

**Team Workload Questionnaire (TWLQ): Development
and Assessment of a Subjective Measure of Team
Workload**

A dissertation submitted in partial fulfilment of the requirements for
the degree of

Master of Science in Applied Psychology

at the

University of Canterbury

By

James Michael Sellers

2013

Research Supervisors:

Dr. William Helton

Dr. Katharina Näswall

Table of Contents

List of Tables and Figures	IV
Acknowledgements	V
1.0 Abstract	1
2.0 Introduction	2
2.1 Workload	4
2.2 Team Workload	5
2.3 Measurement of Team Workload	6
3.0 Study 1	10
3.1 <i>Taskwork</i>	10
3.2 <i>Teamwork</i>	12
3.3 Method	15
3.3.1 <i>Sample</i>	15
3.3.2 <i>Sampling Procedures</i>	15
3.3.3 <i>Measures</i>	16
3.3.4 <i>Procedure</i>	18
3.4 Results	19
3.5 Discussion	21
4.0 Study 2	25
4.1.1 <i>Unmanned Aerial Vehicle</i>	25
4.1.2 <i>Situational Awareness</i>	26
4.1.3 <i>Quantitative Analysis of Situational Awareness</i>	27
4.1.4 <i>Spatial Disorientation</i>	27

4.1.5 <i>Cognitive Map</i>	28
4.1.6 <i>Cohesion</i>	29
4.1.7 <i>Inclusion of Other in Self</i>	29
4.1.8 <i>Research Expectations</i>	30
4.2 Method	30
4.2.1 <i>Participants</i>	30
4.2.2 <i>Design</i>	30
4.2.3 <i>Materials</i>	31
4.2.4 <i>Measures</i>	34
4.2.5 <i>Procedure</i>	36
4.3 Results	37
4.4 Discussion	39
5.0 General Discussion	41
5.1 Practical and Theoretical Applications	43
5.2 Limitations and Future Research	44
5.3 Concluding Statement	45
6.0 References	46
Appendix A — Team Workload Measure	56
Appendix B — Stress Scale	59
Appendix C — Information Sheet and Consent Form	61
Appendix D — Team Workload Questionnaire (TWLQ)	63
Appendix E — Course Configurations	65
Appendix F — Stick Figures	67
Appendix G — Hazard Symbols	68

Appendix H — Guider Reference Sheet	69
Appendix I — Inclusion of Other in Self (IOS) scale	70
Appendix J — Quantitative Analysis of Situational Awareness (QUASA) scale	71
Appendix K — Cognitive Map	73
Appendix L — Experiment Instruction Sheet	74

List of Tables and Figures

<i>Table 1.</i> Rotated Pattern Matrix for Team Workload Items.	20
<i>Table 2.</i> Means, Standard Deviation and t-value of performance variables and workload scale measurements for Session 1 and Session 2.	38
<i>Table 3.</i> Correlations between Workload factors, Composite Performance score and IOS scale (Session 1 correlations are above the main diagonal; Session 2 correlations are below the main diagonal).	39
<i>Figure 1.</i> Representation of the experiment field and positioning of guider and pilot. Numbers represent the position of boxes; ‘H’ represents the helipad; and the two lines indicate the positioning of barriers.	32
<i>Figure 2.</i> Parrot AR. Drone 2.0 Quadrocopter, with hull and guard rings attached.	33
<i>Figure 3.</i> Ipad touch display showing live image from the drone downward facing camera with control overlay visible.	33
<i>Figure 4.</i> Parrot AR. Drone 2.0 being guided around the course.	33

Acknowledgements

I would like to thank everyone who participated or helped in any way with my research. Especially, the guidance and advice from my supervisors Deak Helton and Katharina Näswall.

Finally, I would like to thank my family for supporting me throughout my studies.

1.0 Abstract

The present research developed and assessed the Team Workload Questionnaire (TWLQ). Despite extensive workload studies, little research has been conducted on the workload experienced by teams. Team workload has largely been ignored by research with no validated theory constructed or dedicated team workload measures available to researchers and practitioners. The research was conducted in two studies with study 1 focusing on the development of the TWLQ with 216 members of sports team completing a team workload measure after games or practise. In study 2, 14 dyadic teams performed two sessions of an unmanned aerial vehicle (UAV) search-and-rescue task. The TWLQ was used to measure the team workload demand of the task. Study 1: Principle Axis Factoring method with Direct Oblimin rotation indicated three separate factors for the TWLQ with the factors classified as Task Workload, Team Workload, and Task-Team Balancing. Study 2: The TWLQ exhibited differential sensitivity, with the three factors measuring unique components of the workload demands in teams. The TWLQ is a valid and reliable subjective measure that can be used to assess the workload demand in team tasks. It provides researchers a tool to advance the understating of team workload and gives practitioners the means to assess the workload demands of team tasks in applied settings.

2.0 Introduction

Teams are a fundamental part of society, found in almost all areas of life (Bowers, Braun, & Morgan, 1997). In many settings team are responsible for highly critical decision making with high risk outcomes. Consider teams in medicine and aviation; the team interactions in these industries affect individuals on a daily basis. In medicine for example, surgeries require a team of doctors and nurses working together to ensure the safety of patients. The effectiveness of team work in operating theatres has a direct effect on the success of surgery, with more effective team work associated with reduced mortality rates (Neily et al., 2010). For aviation, team work is essential for the safe operation of aircraft, with 70% of accidents attributed to flight crew actions (Helmreich & Foushee, 1993). In aviation, effective team work provides a redundant system in which individual team members monitor and provide support for other team members, increasing the capability to identify safety threats (Wilson, Guthrie, Salas, & Howse, 2006). Generally, teams are effective, as they provide a broader pool of cognitive resources and task-relevant knowledge than individuals possess (Salas, Shuffler, & DiazGranados, 2010). However, when teams fail mistakes can be disastrous. An example of such is the destruction of Iran Air Flight 655 by the American missile cruiser, the USS Vincennes. The command and control team of the USS Vincennes mistakenly identified Flight 655 as a hostile plane and engaged it killing close to 300 individuals. The disaster was attributed to a breakdown in team processes (Urban, Bowers, Monday, & Morgan, 1995).

The use of teams in all settings has become more prevalent today as many work systems are simply too complex for individuals alone to operate (Bowers & Jentsch, 2005). Much of the complexity in work settings is a result of advancements in technology (Mathieu, Heffner, Goodwin, Salas, & Cannon-Bowers, 2000; Urban, Weaver, Bowers, & Rhodenizer,

1996). Technological advancements, while reducing physical demand has increased mental demand to levels that exceed the capability of individuals. As an example, Nuclear power plants have embraced automated systems, resulting in a shift in the function of the operator from a direct manual controller to a supervisor who monitors the system for anomalies and threats (O'Hara & Hall, 1992). The complicated nature of ensuring the safe operation of a nuclear plant means operators are flooded with information from roughly 1,000 system monitors on 45 displays (Lin, Hsieh, Tsai, Yang, & Yenn, 2011). Due to the vast amount of information that must be processed teams are employed to monitor the systems to ensure safe operations of the plant. However, many of the most serious accidents in Nuclear power plants are attributed to a break down among teams (Frye, 1988). It is such that, while teams possess more resources than individuals alone, team work in of itself adds additional workload above and beyond that which individuals experience (Bowers, Braun, & Morgan, 1992; Burke, Wilson, & Salas, 2003).

A considerable amount is known about teams with researchers having identified many of the inputs, process, and outcome variables that affect team performance (Burke et al., 2003), yet, there has been little research into the concept of workload among teams (Bowers & Jentsch, 2005; Funke, Knott, Salas, Pavalas, & Strang, 2012). With increasing mental demands in work settings an understanding of team workload is important. Unfortunately, team workload has largely been ignored with surprisingly little research regarding workload in team settings. It is surprising team workload has been so severely neglected when teams are primarily used in environments characterised by high levels of workload (Bowers et al., 1997). Therefore, with the increasing use of teams in work settings, it is important to learn more about team workload in order to ensure the safety, and effective performance of teams and those affected by team outcomes. A specific focus of this study is the development of a

subjective team workload measure which can be used to identify the workload demand on teams.

2.1 Workload

The theory underpinning team workload is derived from individual workload literature. Workload is a hypothetical construct which is generally conceptualised as the difference between an individual's available resources and a task's demands (Bowers and Jentsch, 2005; Young & Stanton, 2005). Research has consistently shown high levels of workload have a negative effect on performance (Bowers et al., 1997; Urban et al., 1995). To describe the mental workload phenomenon researchers often point to cognitive resource theories (e.g., Funke et al., 2012; Tsang & Vidulich, 2006; Young & Stanton, 2005). Cognitive resource theories propose mental workload to be a function of the supply and demand of cognitive or processing resources (Young & Stanton, 2005). Cognitive resources are hypothesised to be in limited supply, slowly diminishing during task performance (Norman & Bobrow, 1975; Tsang & Vidulich, 2006). Central to the concept of workload is the recognition that performance is in direct proportion to the resources dedicated to accomplishing a task (Bowers et al., 1997). Thus, when task demands increase individuals are required to input greater levels of resources. If however, the available resources are insufficient to meet task demands, the individual experiences high workload resulting in likely performance impairments (Funke et al., 2012) and associated distress (Helton, Matthew, & Warm, 2009).

Workload not only affects performance at high levels but also at low levels of task demand. In conditions where task demands are below a threshold, individuals may become complacent or bored; resulting in inattentiveness and reduced performance. In some cases individuals may make the task more difficult to increase the level of workload to a more

normal level (Young & Stanton, 2005); there is evidence of this in driving, with drivers increasing the task demands by driving faster when they are under low levels of workload (Zeitlin, 1995).

2.2. Team Workload

Team workload is broadly described as “an index of the ratio of available team resources to task demands” (Bowers & Jentsch, 2005, p. 57-1). It has also been characterised as “the relationship between the finite performance capacities of a team and the demands placed on the team by its performance environment” (Bowers et al., 1997, p. 90). Moreover, Funke et al., (2012) suggest team workload could be “characterized by a nonadditive relationship between finite performance capabilities of a team and the taskwork and teamwork demands placed on the team by its performance environment” (p. 38) or “characterized as a hypothetical construct that represents the linear aggregate cost incurred by all members of a team performance, which reflects the interactions of taskwork and teamwork demands and relevant individual characteristics” (p.38). From the proceeding, it is clear there is yet to be an agreed upon comprehensive and validated theoretical framework for team workload (Funke et al., 2012), with current definitions acting as tentative guides for the team workload construct. The challenges in developing a theoretical framework of team workload has been attributed to the complex nature of measurement in team settings and the difficulty in defining the dimensions of team work (Funke et al., 2012).

Each of the definitions presented above have a common origin in that they are adapted from individual workload theory. This is a major flaw in the team workload construct as researchers are unsure of the relationship between individual workload and team workload (Funke et al., 2012). While workload at the individual level is relatively straight forward, at the team level workload is more complex with it described as being “multiplicative or

nonlinear” (Funke et al., 2012, pg. 38). However, as it is believed individuals have a finite supply of resources, so too is it believed that teams have a finite supply of resources which diminish during tasks (Barnes et al., 2008). While teams by their nature possess more resources than individuals alone (Levi, 2007); there are additional demands inherent to teams that elevate team workload above that of an individual (Bowers, et al., 1992). Researchers believe that team members, in addition to experiencing a level of individual workload, are also loaded with workload brought about by team interactions (e.g., communication, coordination). It is herein why the team environment is likened to a dual-task performance situation with team members required to perform taskwork (i.e., individual efforts to meet task demands) and teamwork (i.e., cooperative efforts of team members required for effective team performance) simultaneously (Bowers et al., 1997; Burke et al., 2003; Funke et al., 2012). In addition to performing taskwork and teamwork, team members must also assign resources to manage the interaction between taskwork and teamwork (Bowers et al., 1997). In other words they must assign resources to ensure the demands of taskwork do not interfere with the demands of teamwork and vice versa. It is this factor which is hypothesised to elevate the workload experienced by the teams above that of simply the sum of individual team members (Bowers et al., 1997; Bowers & Jentsch, 2005; Gopher & Braun, 1984).

2.3. Measurement of Team Workload

Without an agreed upon theoretical framework of team workload, researchers have found it difficult to develop effective team workload measures. Bowers et al. (1997) identified many of the measurement questions facing team workload researchers, explicitly,

[A]t what level of the team does one assess team workload? How should the data from a team of individuals be combined and interpreted? Should team members be

required to rate their subjective perception of overall team workload? Can team workload be defined as the total (or average) of the individual levels of workload experienced by team's members? What dimensions or components of team workload should be measured, and how should these measures be combined? (p.101)

In the years since Bowers et al. (1997) posed these questions there has been little progress in addressing the measurement issues facing team workload. Not only is this concerning for the assessment of team workload in applied settings, but theory-wise, until a psychometrically sound measurement tool is developed, team workload theory will be stuck in a conceptual stage (Baker & Salas, 1997).

Workload is predominately measured in three ways, namely, subjective measures (e.g., self-report scales), performance measures (e.g., primary and secondary tasks), and physio-behavioral measures (Funke et al., 2012; Rubio, Diaz, Martin, & Puente, 2004). Much like the theory of team workload, the methods for measuring team workload are largely an extension of individual workload methods. Funke et al., (2012) questions whether researchers use individual workload measures to assess team workload because of the theoretical connection between the concepts or because of a lack of specific team workload measures, meaning researchers have no other option.

Subjective measures are by far the most commonly used method for measuring team workload, largely due to their ease of implementation and non-intrusiveness (Rubio et al., 2004). The measures are either applied directly from individual workload measures or are modified by altering the instruction set, item set, or both to make them more applicable to teams (Bowers & Jentsch, 2005). At the individual level subjective measures have been shown to be sensitive to differing levels of workload, and generally correlate well with other measures of workload (Rubio et al., 2004).

Performance measures are effective measures of individual workload but have only rarely been used with teams (Bowers & Jentsch, 2005). There are two types of performance measures: primary task, and secondary task measures (Funke et al., 2012). The most effective is secondary tasks performance measures which require an operator to perform primary and secondary tasks simultaneously (Tsang & Vidulich, 2006). Secondary task measures of workload are inherently suited to measuring the conceptual framework of workload, in that, they assess whether an operator has available resources to perform additional tasks. If an individual is under a low level of workload they have the additional resources to attend to the secondary task; in comparison, individuals under high task level simply do not have the resources to respond effectively to the secondary task. Performance measures can be difficult to interpret in the team situation because they do not represent the differences in resource inputs of team members (Funke et al., 2012).

The final method for examining workload are physio-behavioral measures, including processes such as echocardiography (ECG), eye tracking, electroencephalography (EEG), functional magnetic resonance imaging (fMRI), and transcranial doppler ultrasonography (TCD; Funke et al., 2012). These measures are relatively successful in the individual setting but when used in the team setting have been found to interfere with team processes (Funke et al., 2012).

Performance measures and physio-behavioral measures while valid in research settings are in many cases not applicable for use in real world settings where it is simply not possible to rig individuals up to a machine or safe to overload operators with additional tasks. As such, it is essential to develop a subjective team workload measure that can be used in applied settings so that work environments and teams can be designed to facilitate effective performance. The present study will be conducted in two parts. Study one will be concerned

with the development of a subjective measure of team workload. A team workload scale will be developed by examining workload and team literature, to create a theoretically sound measure of team workload. The second study will take the workload measure developed in study one, and experimentally test the scale to analysis the sensitivity and usefulness of the scale.

3.0 *Study 1*

Working on the basis that team performance situations consists of taskwork and teamwork, the scale needs to be developed to address each of these components.

3.1 Taskwork

The National Aeronautic and Space Agency – Task Load Index (NASA-TLX; Hart & Staveland, 1988) has been used in the majority of team workload experiments, to measure task workload, in either its original form or a modified form (Funke et al., 2012). The NASA-TLX is an effective measure of workload, which has been shown to be sensitive to different levels of workload across a number of settings (Nygren, 1991; Rubio et al., 2004). The NASA-TLX consists of six scales that represent independent groups of variables: mental demand, physical demand, temporal demand, frustration, effort, and performance. The assumption is that a combination of these dimensions is likely to represent the workload experienced by individuals (Hart & Staveland, 1988). Since its introduction the NASA-TLX has been widely used in research with few questions raised about its validity. However, recently some researchers have questioned whether measures of frustration and performance should be included in assessments of workload as they are not conceptually related to workload (Bailey & Thompson, 2001; Ramiro, Valdehita, Lourdes, & Moreno, 2010). Frustration, for instance, is an evaluative response to a task, not an indication of the level that one has worked. Likewise, performance is a largely subjective assessment of how well one has completed the demands of a task. For example, take two gamblers playing at a casino, if one were to win and the other were to lose they would report vastly different assessments of performance and frustration, yet the actual workload levels would be relatively similar (assuming neither player was cheating).

Instead of Frustration a more appropriate construct is Emotional Demand. Whenever a task is tied to important goals, an emotional reaction is generated (Austin & Vancouver, 1996). Emotional responses are both positive and negative encompassing a range of different reactions. Cognitive theories propose emotional responses are related to, (1) an appraisal of the emotional threat of the environment or situation, and (2) an appraisal of the available resources required to regulate the emotional demands of the situation, (Lazarus, 1982). Thus, when an individual views a task as being emotionally threatening and perceives they lack the resources to self-regulate their emotions, the task is viewed as being more emotionally demanding (Zohar, Tzischinski, & Epstein, 2003). It is the cognitive demand of controlling emotions that makes Emotional Demand a more appropriate assessment of workload than Frustration; with frustration simply the display of one's negative emotion when they do not have the necessary resources to regulate the emotional demands of a task. Furthermore, Frustration does not take into account the numerous other emotional reactions individuals go through when performing tasks, being too narrow in scope.

Instead of Performance a more appropriate construct is Performance Monitoring Demand. Dickinson and McIntyre (1997) describe performance monitoring as the observation and awareness of one's activities and performance. Monitoring and adjusting of performance is essential in all tasks. To achieve task goals individuals must be aware of their current performance to know if they must adjust their behaviour. When individuals engage in performance monitoring they must use cognitive resources to evaluate their performance (Porter, Gogus, & Yu, 2010). When examining performance in high pressure situations, DeCaro, Thomas, Albert, and Beilock (2011) found that performance in tasks that rely on working memory and attention are negatively affected when individuals are required to ensure they are performing at certain levels. This suggests the demand of monitoring

performance uses resources that would otherwise be dedicated to task demands, and will thus be associated with high task workload.

3.2 Teamwork

There have been a number of different approaches to the development of the teamwork component of team workload measures. Researchers have tried diverse combinations of teamwork characteristics in an attempt to create comprehensive measures. The absence of a unified approach is due to the fact that an agreed upon taxonomy of teamwork does not exist, leading to difficulties defining the characteristics that comprise effective teamwork (Bowers et al., 1997; Porter et al, 2010). While it is unlikely a comprehensive taxonomy of teamwork will be developed in this study, a review of many of the existing team workload measures indicates a common group of items, including: communication, coordination, time share demand, support, performance monitoring, control demand, team success, and team frustration (Funke et al., 2012; Lin et al., 2011; Porter et al., 2010; Yang, Yenn, & Lin, 2010).

The role of coordination and communication in teamwork is clear, put simply, teamwork cannot occur without coordination and communication (Bowers et al., 1997; Jentsch & Bowers, 2005). Coordination is a broad classification of the behaviours performed by team members to achieve common goals (Bowers et al., 1997). In the workload context it is primarily concerned with the process of adjusting team activities (e.g., plans, actions, responsibilities, etc.), to ensure that all team members are working towards the same goal. The demand for coordination is cognitively demanding (Ford & Schmidt, 2000), and can significantly impact performance (Chiocchio, 2007). The performance impact of coordination activities was demonstrated in a study by Stout, Cannon-Bowers, Salas, and Milanovich

(1999). The researchers found teams which were given more time to plan for a task, thus requiring less coordination activities during the task, perform significantly better than teams that had less time planning. The demand of coordinating in teams uses resources that would otherwise be dedicated to taskwork.

Communication is essential for the successful completion of all team activities. Like coordination demands, the requirement of communication with team members produces considerable cognitive demands on an individual (Bowers et al., 1997). Communication requires processing resources, in both the sending and receiving of the message which is cognitively demanding. In experienced teams communication mediates the effects of workload (Urban et al., 1995). However for inexperienced teams communication can become a detriment to performance, with too much communication overloading the processing capabilities of individuals (Lin et al., 2011), and thus increasing workload.

The belief that team environments require a time share between taskwork and teamwork is well established. Without wanting to repeat much of what has already been mentioned on the topic, the demands of managing resources between teamwork and taskwork is cognitively taxing. The greater the demand for managing the requirements of taskwork and teamwork the more difficult the team environment becomes and thus the greater the workload.

Support and team performance monitoring have often been viewed as essentially the same construct; however, they are fundamentally different and should be viewed separately (Porter et al., 2010). Support demand is concerned with the providing and receiving of support between team members. It requires team members have an awareness of other team members' roles and are willing and able to provide and seek support, intervening when required (Porter et al., 2010). Support therefore requires team members allocate both

cognitive (e.g., be aware that a team member is struggling) and physical resources (e.g., intervene in the taskwork; Marks, Mathieu, & Zaccaro, 2001; Porter et al., 2003). Team monitoring, in comparison, requires regular observation and processing of the performance of other team members (Serfaty, Entin, & Johnston, 1998), requiring only the use of cognitive resources (Marks et al., 2001) to monitor the performance of others.

In teams individuals experience a range of emotional responses because teams are characterised by complex interpersonal exchanges (Tse, & Dasborough, 2008). Much like the emotional demand of taskwork, emotional reactions in teams are a function of one's perception of the environment and their ability to regulate their emotions (Compo, Mellalieu, Ferrand, Martinent, & Rosnet, 2012) However, unlike when working alone, the actions of team members can impact the emotional state of other members (Totterdell, 2000). For example, one team member's actions may place another team member in a difficult situation, resulting in a negative emotional reaction. With this comes a perceived reduction in the control of the situation (Mann, Williams, Ward, & Janelle, 2007). This lack of control is viewed negatively and if coupled with a perceived inability to regulate the emotional demands an emotional response will be exhibited with high levels of emotional demand reported.

Other plausible teamwork components of workload measures are constructs related to team success and team frustration, analogous to the NASA-TLX's individual items of Performance and Frustration. While as discussed earlier the constructs of performance and frustration are not conceptually related to workload, they are included due to the NASA-TLX's widespread popularity (e.g., tradition). Similarly, measures of team success and team frustration are included to assess whether they duly load, appropriately, in a factor analysis on team workload factors.

3.3 Method

3.3.1 Sample

Two hundred sixteen (179 men; 37 women) members of sports teams participated in the study. Their average age was 21.9 ($SD = 2.5$) years, with age ranging from 18 to 32 years. All participants were members of amateur club teams from Christchurch, with the level of competition ranging from social sport to Senior Division One competition. Of the sample 78.2 percent were from rugby teams, and 21.8 percent mixed sex touch rugby teams.

3.3.2 Sampling Procedures

Purposive sampling was undertaken, as the study required data was collected from teams. Sports teams were chosen as they provided the greatest access to numbers of participants required for valid statistical procedures. Sports teams provide a suitable sample in which to explore team workload as sports teams typically have been formed for a long time, success or failure can be clearly determined, and the outcomes are meaningful to the individuals.

Participants were recruited by contacting sporting clubs throughout Christchurch. Eight rugby clubs were contacted, six of which gave permission to approach their teams; and one touch rugby organisation was contacted who gave permission. Of the six rugby clubs which granted permission to approach their teams, a total of 22 teams were approached, 14 of which gave permission to survey their players. Eight touch teams were approached, five participated in the study. Therefore, the response rate of teams approached was 63%.

Clubs were contacted through telephone calls which explained the reason for the call and purpose of the study. Those clubs which were happy to participate were sent a copy of

the questionnaire, to be used, and an information sheet detailing the study. Once club managers gave their approval they were asked to supply contact numbers for team managers or coaches. Again, the managers and coaches were contacted by telephone and it was explained why they were contacted and the purpose of the study. For the teams that agreed to be surveyed a time was set to come along to administer the questionnaire. Data was collected at team's trainings, for rugby teams; while it would have been preferred to collect data after competition games, initial enquires suggested players and teams were resistant to the idea. To ensure the data collected was as close as possible to a game situation the surveys were administered after team run-throughs. Team run-throughs are essentially a simulated game that is as representative of a game one could find without being the real event. None of the touch rugby teams approached had training sessions; therefore, the data was collected after their competitive games. Participation was rewarded by entry into a draw to win an Apple iPad2.

3.3.3. Measures

Team Workload Measure.

A number of demographic variables were collected, these included gender, age, position in the team and years of experience in the team.

The team workload measure (Appendix A) was developed in accordance with current conceptualisation of team workload. To measure the taskwork component an adapted version of the NASA-TLX was used. The six scales of the NASA-TLX: Physical Demand, Mental Demand, Temporal Demand, Effort, Frustration, and Performance, were all included. While, as discussed earlier, a number of researchers have questioned the validity of performance and frustration as a measure of workload, both were kept for exploratory reasons, as a goal of this

study is to generate a greater understanding of team workload. In addition to the six scales of the NASA-TLX two additional scales were added. The scales were Emotional Demand and Performance Monitoring Demand.

In line with the literature regarding emotion and performance monitoring a group of potential items were developed for each scale. The items were then reviewed by the researcher and two other independent individuals. The items were reviewed to determine whether, in the eyes of the independent reviewers, they made sense and if they represented the proposed construct. The items were then discussed as to their merits for inclusion, with each item undergoing thorough examination with suggestions for changes and exclusion. Ultimately the following items were included in the taskwork component of the team workload questionnaire:

Emotional Demand – How much did you have to control your emotions (e.g. anger, joy, disappointment)?

Performance Monitoring Demand– How much did the task require you to monitor your performance (i.e., ensure you were performing at specific levels)?

The teamwork component of the scale consisted of eight items, which were considered after a review of the literature. The items were developed following the same procedures as discussed for the creation of taskwork items. The items developed were:

Communication Demand – How much communication activity was required (e.g. discussing, negotiating, sending and receiving messages, etc.)?

Coordination Demand – How much coordination activity was required (e.g. correction, adjustment, etc.)?

Time Share Demand - How difficult was it to share and manage time between task-work (work done individually) and team-work (work done as a team)?

Team Effectiveness - How successful do you think the team was in working together?

Team Support - How difficult was it to provide and receive support (providing guidance, helping team members, providing instructions, etc.) from team members?

Team Dissatisfaction - How irritated and annoyed were you with your team?

Team Emotion Demand - How emotionally demanding was working in the team?

Team Performance Monitoring Demand - How much did the task require you to monitor your team's performance?

Items are measured on an 11-point scale. The response scale ranges from 0 (very low) to 10 (very high), with a high score indicating a higher level of subjective workload. Values are multiplied by a factor of 10 so that each item will result in a rating ranging between 0 and 100.

Also included was an 11 item stress scale (Appendix B). The stress scale was included to further evaluate the measurement properties of the team workload measure by assessing the respondent's state at the time of taking the team workload questionnaire. The stress scale asked respondents to indicate aspects such as motivation, focus, happiness, and interest.

3.3.4. Procedure

The team workload measure, stress scale, information sheet and consent form (Appendix C) were distributed to players at the end of team run-throughs, for rugby, and at

the end of games for touch players. Players completed the paper questionnaires using pens or pencils. The questionnaires were completed independently, generally taking no more than five minutes to finish. When players finished the questionnaire they handed it back to the experimenter.

3.4 Results

Before conducting factor analysis the data were assessed to determine suitability for factor analysis. The Kaiser-Meyer-Olkin measure of sampling adequacy was .73, above the recommended value, and Bartlett's test of sphericity was significant ($\chi^2(120) = 878.08, p < .001$). Based on the results of the tests it was concluded factor analysis was appropriate. Exploratory factor analysis using Principle Axis Factoring method with Direct Oblimin rotation was undertaken. Direct Oblimin rotation was employed as correlation analysis indicated the factors were intercorrelated. Factors were considered based on eigenvalues greater than 1.0 and visual inspection of eigenvalue-plots. Furthermore, individual items were kept if they loaded above .40 on a single factor and did not exhibit serious cross loading (i.e., $> .30$).

Principle Axis Factoring method indicated a three factor model accounting for 57.80% of the variance (Table 1). The eigenvalues were all greater than 1.0 and the eigenvalue-plot indicated an elbow at the forth factor. The first factor accounted for 26.32% of the variance, the second factor accounted for 17.70% of the variance and the third factor accounted for 13.78% of the variance. Four and five factor solutions were also examined; however, the three factor model provided the most theoretically sound representation of the team workload construct.

Table 1. Rotated Pattern Matrix for Team Workload Items

	Factors		
	1	2	3
Physical Demand	.574		
Mental Demand	.734		
Temporal Demand	.595		
Emotional Demand	.446		
Performance Monitor Demand	.653		
Effort	.680		
Communication		.808	
Coordination		.858	
Time Share Demand			.738
Team Emotion Demand			.559
Team Performance Monitoring Demand		.448	
Team Support			.698

During several steps a number of items that did not meet the requirements were removed. Performance and Frustration did not load above .40 on any factors and were thus removed. Team Effectiveness loaded above .40 on a factor but also loaded above .30 on a second factor, thus it was also removed. Team Dissatisfaction loaded above .40 on a single factor, however, exhibited some cross loading with a factor value of .22 on a second factor. While not serious enough to suggest it should immediately be removed, theoretically, Team Dissatisfaction does not match up with the construct of workload, therefore, more analysis was undertaken. Scores on the Team Dissatisfaction item were correlated with the stress items to determine if there was evidence to suggest dissatisfaction is a reaction towards the task, rather than a cause of workload. Pearson's correlation analysis indicated Team Dissatisfaction to be significantly correlated with 10 of the 11 stress items. Team Dissatisfaction was strongest when individuals reported higher levels of unhappiness ($R = .315, p < .001$), were less focused on the task ($R = -.312, p < .001$) and thought about them self more often ($R = .269, p < .001$). The relationship to the stress items coupled with results

from the factor analysis and the theoretical concept of workload lead to the decision to remove Team Dissatisfaction from the final model.

The three factor model consisted of 12 items. The factors were classified in accordance with past theoretical work. The first factor was classified as *Task Workload* and consisted of the scales: Mental Demand, Physical Demand, Temporal Demand, Effort, Emotional Demand and Performance Monitoring Demand. The second factor classified as *Team Workload* and consisted of the scales: Communication Demand, Coordination Demand, and Team Performance Monitoring Demand. The third factor classified as *Task- Team Balancing* and consisted of the scales: Time Share Demand, Team Emotion Demand, and Team Support.

Internal consistency for each of the scales was examined using Cronbach's alpha. The Cronbach alpha was moderate for all three factors, Task Workload, $\alpha = .783$, Team Workload, $\alpha = .739$, and Task-Team Balancing, $\alpha = .692$.

3.5 Discussion

The Team Workload Questionnaire (TWLQ; Appendix D) was developed after analysing the team workload demand in team sports. The 12 scale questionnaire measures three unique components of team workload, with workload in teams consisting of Task Workload, Team Workload and Task-Team Balancing. Overall the items loaded as one would expect with the factors making conceptual sense (Bowers et al., 1997).

Task Workload.

The Task Workload scale of the questionnaire consists of the traditional NASA-TLX scales, Mental Demand, Physical Demand, Temporal Demand and Effort, as well as two new scales, Emotional Demand, and Performance Monitoring Demand. All of the items are theoretically related to taskwork demands and formed as was expected. Interestingly, the results indicated the other traditional NASA-TLX scales, Frustration and Performance, are independent constructs not related to workload. These results support those found by a number of researchers (Bailey & Thompson, 2001; Ramiro et al., 2010), and further advance the belief that frustration and performance are a reaction to a task rather than a cause of high workload. With the increasing evidence regarding frustration and performance, researchers should cease their use in aggregate measures of workload. Instead, researchers should consider including Emotional Demand and Performance Monitoring Demand. The two scales created in the study both appear to account for unique demand associated with taskwork. Conceptually this makes sense as both require the use of cognitive resources that would otherwise be used for performing task demands.

Team Workload.

The Team Workload aspect of the questionnaire is concerned with the demands of team interactions, specifically communicating, coordinating and the monitoring of team performance. Communication and coordination are undoubtedly connected and are essential for teamwork (Bowers et al., 1997). The new scale developed, Team Performance Monitoring was also shown to be a demand of teamwork. The relationship of team performance monitoring to communication and coordination is logical. To coordinate one's behaviour to match that of the team they must be aware how other team members are performing. By monitoring the performance of teammates, teams can efficiently evaluate their performance

and adjust planes accordingly. Effective communication is also reliant on understanding the needs of teammates. Teams that are most effective are efficient in exchanging information. They engage in less question-and-answer communication but rather anticipate the needs of others and provide the required information (Urban et al., 1995). As such it appears Team Performance Monitoring is a unique measure of the demands associated with teamwork and should be included in future team workload measures.

Task-Team Balancing.

Task-Team Balancing is concerned with the management of taskwork demands and teamwork demands. The factor analysis results suggest this involves Time Share Demand, Team Support, and Team Emotion Demand. Seemingly, Time Share Demand is a function of the level of Team Support required for a task with support adding to the requirements of allocating time between taskwork and teamwork. The inclusion of Team Emotion Demand in Task-Team Balancing is interesting, suggesting high demands for task-team balancing are also emotionally demanding. Presumably, much of the emotional demand is related to providing and receiving support. Individuals may become annoyed they have to perform their own taskwork as well as support their teammates. Conversely those individuals who seek support may feel ashamed or embarrassed they need support from their teammates. In both instances the actions of other teammates leads to an emotional evaluation of the environment, with a perceived loss of control over their own performance viewed negatively. For individuals asked to provide support, if they are already under high workload they may not possess the resources required to regulate emotional reactions and react emotionally to requests for support. In the case of individuals seeking support they are already lacking the resources to fulfil their taskwork demands and are unlikely to have necessary resources to regulate the emotional demands, likely becoming annoyed and frustrated.

Performance and Frustration.

The results from this study do not support the inclusion of items related to performance and frustration for the assessment of taskwork or teamwork. In addition to excluding the traditional NASA-TLX items Performance and Frustration the results also suggested excluding measures of team effectiveness and team dissatisfaction in team workload questionnaires. This reinforces the point that assessments of performance outcomes and frustration should be excluded from aggregate measures of team workload. While performance outcomes and frustration may be somewhat related to team workload, they are theoretically distinct constructs.

4.0 Study 2

The development of a theoretically sound team workload questionnaire that can be used for the assessment of team workload is the primary purpose of this research. Therefore, it is important to take the TWLQ, and test it in an applied setting to determine its measurement properties.

4.1.1 Unmanned Aerial Vehicle (UAV)

To test the TWLQ a team task was designed using a UAV. The use of UAVs or ‘Drones’ to test the TWLQ is applicable as research has shown piloting UAVs to be cognitively taxing (Guznov, Matthews, Funke, & Duke, 2011) with operators having to deal with stress, fatigue, and high workload, (Hancock, Mouloua, Gilson, Szalma, & Oron-Gilad, 2007). Furthermore, the use of a synthetic task environment, similar to the task developed in the present study, is considered to be a valuable tool for assessing stress and workload in UAV operation (Guznov et al., 2011).

UAVs are an aircraft without a human pilot on board, instead being controlled autonomously by computers or remotely by a pilot on the ground. UAVs have been extensively used by militaries, mainly for surveillance and attacking enemies (Everaerts, 2008). However, with advancements in technology they are becoming more accessible and are being used by the civilian sector for surveillance, fire-fighting, monitoring crops and wildlife, and search-and-rescue (Benini, Mancini, Minutolo, Longhi, & Montanari, 2012; Guznov et al., 2011). In search-and-rescue situations the use of UAVs has numerous advantages over traditional search-and-rescue methods. First and foremost, UAVs provide rescue teams the ability to explore environments that are too dangerous for people or dogs

(Burke & Murphy, 2004). Also, compared to fixed wing search-and-rescue aircraft, UAVs are more cost efficient (Drury, Richer, Rackliffe, & Goodrich, 2006) and can spend more time in the air searching (Goldberg, 2010), While useful, flying UAVs is difficult as indicated by the fact that UAVs suffer more accidents per 1,000 flight hours than manned aircraft (Drury et al., 2006). The difficulty in flying UAVs is attributed to reduced situational awareness with the use of UAVs likened to exploring environments through a peephole (Burke & Murphy, 2004).

4.1.2 Situational Awareness

Situational Awareness is defined as the “perception of the elements in the environment within a volume of time and space, the comprehension of their meaning, and the projection of their status in the near future” (Endsley, 1988, p. 97). Situational awareness is considered to be the creation of a mental model of the environment, with the model built through available environmental information, and existing cognitive bias (Endsley, 2001; Menda et al., 2011). The mental model that individuals hold guides their decision making and actions (Menda et al., 2011); with the appropriateness of decisions and actions in direct relation to the accuracy of the mental model (Hendy, 1995). In relation to teams, team situational awareness reflects the coordinated awareness that the team possesses as a whole unit (Salmon, Stanton, Walker, Green, 2006). It is not simply the combined situational awareness that individuals hold but rather the degree to which team members have the same interpretation of the environment, with a shared mental model (Gorman, Cooke, Pederson, Connor, 2006; Salmon et al, 2006).

The Development of situational awareness when flying UAVs is difficult as pilots often have to rely solely on information provided by the UAV (e.g., systems, cameras) which

can lack contextual awareness. Research exploring the use of unmanned vehicles for search-and-rescue found rescue workers participating in a ground-based robot experiment spent approximately 30% of the time trying to gain or maintain situational awareness, by trying to understand the robot's location, surroundings, and status (Yanco and Drury, 2004). In another ground-based robot experiment, Burke, Murphy, Coovert, and Riddle (2004) found that operators spent more time collecting information regarding the positioning of the robot and the state of the environment than they did performing the task.

4.1.3 Quantitative Analysis of Situational Awareness

To measure the situational awareness of individuals in the study, the quantitative analysis of situational awareness (QUASA; McGuinness 2004) technique will be used. The QUASA is based on the belief that the extent of an individual's knowledge of a situation can be indicated by his or her ability to judge the truth or falsity of a proposition related to the task (Ebel and Frisbie, 1991). The QUASA technique consists of true/false questions, whereby the individual is presented with a set of statements regarding the situation and asked to indicate in each case whether the statement is true or false.

4.1.4 Spatial Disorientation

Linked to situational awareness is the concept of Spatial Disorientation. Spatial disorientation is a pilot's failure to correctly sense the position, motion, or attitude of their aircraft in relation to the surface of the earth (Benson, 1999). Spatial disorientation is divided into three types: Type I (unrecognised), when the pilot does not know they are disorientated and continues as if all is normal; Type II (recognised), when the pilot is aware of erroneous

orientation information, but can still control the aircraft; Type III (incapacitating), the pilot is aware they are disoriented but are so confused they become afraid and freeze (Previc & Ercoline, 2004; Webb, Estrada III, & Kelley, 2012). When a pilot suffers spatial disorientation they lose situational awareness, as they have lost the spatial cues that are informing their mental model of the situation. However, a loss of situational awareness does not mean the pilot suffers spatial disorientation, as they can lose their geographical orientation without becoming spatially disorientated (Previc & Ercoline, 2004).

Spatial disorientation is perhaps more of an issue in UAVs than manned aircraft as pilots lose many of the senses that warn of potential danger. UAV pilots detached from the aircraft do not get to hear ambient noise, or feel the sensation of vibrations or changes in pitch and yaw (Mola, 2008). In addition pilots do not receive feedback from the vestibular system. Vestibular information is derived from organs within the inner ear, the semicircular canals, and the otolith organ; movement of the body through space is sensed through the vestibular system so as to allow spatial orientation to be maintained (Williams, 2008).

4.1.5 Cognitive Map

To measure the level of spatial disorientation, or spatial orientation, of teams during the experiment, a cognitive map task will be used which requires individuals to indicate on a representative map where they believe the drone was located in relation to the experiment field.

4.1.6 Cohesion

Cohesion is “a dynamic process that is reflected in the tendency for a group to stick together and remain united in the pursuit of its instrumental objectives and/or for the satisfaction of member affective needs” (Carron, Brawley, & Widmeyer, 1998, p. 213). It has been considered one of the most important variables in small groups (Carron & Brawley, 2012). Research on cohesion has shown that a team’s level of cohesion influences important team processes and outcomes (Mullen & Copper, 1994). For instance Reagans and McEvily (2003) indicated individuals who have stronger emotional ties are more likely to share knowledge, and are more trusting of those who they have strong emotional connections with. Individuals also remember more information when they receive it from someone they are emotionally closer to than strangers (Brenner, 1973). Furthermore, Mullen and Copper (1994), in a meta-analysis of cohesion, confirmed a significant relationship between team cohesion and team performance.

4.1.7 Inclusion of Other in Self

Cohesion in teams will be measured using the Inclusion of Other in Self (IOS; Aron, Aron, & Smollan, 1992) scale. The IOS scale is hypothesised to tap an individual’s sense of being interconnected with others (Aron et al., 1992). The scale, asks respondents to select from a set of Venn-like diagrams that indicate different degrees of closeness that best describes their relationship. The diagrams are designed so that the total area of each is constant (thus as the overlap of the circles increases, so does the diameter), and the degree of overlap progresses linearly, creating a seven-step, interval-level scale (Aron et al., 1992). When averaged at the team level, the IOS indicates individual’s perceptions about what the team believes about its level of closeness.

4.1.8 Research Expectations

It is expected that the three workload factors will be differentially impacted by team experience. With increased experience working with a team, and more time to develop cohesion, the demand for task-team balancing workload should decrease. While individual workload and team workload may also decrease, this may not be as notable with the minimal amount of experience provided by only two team sessions. Performance should improve with increased experience (situational awareness increase, mission time decrease, and spatial disorientation decrease), although this is not likely to be dramatic given only two team sessions, with research showing teams are often still learning to work together after two sessions (Cooke, Pederson, Connor, Gorman, & Andrews, 2006). Also teams with lower workload, in particular those who work well together (thus, low in task-team balancing demands) should perform overall better than non-cohesive teams.

4.2 Method

4.2.1 Participants

28 individuals (8 Male and 20 Female) participated in this study, some for course credit, and others for \$15 worth of vouchers. Their average age was 24.1 ($SD = 5.07$) years, and ranged from 21 to 48 years. Fourteen dyadic teams were formed. Participants switched roles between sessions. No participants had any previous experience with quadcopter UAVs.

4.2.2. Design

A mixed experimental design was employed. Course configurations were randomly generated. All teams completed both configurations, however, their order was

counterbalanced across teams, so that half of the teams performed configuration A for their first session while the other half performed configuration B for their first session.

4.2.3. Materials

The experiment was conducted in a quiet laboratory room measuring 12 m by 10 m, with plenty of natural light, and overhead florescent lights, ensuring clear visibility. An 8.0 m by 8.0 m area was marked in the middle of the room using white masking tape (see figure 1). A distance of at least 0.7 m was kept between the walls of the room and the marked area to reduce the chance of crashes. A helipad was marked, by a large 'H', at the front centre of the course, with the Drone taking off and landing in the designated area. A designated area for the guider was marked with chalk, at the front of the room; this area was 0.8 m from the front edge of the marked area. The guiders were positioned here so they could see the entire course, but were far enough back that they could not see the contents of the boxes. The pilot was seated directly behind the guider, 2.4 m from the front of the course. The guider faced forwards out towards the course but their view was obscured by a barrier that completely blocked the view of the course. This was to ensure the pilot had to rely on directions from the guider and feedback from the drone camera. Another barrier was set up to the right of the pilot, for the guider to use when completing questionnaires. This was important as a number of the questionnaires required the team members describe the course; therefore, it was important neither team member could see they lay out of the course when completing the questionnaires.

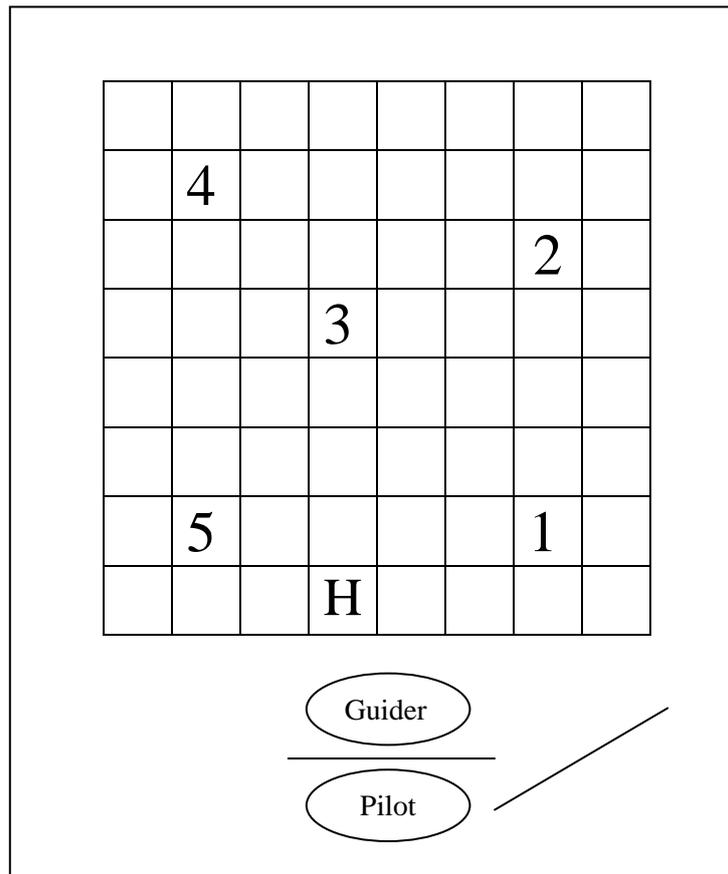


Figure 1. Representation of the experiment field and positioning of guider and pilot. Numbers represent the position of boxes; 'H' represents the helipad; and the two lines indicate the positioning of barriers.

Parrot AR.Drone.

Pilots flew the Parrot AR.Drone 2.0 (Figure 2) fitted with Hull and guard rings, for indoor flight. The AR.Drone is a quadcopter remotely controlled UAV, weighing 420g. The drone has two HD 720p cameras, one forward facing, capable of producing images at 30 frames per second, and another downward facing, capable of producing images at 60 frames per second. Only the downward facing camera was operating during the experiment. The AR.Drone was controlled on an Apple Ipad Touch third generation with 3.5-inch widescreen, 480-by-320-pixel resolution multi-touch display. The Ipad Touch was installed with AR FreeFlight 2.0 software. The drone is controlled with virtual joysticks, one joystick moves the drone forwards, backwards, left, and right; the other joystick controls the elevation and spins

the drone clockwise or counter clockwise. There is also a take-off and land button. Camera images are displayed in real time on the Ipod screen with the controls superimposed on top (See figure 3).



Figure 2. Parrot AR. Drone 2.0 Quadrocopter, with hull and guard rings attached

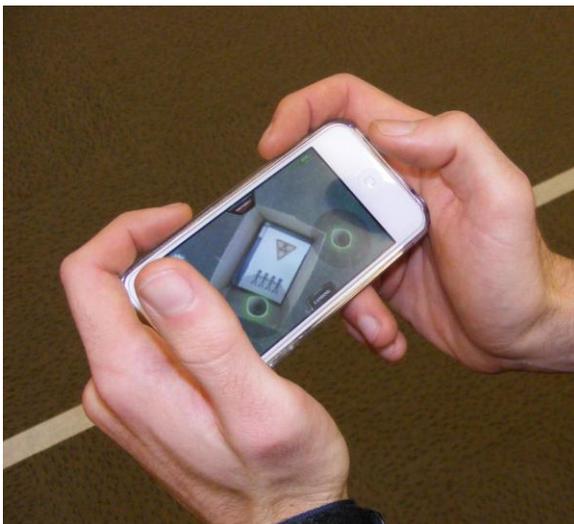


Figure 3. Ipad Touch display showing a live image from the drone downward facing camera with control overlay visible.

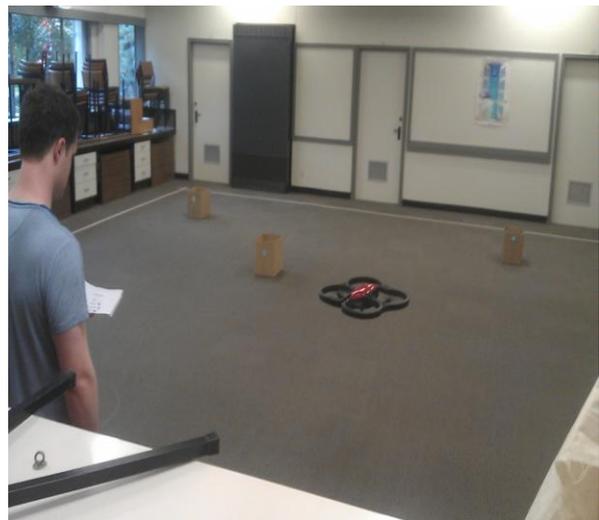


Figure 4. Parrot AR.Drone 2.0 being guided around the course.

Box Contents and Configuration.

Two box configurations (Appendix E) were designed by randomly generating a box position, and box contents. Each configuration consisted of five boxes, which represented sea vessels (see figure 4). The boxes measure 0.35 m X 0.27 m X 0.41 m. Each box contained an A4-sized print-out placed at the bottom of the box. The print-outs featured one to four stick figure people (see appendix F), representing the number of people aboard the boat. One figure equated 10 people. The print-outs also featured a hazard symbol, if there was an issue aboard the boat, or no symbol if there was no issue aboard the boat. The hazard symbols were easily recognisable characterisations of potential disasters and included a flame for fire, a fallen oil barrel with slick for an oil spill, and a bio hazard sign for a biological disaster (Appendix G). The print outs were not visible to the guider; only the pilot could see this by hovering over the box and attaining a visual with the drone's down-facing camera (see figure 3). On the outside of each box was a coloured tag. This was invisible to the pilot (due to it being stuck on a vertical surface) but readily visible to the guider. The coloured tags represented a type of vessel (i.e., recreation, fishing, military, and cargo); the guider was given a reference sheet to identify the type of vessel (Appendix H).

4.2.4. Measures

Team Workload.

The TWLQ developed in study 1 was used to assess the team workload demand during the task. Age, sex, and role (i.e., pilot or guider) demographics were also collected.

Team Cohesion.

To examine the relationship that exists between team members an adapted version of The Inclusion of Other in Self (IOS) scale (Aron et al., 1992; Appendix I), was used.

Participants indicated the degree of their relationship with their team mate by circling one of the seven possible diagrams. The diagrams ranged from not close, indicated by two circles that do not intercept, to extremely close indicated by two circles which almost envelop each other. The level of cohesion for the team is calculated by assigning a number to each degree of relationship (e.g., 1 for not close, 7 for extremely close) summing the scores for each team member then averaging it across the team.

Situational Awareness.

The quantitative analysis of situational awareness (QUASA; Appendix J) was used to assess the situational awareness of teams. The questionnaire consisted of twelve true-false questions which probed the team member's ability to recall key information they encountered during the task. For example, "A fishing vessel had a fire aboard." Two QUASA scales were created for the separate configurations. The scale was scored by calculating the number of correct answers with higher scores indicating a high level of situational awareness.

Spatial Disorientation.

Spatial Disorientation was assessed using a cognitive map method. A map was created which divided the course into 64 square segments, equating a ratio of 2:100 centimetres (See Appendix K). Participants were required to indicate the positioning of the boxes they discovered during the experiment by writing a number which indicated the order of the box (i.e., a '4' for the fourth box they examined). The only reference points on the map were a label indicating the front of the room and an 'H' indicating the position of the helipad. The cognitive map was scored by calculating the number of squares the box was from the actual position. The score on the cognitive map indicated the distance from the actual target, therefore, a higher score represents worse performance and greater spatial disorientation.

4.2.5. Procedure

Upon entering the room participants were given an instruction sheet to read (Appendix L), detailing the expectations of the experiment. They were also given an example of the box contents and shown the possible symbols that could be included in the boxes. Following this the experimenter verbally explained the instruction sheet to ensure participants understood the experiment. Once participants were clear on the details of the experiment they were assigned the role of pilot or guider, based on their last name. The participant with the last name closest to the beginning of the alphabet was assigned as the pilot for the first session. The participant who was assigned the role of pilot was then given a demonstration of how to fly the drone, with an explanation of the controls and capabilities of the drone. They were then given approximately five minutes of practice flying the drone around the course. During the practice period the course was bare except for one box (with no contents) situated in the centre of the room. After initially being given free will to fly the drone, the experimenter instructed the pilot to only look at the screen while flying, using just the visual feed from the downwards camera. The experimenter gave them navigational directions to simulate the experiment; examples of these included “Go forward. Move left slightly. Spin to your right. Move upwards.” It was important to ensure the pilots were able to successfully comply with the directions before beginning the experiment as the task was reliant on the ability of the pilot to respond to instructions from the guider. When the pilot was comfortable, and the experimenter was satisfied with the pilot’s control, both the pilot and guider were asked to move behind separate barriers. At this time they completed an IOS questionnaire. While participants were completing the questionnaire the experimenter set up the course to one of the two configurations. Once the course was set up the guider was asked to come out from behind their barrier and stand in the assigned spot at the front of the room. When teams were ready, the pilot would take off from the helipad and hover the drone until the guider issued

instructions. At this point timing began. The teams then set about checking all the boxes to determine the type of vessel they were encountering, the number of people aboard and the type of hazard (if there was a hazard). The experimenter listened for communication between team members to check that all vessel contents and type were relayed between the pilot and guider, as well as that they were correct. When the drone was landed at the completion of the flying, the time was stopped and recorded. The guider was immediately told to go back behind the guider barrier, while the pilot was told to remain seated behind the pilot barrier. No feedback was given to the team. Several questionnaires were then issued to both team members. These included the TWLQ, QUASA, cognitive map, and another IOS. At this point the course was cleared and a single practice box was put in place in preparation for the second session. Following the participants' completion of the questionnaires, they were asked to emerge from behind their barriers to move on to session number 2. Their roles within the team were swapped and the new pilot commenced flying practice. From here the procedure repeated, as described above, until the experiment ended at the end of session 2.

4.3 Results

Based on the factor analysis of study 1, workload across the three scales was calculated for each individual. The workload score was calculated by averaging the items for each workload factor (i.e., Team Workload was calculated by averaging the scores from Communication Demand, Coordination Demand, and Team Monitoring Demand). For each of the two session's, scores for the two team members (pilot and guider) were averaged to form team level composite scores of Task Workload, Team Workload, Task-Team Balancing, Cohesion, Situational Awareness, and Spatial Disorientation. Time to complete the simulated mission was already a team-level metric.

Paired samples t-tests (two-tailed) were performed for each of the team level items comparing the first and second sessions (See Table 2). As is displayed in Table 3, Task-Team Balancing significantly decreased for the second session, and cohesion significantly increased for the second session. While none of the performance metrics, when compared for session 1 and 2 reached statistical significant at an alpha < .05 (spatial disorientation was close, $p = .054$), the overall trend was for improved time performance (e.g. they finished the second mission more quickly) and a slight trend in improved situational awareness, but this came at the cost of increased spatial disorientation for the second session.

Table 2. Means, Standard Deviation and t-value of performance variables and workload scale measurements for Session 1 and Session 2

	Session 1		Session 2		T-Value
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	
Time	5.02	2.21	4.02	1.42	1.65
Situational Awareness	6.43	1.71	7.07	1.62	-0.93
Spatial Disorientation	12.9	4.36	15.4	3.16	-2.10
Task Workload	5.18	0.92	5.55	1.12	-1.54
Team Workload	7.20	0.96	7.39	1.37	-0.63
Task-Team Balancing	4.37	1.41	3.68	1.22	2.26*
IOS	3.30	1.79	3.86	1.63	-3.42*

* $p < .05$

In order to further explore the relationship between the workload factors and performance, composite performance metrics were calculated. Overall task performance was a combination of time to complete the mission, situational awareness, and spatial disorientation. Situation awareness was reversed scored to align all the performance metrics, thus a higher performance score indicates overall worse performance (greater spatial disorientation, longer mission time, and reduced situational awareness). Each of the performance items were then transformed into z-scores (as they are on different scales) and summed to form a composite performance score. For both sessions, correlation analysis was conducted to examine the relationship between the workload factors and the composite performance metric. Given the low sample size (only 14 teams) a more lenient alpha = .10 was used as criteria. While the more lenient alpha will increase the chances of Type I error;

the exploratory nature of the study means that a more lenient alpha can identify potentially important relationships that future studies, employing larger sample sizes, can explore further. At the $p < .10$ significant correlations were identified between Task-Team Balancing and the performance composite score. The relationship between Task Workload and Team Workload was weaker. See Table 3 for a display of these correlations. Those teams reporting lower demands of balancing taskwork and teamwork had better overall performance, $p < .10$.

Table 3. Correlations between Workload factors, Composite Performance score and IOS scale (Session 1 correlations are above the main diagonal; Session 2 correlations are below the main diagonal)

	1	2	3	4	5
1. Task Workload		.244	.315	.337	-.191
2. Team Workload	.673**		.495*	.293	-.308
3. Task-Team Balancing	.281	.484*		.483*	-.379
4. Performance	.278	.216	.465*		-.464*
5. IOS	-.255	-.039	-.323	-.268	

* $p < 0.1$; ** $p < 0.01$

4.4 Discussion

The results of this experiment indicate the three scales of the TWLQ exhibit differential sensitivity, with the Task-Team Balancing measure acting independently from the Task Workload and Team Workload scales. Task-Team Balancing significantly decreased from session 1 to session 2 indicating teams became more adapt at managing the interaction between taskwork and teamwork with experience in the team. This result is coupled with an increase in cohesiveness, with teams becoming more cohesive in session 2. Presumably as team members become more cohesive and feel closer to their teammate they are more trusting of the abilities of their teammate (Reagans & McEvily, 2003), and thus have to balance less between taskwork and teamwork. This is further supported by the improved performance in session 2, indicating the teams are able to focus more on the taskwork than teamwork. While the performance results were not significant teams showed performance

improvements in average time, and increased situational awareness, however, there is a trade-off with greater spatial disorientation. It is not entirely surprising non-significant results were found for the performance metrics as much of the initial focus in new teams is about learning how to interact, with the performance potential of new teams not revealed until they have developed a shared understanding (Cooke, et al., 2006). With more time together the teams would develop a greater shared understanding which I believe would translate into significant performance improvements. Furthermore, the task was very difficult with the participants expected to perform a number of tasks simultaneously. For experienced operators of UAVs it would not have been as demanding, but all participants were novices. Therefore, much of their attention would have been focused on operation of the drone rather than directing full attention to the task.

When looked at between-teams those teams that were better able to balance taskwork and teamwork performed significantly better across both sessions (although given the lenient alpha level, this is a tentative conclusion). This indicates the importance of task-team balancing in teams and the demand that it places on individuals. There was also a significant relationship between Task-Team Balancing and Team Workload with high team workload demands associated with high demand for balancing taskwork and teamwork. This would be expected, as a high demand of Team Workload should lead to greater requirements for Task-Team Balancing. Conversely however, Task Workload was not shown to be related to Task-Team Balancing. This is surprising as it would be expected high levels of Task Workload would result in an increased demand for Task-Team Balancing. A possible explanation for this is that the task is the primary focus of teams; therefore, when there is a high level of Team Workload they notice that they are spending more time interacting with teammates and thus report an associated level of Task-Team Balancing.

5.0 General Discussion

The main purpose of this study was the development of a subjective measure of team workload, which would aid researchers and practitioners in the measurement of workload in teams. The TWLQ was developed in study 1 drawing on current literature in the creation of the questionnaire. The TWLQ was then experimentally tested in study 2 to examine its sensitivity and ability to be used in an applied setting. Combining the results from both studies allows an analysis of the validity and reliability of the TWLQ.

Construct Validity.

Construct validity is the principle concern in evaluating the validity of the TWLQ, as the credibility to represent a latent construct is recognised by its construct validity (Haig, 1999). To assess the construct validity of the TWLQ the convergent and discriminant validity of the scale was examined.

Convergent Validity. The relationship between the workload factors and performance provides evidence of convergent validity. High levels of workload have consistently been shown to negatively impact performance (Bowers et al., 1997; Urban et al., 1995). In both sessions high levels of workload are associated with worse performance. While the relationship was only significant for task-team balancing the general trend is encouraging.

Discriminant validity. The discriminant validity of the TWLQ is strong. Factor analysis identified three clear constructs, with minimal cross loading on the other factors. This indicates each construct is measuring a distinct component of team workload. Furthermore, there was no crossover of teamwork and taskwork items; all taskwork items loaded on one factor and teamwork items loaded on two unique factors. This is encouraging as it indicates the TWLQ is able to differentiate between taskwork and teamwork. The scale

also showed sensitivity to differing levels of workload with the three scales acting independently in assessing levels of workload in relation to performance.

Content Validity.

Content validity of the TWLQ was also examined. Content validity is concerned with the representativeness of the scale, in that it encompasses all facets of the construct. A content valid scale is directly related to the procedures used to create the scale (Nunnally, 1976). In developing the TWLQ considerable care was taken to ensure representativeness of the team workload construct by examining past research and literature, and following best practice in the development of new items. There is a considerable amount of uncertainty regarding team workload (Funke et al., 2012) which coupled with the uncertainty regarding essential team behaviours (Porter et al, 2010) makes it difficult to discern the representativeness of the TWLQ. It is however encouraging the scale formed along the conceptual lines of team workload with separate taskwork, teamwork and task-team balancing scales. Overall the stringent procedures used in the development of the TWLQ leads to the conclusion that it is a content valid measure, up to a standard currently possible considering the lack of certainty in the team workload construct.

Reliability.

The reliability of the scales is acceptable especially considering the Team Workload and Task-Team Balancing scales only consisted of three items, which is considered a minimum threshold required for producing acceptable levels of internal consistency (Hinkin, 1995). Two of the scales were above the recommended Cronbach Alpha of .70, with Task-Team Balancing very close ($\alpha = .692$). Hinkin (1995) reports the common cause of low reliabilities is as a result of poor item generation and lenient scale development methods. The

reported reliabilities positively reflect the steps taken throughout, and endorse the stringent statistical minimums required during the scale development phase.

5.1 Practical and Theoretical Applications

In the absence of a dedicated and viable team workload measure, this study provides researchers and practitioners with an effective tool for the measurement of workload in teams. The TWLQ allows researchers to utilise a measure of team workload that is easy to use and interpret, and does not require the use of expensive and complex machines, like those required for performance and physio-behavioral measures (Funke et al., 2012). For practitioners the TWLQ can be used in the design and assessment of work settings, to understand the levels of workload work teams are experiencing. It is especially applicable in situations where the use of other workload measures is not possible. The easy and simplicity of using the TWLQ means it can be used in a range sectors providing practitioners with easily interpretable results, that identifies the specific aspects causing high workload.

Theoretically this study adds to the concept of workload. First and foremost the study adds to the growing evidence that measures of performance and frustration should not be included in the assessment of task workload (Bailey & Thompson, 2001; Ramiro et al., 2010). Further advancing this concept, the study indicates performance and frustration are not appropriate as workload items in the team workload context either. While frustration and performance appraisals are to some extent related to workload (perhaps, responses to or preconditions for), they are theoretically not workload per se (assessments of how hard one or a team is working). The study has also identified two new demands related to workload, specifically, emotion demand and performance monitoring demand. Emotion and performance monitoring are associated with both taskwork and teamwork and offer a new

perspective on the workload demands. Indicating the emotional and performance monitoring aspects of tasks are indeed workload (i.e., that require effort).

5.2 Limitations and Future Research

This study is not without its limitations. In study 1 the team workload questionnaire was administered solely to sports teams. This could be considered a limitation as the ability to generalise from sports teams to work teams may be difficult. Sports teams operate in a unique environment with many different characteristics than those found in work settings. Although the results were consistent with similar studies conducted with non-sports teams, a replication of study 1 with work teams would be interesting.

The sample of study 1 is also a limitation. Not so much the sample size, which was acceptable for performing exploratory factor analysis (Hinkin, 1995), but rather the disproportionate number of males to females. Males made up over 80% of the sample, and it is reasonable to question whether different results may have been found if there was a more even split of male and females.

The sample size of study 2 was also a concern with only 14 teams participating in the study. Due to time and resource constraints it was not possible to conduct this experiment with a larger number of participants. Given a larger sample size the relationships found in the experiment could be examined with more power. An associated issue related to the sample size of study 2 is the use of a lenient alpha level. The small sample size coupled with the exploratory nature of the research lead to the decision to employ a more lenient alpha level. While employing a more lenient alpha allowed the identification of interesting relationships, Type I error is a concern and must be considered when interpreting the results from study 2. It would be interesting to see if the results identified would be similar with a larger sample size and a stricter alpha level.

Future research should also look to rigorously test the TWLQ exploring the measurement characteristics of the questionnaire with a more in-depth analysis of the validity of the scales. One possible avenue for future evaluation of the TWLQ is a comparison to other measures of workload (e.g., performance measures and physio-behavioral measures). This could provide worthwhile information regarding the convergent validity of the scale, and the sensitivity to differing levels of workload.

5.3 Concluding Statement

Given the relative lack of research into team workload it is important researchers continue to explore the concept further. It is amazing so little workload research has been conducted in collaborative settings when considering the importance of teams. With the team workload questionnaire developed in this study researchers have a viable tool that they can use to further explore workload in teams, and practitioners have an easy to use subjective measure of team workload that can be used in a variety of settings. While further research is needed to explore the measurement properties of the team workload questionnaire it opens many avenues for the advancement of team workload research.

References

- Aron, A., Aron, E. N., & Smollan, D. (1992). Inclusion of Other in the Self Scale and the Structure of Interpersonal Choices. *Journal of Personality and Social Psychology*, 63(4), 596-612.
- Austin, J. T., & Vancouver, J. B. (1996). Goal Constructs in Psychology: Structure, Processes, and Content. *Psychological Bulletin*, 120(3), 338-375. DOI: 10.1037/0033-2909.120.3.338
- Bailey, L., & Thompson, R. (2001). The TLX: One or more constructs. Paper presented at the Eleventh International Symposium for Aviation Psychology
- Baker, D. P., & Salas, E. (1997). Principles for Measuring Teamwork: A Summary and Look Toward the Future. In M. T. Brannick, E. Salas, & C. Prince (Eds.), *Team Performance Assessment and Measurement: Theory, Methods, and Applications* (pp. 85-108). Mahwah, NJ: Lawrence Erlbaum.
- Barnes, C. M., Hollenbeck, J. R., Wagner, D. T., DeRue, D. S., Nahrgang, J. D., & Schwind, K. M. (2008). Harmful help: The costs of backing-up behavior in teams. *Journal of Applied Psychology*, 93, 529-539. DOI: 10.1037/0021-9010.93.3.529
- Benini, A., Mancini, A., Minutolo, R., Longhi, D., & Montanari, M. (2012). A Modular Framework for Fast Prototyping of Cooperative Unmanned Aerial Vehicle. *Journal of Intelligent and Robotic Systems*, 65, 507-520. DOI: 10.1007/s10846-011-9577-1
- Benson, A. J. (1999). Spatial disorientation—General aspects. In J. Ernsting, A. N. Nicholson, & D. J. Rainford (Eds.), *Aviation medicine* (3rd ed., pp. 419–436). Oxford, UK: Butterworth Heinemann
- Bowers, C. A., Bruan, C. C., & Morgan Jr, B. B. (1992). The Study of Crew Coordination and performance in hierarchical team decision making (Tech. Rep. No. 92-1). Orlando: university of Central Florida, Team Performance Laboratory.

- Bowers, C. A., Bruan, C. C., & Morgan Jr, B. B. (1997). Team Workload: Its Meaning and Measurement. In M. T. Brannick, E. Salas, & C. Prince (Eds.), *Team Performance Assessment and Measurement: Theory, Methods, and Applications* (pp. 85-108). Mahwah, NJ: Lawrence Erlbaum.
- Bowers, C. A., & Jentsch, F. (2005). Team Workload. In N. Stanton, A. Hedge, K. Brookhuis, E. Salas, & H. Hendrick (Eds.), *Handbook of Human Factors and Ergonomics Methods* (pp.57-1-57-3). Boca Raton, FL: CRC Press. DOI: 10.1201/9780203489925.ch57
- Brenner, M. (1973). The next-in-line effect. *Journal of Verbal Learning and Verbal Behavior*, 12, 320-323. DOI: 10.1016/S0022-5371(73)80076-3
- Burke, J. L., & Murphy, R. R. (2004). Human-Robot Interaction in USAR Technical Search: Two Heads Are Better Than One. Proceedings of the 2004 IEEE International Workshop on Robot and Human Interactive Communication Kurashiki, Okayama Japan September 20-22
- Burke, J. L., Murphy, R.R., Coovert, M. D., & Riddle, D. L. (2004). Moonlight in Miami: A Field Study of Human-Robot Interaction in the Context of an Urban Search and Rescue Disaster Response Training Exercise. *Human-Computer Interaction*, 19(1-2), 85 - 116.
- Burke, S. C., Wilson, K. A., & Salas, E. (2003). Teamwork at 35,000 Feet: Enhancing safety through Team Training. *Human Factors and Aerospace Safety*, 3(4), 287-312.
- Carron, A. V., & Brawley, L. R. (2012). Cohesion: Conceptual and Measurement Issues. *Small Group Research*, 43(6), 726-743. DOI: 10.1177/1046496412468072
- Carron, A. V., Brawley, L. R., & Widmeyer, W. N. (1998). The Measurement of Cohesiveness in Sport Groups. In J. L. Duda (Ed.), *Advances in Sport and Exercise*

- Psychology measurement (pp. 213-226). Morgantown, WV: Fitness Information Technology.
- Chiocchio, F. (2007). Project Team Performance: A Study of Electronic Task and Coordination Communication. *Project Management Journal*, 38(1), 97-109.
- Cooke, N. J., Pederson, H. K., Connor, O., Gorman, J. C., & Andrews, D. (2006). Acquiring Team-Level Command and Control Skill for UAV Operation. In N. J. Cooke, H. L. Pringle, H. K. Pedersen, & O. Connor (Eds.), *Human Factors of Remotely Operated Vehicles (Advances in Human Performance and Cognitive Engineering Research, Volume 7)* (pp. 285 – 297). Burlington, MA : Elsevier. DOI: 10.1016/S1479-3601(05)07020-7
- Compo, M., Mellalieu, S., Ferrand, C., Martinet, G., & Rosnet, E. (2012) Emotions in Team Contact Sports: A Systematic Review. *The Sport Psychologist*, 26, 62-97.
- DeCaro, M. S., Thomas, R. D., Albert, N. B., & Beilock, S.L. (2011). Choking Under Pressure: multiple Routes to Skill Failure. *Journal of Experimental Psychology*, 140(3), 390-406. DOI: 10.1037/a0023466
- Dickinson, T. L., & McIntyre, R. M. (1997). A conceptual framework for teamwork measurement. In M. T. Brannick, E. Salas, & C. Prince (Eds.), *Team Performance Assessment and Measurement: Theory, Methods, and Applications* (pp. 19-43). Mahwah, NJ: Lawrence Erlbaum.
- Drury, J. L., Richer, J., Rackliffe, N & Goodrich, M. A. (2006). Comparing Situation Awareness for Two Unmanned Aerial Vehicle Human Interface Approaches. Proceedings of the IEEE International Workshop on Safety, Security and Rescue Robotics (SSRR).
- Ebel, R.L. & Frisbie, D.A. (1991). *Essentials of Educational Measurement* (5th ed.). Englewood Cliffs, NJ: Prentice-Hall.

- Endsley, M. R. (1988). Design and Evaluation for Situation Awareness Enhancement. *Proceedings of the Human Factors Society 32nd Annual Meeting*, 32, 97-101
- Endsley, M. R. (2001). Designing for Situation Awareness in Complex Systems. Proceedings of the Second International Workshop on Symbiosis of Humans, Artifacts and Environment, Kyoto, Japan.
- Everaerts, J. (2008). The Use of Unmanned Aerial Vehicles (UAVs) for Remote Sensing and Mapping. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XXXVII(B1), 1187-1192.
- Ford, J. K., & Schmidt, A. M. (2000). Emergency response training: Strategies for enhancing real-world performance. *Journal of Hazardous materials*, 75(2-3), 195-215. DOI: 10.1016/S0304-3894(00)00180-1
- Frye, S. R. (1988). An Approach to Enhanced Control Room Crew Performance. *IEEE 4th Conference on Human Factors and Power Plants*. Monterey, CA: IEEE.
- Funke, G. J., Knott, B. A., Salas, E., Pavlas., & Strang, A. J. (2012). Conceptualization and Measurement of Team Workload: A Critical Need. *Human Factors*, 54(1), 36-51. DOI: 10.1177/0018720811427901
- Gertler, J. (2012). *U.S. Unmanned Aerial Systems* (Congressional Research Service Report No. R42136). Retrieved from www.fac.org/sgp/crs/natsec/R42136.pdf
- Goldberg, B. (2010) Unmanned Aerial Systems: The Role of the Operator and Human Factor. Proceedings of the 2010 Spring Simulation Multiconference, Orlando, FL. DOI: 10.1145/1878537.1878776
- Gorman, J. C., Cooke, N. J., Pederson, H. K., & Connor, O. O. (2005). Coordinated Awareness of Situation by Teams (Cast): Measuring Team Situation Awareness of a Communication Glitch. Proceedings of the Human Factors And Ergonomics Society 49th Annual Meeting, 274-276. DOI: 10.1177/154193120504900313

- Gopher, D., & Braune, R. (1984). On the Psychophysics of Workload: Why bother with Subjective Measures? *Journal of Human Factors and Ergonomics*, 26(5), 519-532.
DOI: 10.1177/001872088402600504
- Guznov, S., Matthews, G., Funke, G., & Duke, A. (2011). Use of the RoboFlag Synthetic Task Environment to Investigate Workload and Stress Responses in UAV Operation. *Behaviour Research Methods*, 43(3), 771–780. DOI 10.3758/s13428-011-0085-9
- Haig, B. D. (1999). Construct Validation and Clinical Assessment. *Behaviour Change*, 16(1), 64-73.
- Hancock, P. A., Mouloua, M., Gilson, R. D., Szalma, J., & Oron-Gilad, T. (2007). Is the UAV control ratio the right question? *Ergonomics in Design*, 15(1), 30–31. DOI: 10.1177/106480460701500104
- Hart, S. G., & Staveland, L. E. (1988). Development of a Multi-Dimensional Workload Scale: Results of Empirical and Theoretical Research. In P.A. Hancock & N. Meshkati (Eds.), *Human Mental Workload* (pp. 139-183). Amsterdam, Netherlands: North-Holland.
- Helton, W. S., Matthews, G., & Warm, J. S. (2009). Stress State Mediation between Environmental Variables and Performance: The Case of Noise and Vigilance. *Acta Psychologica*, 130(3), 204-213. DOI:10.1016/j.actpsy.2008.12.006
- Helmreich, R. L., & Foushee, H. C. (1993). Why Crew Resource Management? Empirical and Theoretical bases of Human Factors Training in Aviation. In E. L. Weiner, B. J. Kanki , & R. L. Helmreich (Eds.), *Cockpit Resource Management* (3-45). San Diego, CA: Academic Press.
- Hendy, K. C. (1995). Situation Awareness and Workload: Birds of a Feather? AGARD AMP Symposium on ‘Situational Awareness: Limitations and Enhancements in the Aviation Environment, 21-1 – 21-7, Brussels: Belgium

- Hinkin, T. R. (1995). A Review of Scale Development Practices in the Study of Organizations. *Journal of Management*, 21(5), 967-988.
- Jentsch, F., & Bowers, C. A. (2005). Team Communication Analysis. In N. Stanton, A. Hedge, K. Brookhuis, E. Salas, & H. Hendrick (Eds.), *Handbook of Human Factors and Ergonomics Methods* (p.50-1-50-4). Boca Raton, FL: CRC Press. DOI: 10.1201/9780203489925.ch50
- Lazarus, R. S. (1982). Thoughts on the Relations between Emotion and Cognition. *American Psychologist*, 37(9), 1019–1024. DOI: 10.1037/0003-066X.37.9.1019
- Levi, D. (2007). *Group Dynamics for Teams* (2nd ed.). Thousand Oaks, CA: Sage Publications.
- Lin, C. J., Hsieh T., Tsai, T., Yang, C., & Yenn, T. (2011). Development of a Team Workload Assessment Technique for the Main Control Room of Advanced Nuclear Power Plants. *Human Factors and Ergonomics in Manufacturing & Service Industries* 21(4), 397-411. DOI: 10.1002/hfm.20247
- Mann, D.T.Y., Williams, A.M., Ward, P., & Janelle, C.M. (2007). Perceptual-cognitive expertise in sport: A meta-analysis. *Journal of Sport & Exercise Psychology*, 29, 457–478.
- Marks, M. A., Mathieu, J. E., & Zaccaro, S. J. (2001). A Temporally Based Framework and Taxonomy of Team Processes. *Academy of Management Review*, 26(3), 356-376. DOI:10.5465/AMR.2001.4845785
- Mathieu, J. E., Heffner, T. S., Goodwin, G. F., Salas, E., & Cannon-Bowers, J. A., (2000). The Influence of Shared Mental Models on Team Processes and Performance. *Journal of Applied Psychology*, 85(2), 273-283. DOI: 10.1037//0021-9010.85.2.273
- McGuinness, B. (2004). Quantitative Analysis of Situational Awareness (QUASA): Applying Signal Detection Theory to True/False Probes and Self-Ratings. *9th International*

Command and Control Research and Technology Symposium. Copenhagen, Denmark.

- Menda, J., Hing, J. T., Ayaz, H., Shewokis, P. A., Izzetoglu, K., Onaral, B., & Oh, P. (2011). Optical Brain Imaging to Enhance UAV Operator Training, Evaluation, and Interface Development. *Journal of Intelligent and Robotic Systems*, *61*(1-4), 423-443. DOI 10.1007/s10846-010-9507-7
- Mola, R. A. (2008). Do Drones Get Vertigo, Too? Up there or down here, it can be a struggle to maintain “situational awareness.” *Air & Space*. Retrieved from http://www.airspacemag.com/flight-today/Do_Drones_Get_Vertigo_Too.html?c=y&page=1
- Mullen, B., & Cooper, C. (1994). The relation between group cohesiveness and performance: An integration. *Psychological Bulletin*, *115*(2), 210-227.
- Neily, J., Mills, P. D., Young-Xu, Y., Carney, B. T., West, P., Berger, D. H., Mazzia, L. M., Paull, D. E., & Bagian, J. P. (2010). Association between Implementation of a Medical Team Training Program and Surgical Mortality. *Journal of American Medical Association* *304*(15), 1693-1700
- Norman, D. A., & Bobrow, D. G. (1975). On Data-Limited and Resource-Limited Processes. *Cognitive Psychology*, *7*, 44-64.
- Nunnally, J.C. (1976). *Psychometric Theory*, 2nd ed. New York: McGraw-Hill.
- Nygren, T. E. (1991). Psychometric Properties of Subjective Workload Measurement Techniques: Implications for their Use in the Assessment of Perceived Mental Workload. *Human Factors*, *33*(1) 17-33.
- O’Hara, J. M., & Hall, R. E. (1992). Advanced Control Rooms and Crew Performance Issues: Implications for Human Reliability. *IEEE Transactions on Nuclear Science*, *39*(4), 919-923. DOI: 0018-9499/92\$03.00

- Porter, C. O. L. H., Gogus, C. I., & Yu, R. C. (2010). When Does Teamwork Translate Into Improved Team Performance? A Resource Allocation Perspective. *Small Group Research, 41*(2), 221-248. DOI: 10.1177/1046496409356319
- Porter, C. O. L. H., Hollenbeck, J. R., Ilgen, D. R., Ellis, A. P. J., West, B. J., & Moon, H. (2003). Backing up Behaviours in Teams: The role of Personality and Legitimacy. *Journal of Applied Psychology, 88*(3), 391-403. DOI: 10.1037/0021-9010.88.3.391
- Previc, F. H., & Ercoline, W. R. (2004). Spatial Disorientation in Aviation: Historical Background, Concepts, and Terminology. In F. H. Previc & W. R. Ercoline (Eds.), *Spatial Disorientation in Aviation* (Vol. 203; pp. 1–36). Lexington, MA: American Institute of Aeronautics and Astronautics.
- Ramiro, E. D., Valdehita, S. R., Lourdes, J. M. G., & Moreno, L. (2010). Psychometric Study of NASA-TLX Mental Workload Index in a Sample of Spanish Workers. *Revista de Psicología del Trabajo y de las Organizaciones, 26*(3), 191-199. DOI: 10.5093/tr2010v26n3a3
- Reagans, R., & McEvily, B. (2003). Network Structure and Knowledge Transfer: The Effects of Cohesion and Range. *Administrative Science Quarterly, 48*(2), 240-267.
- Rubio, S., Diaz, E., Martin, J., & Puente, J. M. (2004). Evaluation of Subjective Mental Workload: A Comparison of SWAT, NASA-TLX, and Workload Profile Methods. *Applied Psychology: An International Review, 53*(1), 61-86.
- Salas, E., Shuffler, M. L., & DiazGranados, D. (2010). Team Dynamics at 35,000 Feet. In E. Salas & D. Maurino (Eds.), *Human Factors in Aviation* (2nd ed., pp. 249-329). Shuffler Burlington, MA: Elsevier.
- Salmon, P., Stanton, N., Walker, G., & Green, D. (2006). Situation Awareness Measurement: A Review of Applicability for C4i Environments. *Applied Ergonomics, 37*(2), 225-238.

- Serfaty, D., Entin, E. E., & Johnston, J. H. (1998). Team coordination training. In J. A. Cannon-Bowers & E. Salas (Eds.), *Making decisions under stress: Implications for individual and team training* (pp. 221-245). Washington, DC: American Psychological Association.
- Stout, R. J., Cannon-Bowers, J. A., Salas, E., & Milanovich, D. M. (1999). Planning, Shared Mental Models, and Coordinated Performance: An Empirical Link Is Established. *Human Factors*, *41*(1), 61-71. DOI: 10.1518/001872099779577273
- Totterdell, P. (2000). Catching moods and hitting runs: Mood linkage and subjective performance in professional sport teams. *The Journal of Applied Psychology*, *85*, 848–859
- Tsang, P., & Vidulich, M. (2006). Workload and Situational Awareness. In G. Salvendy (Ed.), *Handbook of Human Factors and Ergonomics* (3rd ed., pp. 243-269). New York, NY: John Wiley.
- Tse, H. H. M., & Dasborough, M. T. (2008). A Study of Exchange and Emotions in Team Member Relationships. *Group & Organization Management*, *33*(2): 194- 215. DOI: 10.1177/1059601106293779
- Urban, J. M., Bowers, C. A., Monday, S. D., & Morgan Jr, B. B. (1995). Workload, Team Structure, and Communication in Team Performance. *Military Psychology*, *7*(2), 123-139. DOI: 10.1207/s15327876mp0702_6
- Urban, J.M., Weaver, J.L., Bowers, C.A., & Rhodenizer, L. (1996). Effects of Workload and Structure on Team Processes and Performance: Implications for complex team decision making. *Human Factors*, *38*(2), 300-310. DOI: 10.1177/001872089606380210

- Webb, C. M., Estrada III, A., & Kelley, A. M (2012). The Effects of Spatial Disorientation on Cognitive Processing. *The International Journal of Aviation Psychology*, 22(3), 224-241. DOI: 10.1080/10508414.2012.689211
- Williams, K. W. (2008). Documentation of Sensory Information in the Operation of Unmanned Aircraft Systems. *Federal Aviation Administration, Office of Aerospace Medicine* (1-23). Oklahoma City, OK. Retrieved from <http://www.dtic.mil/cgi-bin/GetTRDoc?AD=ADA490325>
- Wilson, K. A., Guthrie, J. W., Salas, E., & Howse, W.R. (2006). Team Processes. In J.A. Wise, V.D. Hopkin, & D.J. Garland (Eds.), *Handbook of Aviation Human Factors* (2nd ed., pp 9.1-9.22). Hillsdale, NJ: CRC Press.
- Yanco, H. A., & Drury, J. Where Am I? Acquiring Situation Awareness Using a Remote Robot Platform. In Proceedings of the IEEE Conference on Systems, Man and Cybernetics, October 2004
- Yang, C. W., Yenn, T. C., & Lin, C. J. (2010). Assessing Team Workload Under Automation Based on a Subjective Performance Measure. *Safety Science*, 48(7), 914-920. DOI:10.1016/j.ssci.2010.03.011
- Young, M. S. and Stanton, N. A. (2005). Mental Workload. In N. Stanton, A. Hedge, K. Brookhuis, E. Salas, & H. Hendrick (Eds.), *Handbook of Human Factors and Ergonomics Methods* (39-1- 39-9). Boca Raton, FL: CRC Press. DOI: 10.1201/9780203489925.ch39
- Zeitlin, L. R. (1995). Estimates of Driver Mental Workload: A long-term feild Trial of Two Subsidiary tasks. *Human Factors*, 37, 611-621.
- Zohar, D., Tzischinski, O., & Epstein, R. (2003). Effects of Energy Availability on Immediate and Delayed Emotional Reactions to Work Events. *Journal of Applied Psychology*, 88(6), 1082-1093. DOI: 10.1037/0021-9010.88.6.1082

Appendix A – Team Workload Measure

The following questions relate to your experiences during the task. Use the response scale below the question by ticking the circle closest to your answer; the scale goes from 0 (**very low**) to 10 (**very high**).

1. How physically activity was required (e.g., pushing, pulling, turning, controlling, activating, etc.)?

0 1 2 3 4 5 6 7 8 9 10

2. How much mental and perceptual activity was required (e.g. thinking, deciding, calculating, remembering, looking, searching, etc.)?

0 1 2 3 4 5 6 7 8 9 10

3. How much time pressure did you feel due to the rate or pace at which the tasks or task elements occurred?

0 1 2 3 4 5 6 7 8 9 10

4. How much did you have to control your emotions (e.g., anger, joy, disappointment)?

0 1 2 3 4 5 6 7 8 9 10

5. How much did the task require you to monitor your performance (i.e., ensure you were performing at specific levels)?

0 1 2 3 4 5 6 7 8 9 10

6. How successful do you think you were in accomplishing the task?

0 1 2 3 4 5 6 7 8 9 10

7. How hard did you have to work (mentally and physically) to accomplish your level of performance?

0 1 2 3 4 5 6 7 8 9 10

8. How frustrated (e.g., insecure, discouraged, irritated, stressed, and annoyed) versus satisfied (e.g., secure, gratified, content, relaxed and complacent) did you feel during the task?

0 1 2 3 4 5 6 7 8 9 10

The following questions relate to the team's performance in the task, you should answer them by considering the team as a whole. Use the response scale below the question by ticking the circle closest to your answer; the scale goes from 0 (**very low**) to 10 (**very high**).

9. How much communication activity was required by your team (e.g. discussing, negotiating, sending and receiving messages, etc.)?

0 1 2 3 4 5 6 7 8 9 10

10. How much coordination activity was required by your team (e.g. changing or adjusting, plans, etc.)?

0 1 2 3 4 5 6 7 8 9 10

11. How difficult was it to share and manage time between task-work (work done individually) and team-work (work done as a team)?

0 1 2 3 4 5 6 7 8 9 10

12. How emotionally demanding was working in the team?

0 1 2 3 4 5 6 7 8 9 10

13. How much did the task require you to monitor your team's performance?

0 1 2 3 4 5 6 7 8 9 10

14. How successful do you think the team was in working together?

0 1 2 3 4 5 6 7 8 9 10

15. How difficult was it to provide and receive support (providing guidance, helping team members, providing instructions, etc.) from team members?

0 1 2 3 4 5 6 7 8 9 10

16. How irritated and annoyed were you with your team?

0 1 2 3 4 5 6 7 8 9 10

Age? _____

Sex? Male Female

What position do you play? _____

How long have you been in the team? _____

The End

Thank you for your participation. If you would like to enter the draw to win an iPad2, please leave a name and contact details in the space provided.

Please be aware that the information provided will only be used to contact you in the event that you are drawn as the winner of the prize. No personal information provided will be used in the study or made available. The draw will occur on 19/11/12.

Name: _____

Contact information (email or phone number): _____

Appendix B – Stress Scale

The following questions relate to how you felt during the task. Use the response scale below the question by ticking the circle closest to your answer; the scale goes from 0 (**very low**) to 10 (**very high**).

1. How mentally tired did you feel during the task?

0 1 2 3 4 5 6 7 8 9 10

2. How physically tired did you feel during the task?

0 1 2 3 4 5 6 7 8 9 10

3. How tense or anxious did you feel during the task?

0 1 2 3 4 5 6 7 8 9 10

4. How unhappy did you feel during the task?

0 1 2 3 4 5 6 7 8 9 10

5. How motivated were you to do well?

0 1 2 3 4 5 6 7 8 9 10

6. How interesting was the task?

0 1 2 3 4 5 6 7 8 9 10

7. How much did you think about yourself while doing the task?

0 1 2 3 4 5 6 7 8 9 10

8. How focused were you on the task?

0 1 2 3 4 5 6 7 8 9 10

9. How confident were you during the task?

0 1 2 3 4 5 6 7 8 9 10

10. How often did you think about the task?

0 1 2 3 4 5 6 7 8 9 10

11. How often did you think about things unrelated to the task?

- 0** **1** **2** **3** **4** **5** **6** **7** **8** **9** **10**

Appendix C - Information Sheet and Consent Form

Department of Psychology



Information Sheet

You are invited to participate as a subject in the Masters Research project titled, 'Team workload: Subjective Measure Development'.

The aim of this project is to develop a measure of team workload.

Your involvement in this project will be to complete a team workload questionnaire, which should take no longer than five minutes. You have the right to withdraw from the project at any time, including the ability to withdraw any information you may have already provided.

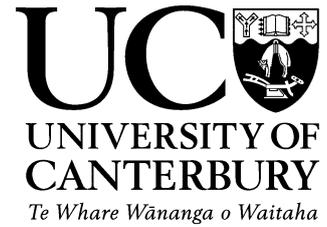
To thank you for taking the time to complete the questionnaire you can choose to enter a draw to win an Apple iPad 2 16GB

Please be aware, that a Masters is a public document that will be accessible through the University of Canterbury, library website. The results of the project may also be published, but you are assured of complete confidentiality of data gathered in this investigation. The identity of participants will not be made public without consent.

The project is being carried out as requirement for Masters of Science by James Sellers under the supervision of, Associate Professor Deak Helton and Dr. Katharina Näswall, who can be contacted by phone on, 03 3 364 2998 or 03 3 364 2552, respectively. They will be pleased to discuss any concerns you may have about the participation in the project. Alternatively, if you have any questions or concerns, you can contact James through his email jms282@pg.canterbury.ac.nz

The project has been reviewed **and approved** by the University of Canterbury Human Ethics Committee.

Department of Psychology



Consent Form

Team Workload: Subjective Measure Development

I have read and understood the description of the above-named project. On this basis I agree to participate as a subject in the project, and I consent to publication of the results of the project with the understanding that anonymity will be preserved.

I understand also that I may, at any time, withdraw from the project, including withdrawal of any information I have provided.

I note that the project has been reviewed **and approved** by the University of Canterbury Human Ethics Committee.

Name (Please Print):

Signature:

Date:

Appendix D – Task Workload Questionnaire (TWLQ)

Team Workload Questionnaire (TWLQ)

The following questions relate to your experiences during the task. Use the response scale below the question by ticking the circle closest to your answer; the scale goes from 0 (**very low**) to 10 (**very high**).

1. How much physical activity was required (e.g., pushing, pulling, turning, controlling, activating, etc.)?

0 1 2 3 4 5 6 7 8 9 10

2. How much mental and perceptual activity was required (e.g. thinking, deciding, calculating, remembering, looking, searching, etc.)?

0 1 2 3 4 5 6 7 8 9 10

3. How much time pressure did you feel due to the rate or pace at which the tasks or task elements occurred?

0 1 2 3 4 5 6 7 8 9 10

4. How much did you have to control your emotions (e.g., anger, joy disappointment)?

0 1 2 3 4 5 6 7 8 9 10

5. How much did the task require you to monitor your performance (i.e., ensure you were performing at specific levels)?

0 1 2 3 4 5 6 7 8 9 10

6. How hard did you have to work (mentally and physically) to accomplish your level of performance?

0 1 2 3 4 5 6 7 8 9 10

The following questions relate to the team's performance in the task, you should answer them by considering the team as a whole. Use the response scale below the question by ticking the circle closest to your answer; the scale goes from 0 (**very low**) to 10 (**very high**).

7. How much communication activity was required by your team (e.g. discussing, negotiating, sending and receiving messages, etc.)?

○ 0 ○ 1 ○ 2 ○ 3 ○ 4 ○ 5 ○ 6 ○ 7 ○ 8 ○ 9 ○ 10

8. How much coordination activity was required by your team (e.g. changing or adjusting, plans, etc.)?

○ 0 ○ 1 ○ 2 ○ 3 ○ 4 ○ 5 ○ 6 ○ 7 ○ 8 ○ 9 ○ 10

9. How difficult was it to share and manage time between task-work (work done individually) and team-work (work done as a team)?

○ 0 ○ 1 ○ 2 ○ 3 ○ 4 ○ 5 ○ 6 ○ 7 ○ 8 ○ 9 ○ 10

10. How emotionally demanding was working in the team?

○ 0 ○ 1 ○ 2 ○ 3 ○ 4 ○ 5 ○ 6 ○ 7 ○ 8 ○ 9 ○ 10

11. How much did the task require you to monitor your team's performance?

○ 0 ○ 1 ○ 2 ○ 3 ○ 4 ○ 5 ○ 6 ○ 7 ○ 8 ○ 9 ○ 10

12. How difficult was it to provide and receive support (providing guidance, helping team members, providing instructions, etc.) from team members?

○ 0 ○ 1 ○ 2 ○ 3 ○ 4 ○ 5 ○ 6 ○ 7 ○ 8 ○ 9 ○ 10

Appendix E - Course Configurations

Configuration A

			4				
	5					3	
	1						
							2
			H				

Location 1: Recreational Boat, 10 passengers, no hazard

Location 2: Fishing Boat, 40 passengers, fire hazard.

Location 3: Fishing Boat, 40 passengers, bio hazard.

Location 4: Military Boat, 30 passengers, no hazard

Location 5: Cargo Ship, 30 passengers, oil spill

Configuration B

	4						
						2	
			3				
	5					1	
			H				

Location 1: Military boat, 30 passengers, fire hazard

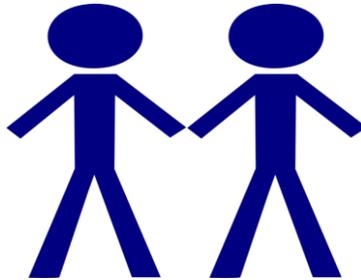
Location 2: Recreational, 40 passengers, oil spill

Location 3: Fishing boat, 10 passengers, bio hazard

Location 4: Cargo boat, 30 passengers, fire hazard

Location 5: Military boat, 20 passengers, no hazard

Appendix F – Stick Figures

10 People on Board**20 People on Board****30 People on Board****40 People on Board**

Appendix G - Hazard Symbols

Oil Spill



Fire



Biological Hazard



Appendix H – Guider Reference Sheet

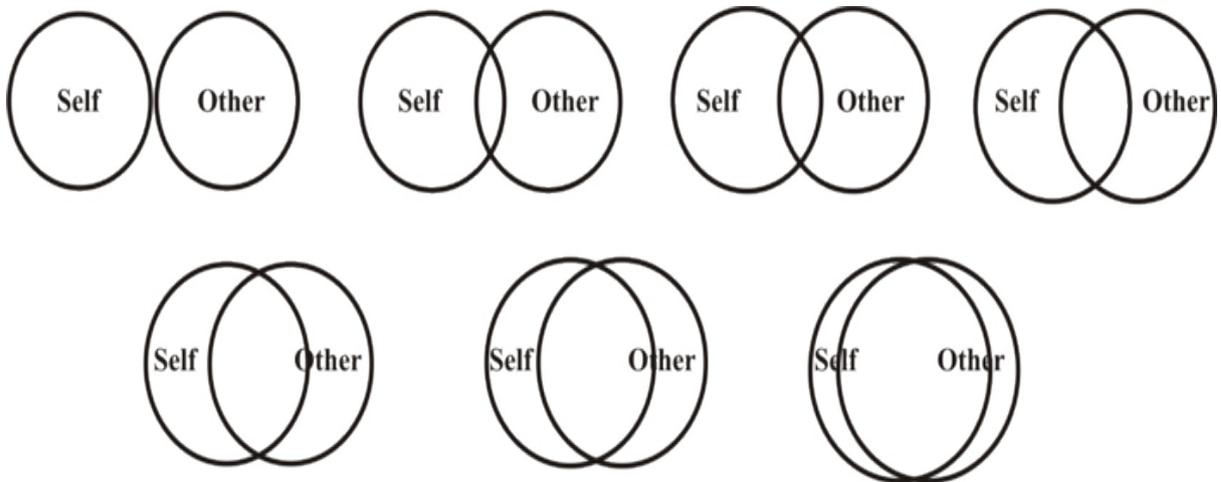
Guider Reference Sheet

 Pink	Fishing Vessel
 Blue	Cargo Vessel
 Green	Military Vessel
 Yellow	Recreational Vessel

Appendix I – Inclusion of Other in Self (IOS) Scale

Role Pilot Spotter

Instructions: Please circle the picture which best describes your current relationship with your teammate.



Appendix J- Quantitative Analysis of Situational Awareness (QUASA) scale

Drone Team Task

Situational Awareness Configuration A

True/False:

(Circle T or F beside the statement)

- | | |
|--|-------|
| There were two cargo boats | T / F |
| Two boats were dealing with fire hazards | T / F |
| There was a recreational boat with 10 passengers | T / F |
| A military boat had 20 passengers | T / F |
| A fishing boat had a biohazard issue | T / F |
| Most boats had 30 or more passengers | T / F |
| Only one boat had no issues/problems | T / F |
| There were two fishing boats | T / F |
| A recreational boat had no issues/problems | T / F |
| A biohazard issue was present on more than one boat | T / F |
| An oil spill issue was present on more than one boat | T / F |
| Only one boat had 10 passengers | T / F |

Drone Team Task

Situational Awareness Configuration B

True/False:

(Circle T or F beside the statement)

- | | |
|--|-------|
| Two boats had fire hazard issues | T / F |
| No boats had more than 40 passengers | T / F |
| A recreational boat had an oil spill issue | T / F |
| A cargo boat had 30 passengers | T / F |
| Most boats had 30 passengers or more | T / F |
| A military boat had no issues | T / F |
| A fishing boat had a fire issue | T / F |
| There was more than one military boat | T / F |
| Two boats had 20 passengers | T / F |
| There were at least 2 recreational boats | T / F |
| A cargo boat had a fire hazard issue | T / F |
| A military boat had 10 passengers | T / F |

Appendix K – Cognitive Map

Drone Team TaskCognitive Map

			H				

Front

The grid you are presented represents the experiment field.

Each “ship” you explored during the experiment was located within one of the squares. For each ship you examined mark its location within the square by drawing the number of the ship relative to the order you identified it. For example, the third ship you examined you would draw a ‘3’ within the box you thought it was located.

Appendix L - Experiment Instruction Sheet

Drone Team Task

Participant Instruction Sheet

You are in a Search-and- Rescue Team. Your aim is to fly the drone across the sea to account for all boats and their contents.

One person in the team is a 'guider' while the other is the 'pilot.' You will swap roles at the end of the first session.

The **pilot will be behind a barrier** and unable to see the drone. They will instead have a view from the drone's down-facing camera projected as a live video feed on to the iPod controller. The **guider's job is to quickly and efficiently guide the pilot from boat to boat by using verbal commands** eg. "go forward... now go to your right and back."

Each boat contains:

- a) Number of passengers on board (1 stick figure = 10 ppl, 2 figures = 20 ppl etc.)
- b) Potentially some sort of problem/issue

These pieces of information can only be seen by the pilot.

The exterior of the boat will be stickered with a colour indicating what type of vessel it is (see reference sheet). This information can only be seen by the guider.

Some information is only available to the guider (type of boat – classified by coloured bit of paper on ship exterior) and **some is only available to the pilot** (number of passengers on board and the type of hazard/issue on the boat, if there is one).

All information must be communicated between team members, e.g., the guider will inform the pilot of the boat type, and the pilot will inform the guider of the passenger numbers and hazards.

Try to remember this information as you will be tested on some of it at the end. You will also be asked to identify locations where boats were present, by drawing on a map. You must not write any of this information down during the flying tasks.

The task is timed and should be completed as quickly and accurately as possible. Timing stops as soon as all 5 boats have been accounted for and the pilot has landed in the heli-pad.

Immediately following the task you will both be presented with some questionnaires.

Also:

You must avoid taking the drone beyond the specified boundaries and crashing.

If the drone crashes there will be a time penalty.

Pilots will have ~5min of flying practice beforehand.

The location of boats and their contents will be randomly reallocated in the pilot-guider changeover.