenarios main effect was found, $F(2,64) = 6.086$, $p = .004$, $\eta_p^2 = .16$.

Ranch Scenario ($M = 10.472$, $SE = .287$)

Colonial Scenario ($M = 10.274$, $SE = .248$)

California Scenario ($M = 9.489$, $SE = .238$).

Pairwise comparisons showed a significant difference between the Ranch and California scenario, p = .009, and the Colonial and California Scenario, p = .038

A significant main effect of TSA levels, $F(2,64) = 18.536$, $p = .000$, $\eta_p^2 = .367$

Level 1 TSA ($M = 11.130$, $SE = .258$)

Level 2 TSA ($M = 9.938$, $SE = .293$)

Level 3 TSA (M = 9.148, SE = .251).

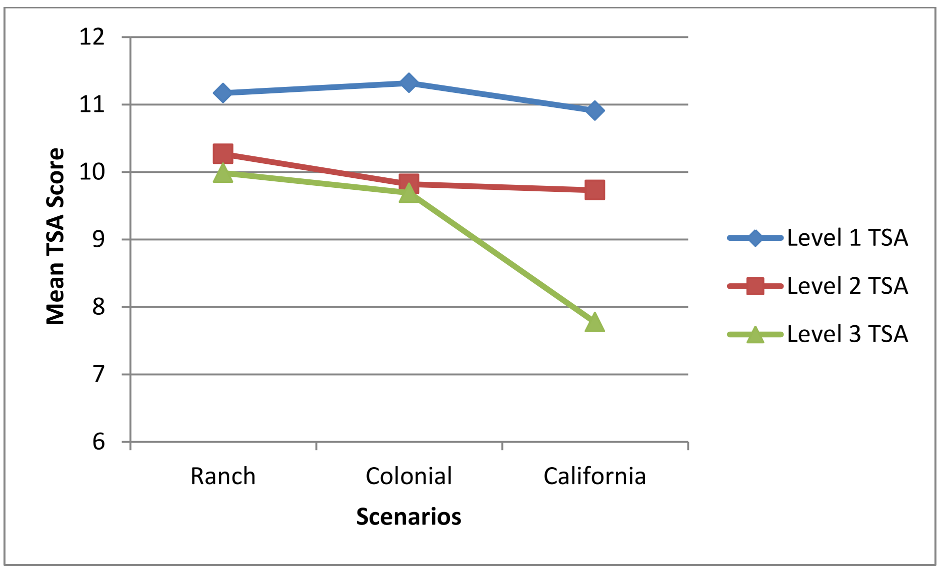
Pairwise comparisons showed a significant difference between the Ranch and California scenario, and the Colonial and California Scenario, both p = .000

A significant Main effect of WMC group, $F(1,32) = 10.457$, $p = .003$, $\eta_p^2 = .246$

Low WMC (M = 9.458, SE = .276),

High WMC (M = 10.685, SE = .260).

A significant interaction between Scenarios and TSA Levels, $F(4,128) = 2.652$, $p = .036$, $\eta_p^2 =$ .077



Team Queries. A significant main effect of TSA Levels $F(2,64) = 10.856$, $p = .000$, $\eta_p^2 = .253$.

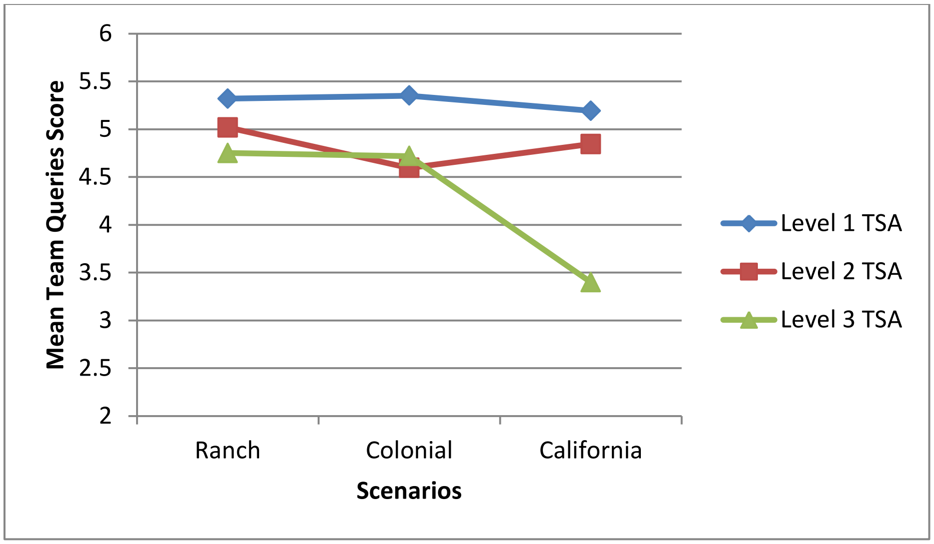
Level 1 TSA (M = 5.287, SE = .155)

Level 2 TSA (M = 4.818, SE = .182)

Level 3 TSA (M = 4.289, SE = .155)

Pairwise comparisons showed a significant difference between the Ranch and Colonial scenarios, p = .045, and the Ranch and California scenarios, p = .000

A significant interaction between Scenarios and TSA Levels, $F(2,64) = 2.878$, $p = .025$, $\eta_p^2 =$ 083



A significant Main effect of WMC group, $F(1,32) = 8.694$, $p = .006$, $\eta_p^2 = .214$

Low WMC ($M = 4.479$, $SE = .157$)

High WMC ($M = 5.117$, $SE = .148$)

Both Correct. A significant Scenarios main effect was found, $F(2,64) = 7.632$, $p = .001$, $\eta_p^2 = .193$

Ranch Scenario ($M = 3.839$, $SE = .194$)

Colonial Scenario ($M = 3.618$, $SE = .187$)

California Scenario ($M = 2.994$, $SE = .156$).

Pairwise comparisons showed a significant difference between the Ranch and California scenarios, p = .004, and the Colonial and California scenarios, p = .021

A significant main effect of TSA Levels, $F(2,64) = 18.713$, $p = .000$, $\eta_p^2 = .369$

Level 1 TSA ($M = 4.248$, $SE = .176$)

Level 2 TSA (*M* = 3.280, *SE* = .189)

Level 3 TSA (*M* = 2.924, *SE* = .173).

Pairwise comparisons showed a significant difference between the Ranch and Colonial scenarios, and the Ranch and California scenarios, both p = .000

Significant Main effect of WMC group, $F(2,32) = 12.607$, $p = .001$, $\eta_p^2 = .283$

Low WMC (*M* = 3.042, *SE* = .181)

High WMC (*M* = 3.926, *SE* = .171)

Quartile Split

Using a quartile split, 9 teams were in the high WMC group and 9 were in the low WMC group. The following results were found;

Total TSA. A significant main effect of TSA Levels, $F(2,32) = 16.165$, $p = .000$, $\eta_p^2 = .503$

Level 1 TSA (*M* = 11.5, *SE* = .308)

Level 2 TSA (*M* = 9.870, *SE* = .421)

Level 3 TSA (*M* = 9.167, *SE* = .327).

Pairwise comparisons showed a significant difference between the Ranch and Colonial scenarios, and the Ranch and California scenarios, both p = .000

A significant main effect of WMC group, $F(1,16) = 11.243$, $p = .004$, $\eta_p^2 = .413$

Low WMC (*M* = 9.309, *SE* = .367)

High WMC (*M* = 11.049, *SE* = .367)

Team Queries. Significant main effect of TSA Levels, $F(2,32) = 9.170$, $p = .001$, $\eta_p^2 = .364$

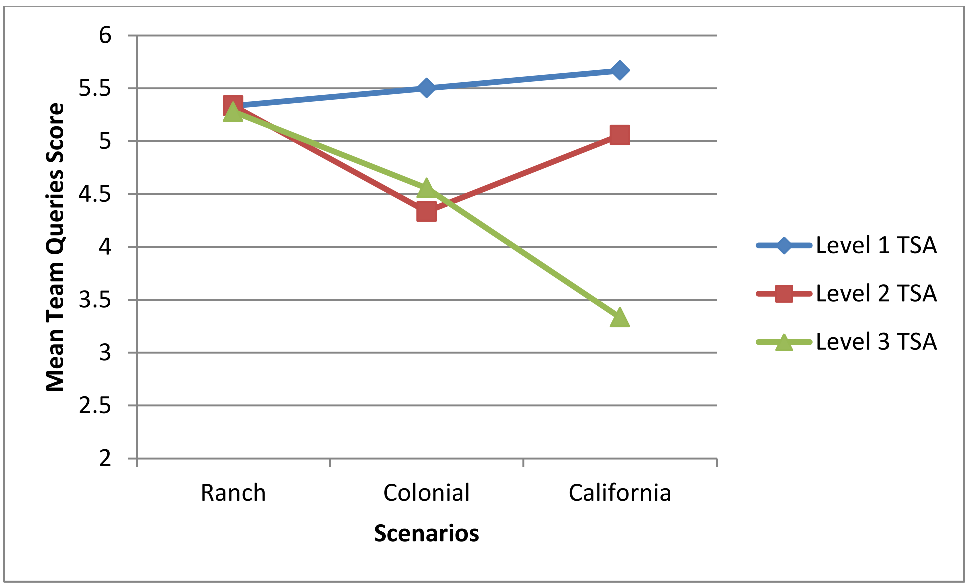
Level 1 TSA (*M* = 5.50, *SE* = .176)

Level 2 TSA (*M* = 4.907, *SE* = .238)

Level 3 TSA (*M* = 4.389, *SE* = .186).

Pairwise comparisons showed a significant difference between the Ranch and Colonial scenarios, p = .045, and the Ranch and California scenarios, p = .001

A significant interaction between Scenarios and TSA Levels, $F(4,64) = 3.746$, $p = .008$, $\eta_p^2 = .190$



A significant main effect of WMC group, $F(1,16) = 13.633$, $p = .002$, $\eta_p^2 = 46$

Low WMC (*M* = 4.432, *SE* = .192)

High WMC (*M* = 5.432, *SE* = .192)

Both Correct. A significant Scenarios main effect was found, $F(2, 32) = 3.377$, $p = .047$, $\eta_p^2 = .174$

Ranch Scenario (*M* = 3.889, *SE* = .309)

Colonial Scenario (*M* = 3.667, *SE* = .261)

California Scenario (*M* = 3.074, *SE* = .164).

Due to the Bonferroni adjustment, none of the pairwise comparisons were significant.

Significant main effect of TSA Levels, $F(2,32) = 15.181$, $p = .000$, $\eta_p^2 = .487$

Level 1 TSA ($M = 4.463$, $SE = .226$)

Level 2 TSA ($M = 3.296$, $SE = .253$)

Level 3 TSA ($M = 2.870$, $SE = .245$).

Pairwise comparisons showed a significant difference between the Ranch and Colonial scenarios, p = .001, and the Ranch and California scenarios, p = .001

Significant Main effect of WMC group, $F(1,16) = 13.382$, $p = .002$, $\eta_p^2 = .455$

Low WMC ($M = 2.926$, $SE = .239$)

High WMC ($M = 4.160$, $SE = .239$)

Both Teammates above the median (WMC High-high) vs. both below (WMC low-low)

Using only teams that had both teammates above the median (n = 7) and both below the median (n = 7), the following results were found;

Total TSA. A significant main effect of TSA Levels, $F(2,24) = 8.026$, $p = .002$, $\eta_p^2 = .401$

Level 1 TSA ($M = 11.357$, $SE = .373$)

Level 2 TSA ($M = 9.976$, $SE = .483$)

Level 3 TSA ($M = 9.310$, $SE = .408$).

Pairwise comparisons showed a significant difference between the Ranch and Colonial scenarios, p = .017, and the Ranch and California scenarios, p = .008

Significant main effect of WMC group, $F(1,12) = 5.343$, $p = .039$, $\eta_p^2 = .308$

Low WMC (M = 9.524, SE = .422)

High WMC (M = 10.905, SE = .422)

Team Queries. A significant main effect of TSA Levels, $F(2,24) = 3.454$, $p = .048$, $\eta_p^2 =$ .224

Level 1 TSA (M = 5.405, SE = .216)

Level 2 TSA (M = 4.952, SE = .271)

Level 3 TSA (M = 4.524, SE = .234).

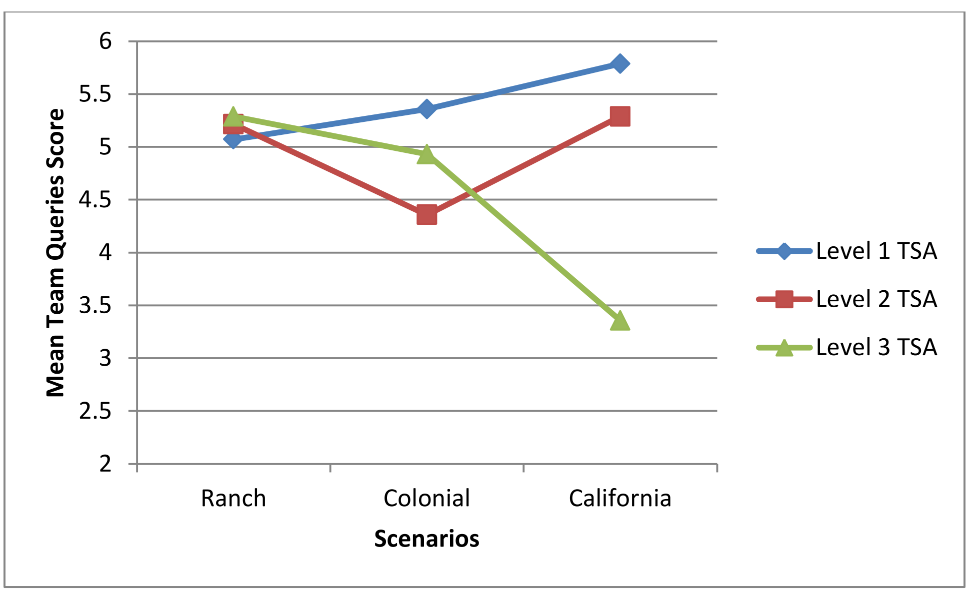
Pairwise comparisons showing a significant difference between Ranch and California, p = .045

Significant main effect of WMC group, $F(1,12) = 5.597$, $p = .036$, $\eta_p^2 = .318$

Low WMC (M = 4.619, SE = .204)

High WMC (M = 5.302, SE = .204)

A significant interaction between Scenarios and TSA Levels, $F(4,48) = 3.790$, $p = .009$, $\eta_p^2 = .24$



Both Correct. A significant Scenarios main effect was found, $F(2,24) = 3.597$, $p = .043$, $\eta_p^2 = .231$

Ranch Scenario ($M = 3.738$, $SE = .342$)

Colonial Scenario ($M = 3.929$, $SE = .287$)

California Scenario ($M = 3.071$, $SE = .181$).

Due to the Bonferroni adjustment, none of the pairwise comparisons were significant.

A significant main effect of TSA Levels, $F(2,24) = 7.654$, $p = .003$, $\eta_p^2 = .389$

Level 1 TSA ($M = 4.357$, $SE = .283$)

Level 2 TSA ($M = 3.310$, $SE = .283$)

Level 3 TSA ($M = 3.071$, $SE = .295$).

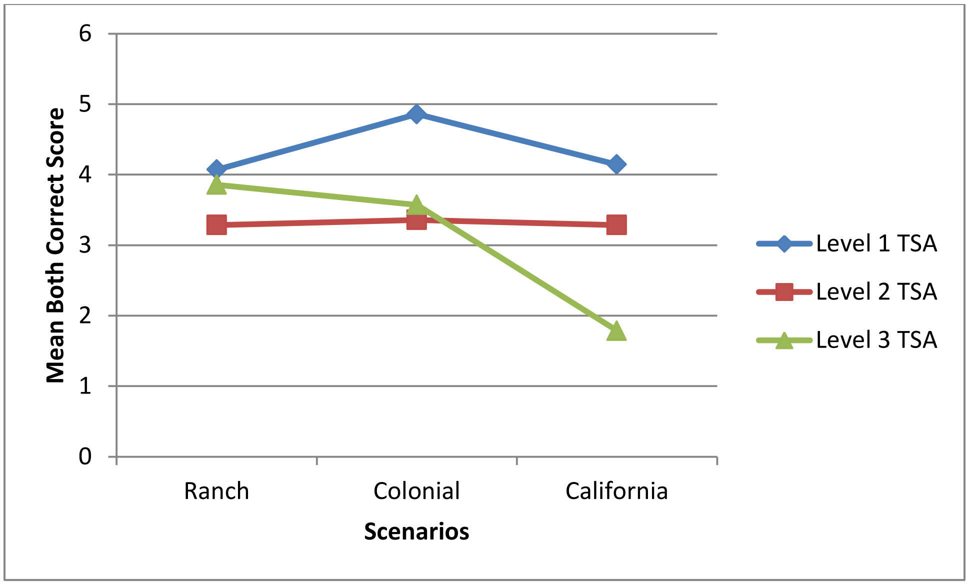
Pairwise comparisons showed a significant difference between the Ranch and Colonial scenarios, p = .016, and the Ranch and California scenarios, p = .023

A significant Main effect of WMC group, $F(1,12) = 7.103$, $p = .021$, $\eta_p^2 = .372$

Low WMC ($M = 3.048$, $SE = .282$)

High WMC ($M = 4.111$, $SE = .282$)

A significant interaction between Scenarios and TSA Levels, $F(4,48) = 2.651$, $p = .044$, $\eta_p^2 = .181$


6
5
4
3
2
1
0
Mean Both Correct Score
Ranch
Colonial
California
Scenarios
Level 1 TSA
Level 2 TSA
Level 3 TSA