UNIVERSITY OF CINCINNATI

Date: 02/26/2009

I, Svyatoslav Guznov, hereby submit this original work as part of the requirements for the degree of:
Master of Arts
in Psychology

It is entitled: Teamwork in a RoboFlag Synthetic Task Environment

Student Signature: Svyatoslav Guznov

This work and its defense approved by:
Committee Chair: Gerald Matthews
Joel Warm
Michael Riley

Approval of the electronic document:

I have reviewed the Thesis/Dissertation in its final electronic format and certify that it is an accurate copy of the document reviewed and approved by the committee.

Committee Chair signature: Gerald Matthews
Teamwork in a RoboFlag Synthetic Task Environment

A thesis submitted to the
Division of Research and Advanced Studies
of the University of Cincinnati

in partial fulfillment of the
requirements for the degree of

MASTER OF ARTS

in the Department of Psychology
of the College of Arts and Sciences

by

Svyatoslav Y. Guznov

B.A., Utica College, 2005

Committee Chair:
Gerald Matthews, Ph.D.
ABSTRACT

Effective teamwork is an important component of successful military missions, particularly when multiple unmanned vehicles (UVs) are employed. In this study, multiple UV control was simulated by use of the RoboFlag game in which participants supervised robots in capturing the flags of opponents while protecting their own. Participants played the RoboFlag game individually ($N = 50$) or in two-member teams ($N = 100$). In the team condition, one member actively controlled the robots while the other served as an arbiter or advisor. In both the single player and team conditions, players were presented with the following sequence of task-phases: RoboFlag game by itself, RoboFlag with a secondary task, RoboFlag game by itself. In the secondary task, players were required to listen for and respond appropriately to individually coded auditory messages. It was predicted that the performance effectiveness of the active team players would exceed that of the single players, especially when players were exposed to the need for multi-tasking. It was also expected that single players would have elevated levels of subjective stress and workload in comparison to both the active and arbiter team players. Finally, an attempt was made to determine if a general model of personality traits known as the Big Five -- Openness, Conscientiousness, Extraversion, Agreeableness, and Neuroticism (Goldberg, 1992) -- influenced performance, stress, and workload in the team condition of the RoboFlag game.

There were no significant differences between single players and the active team players in the RoboFlag game. However, single players performed more effectively on the secondary task than active team players. Differential stress and workload patterns were found for single and team players that depended upon task conditions. As measured by the Dundee Stress State Questionnaire (Matthews et al., 2000), single players evidenced lower levels of task engagement in the multi-task and the second exposure to the RoboFlag game by itself task conditions than
did the active team players and lower levels of engagement than arbiter team players on the second exposure to the RoboFlag game by itself. As measured by the NASA-Task Load Index (Hart & Staveland, 1988), both the active and arbiter team players had higher levels of workload than single players on the RoboFlag alone conditions but there were no differences among the players in the multi-task condition. In sum, the results portray a complex picture of single player, team, and task combinations in regard to performance, stress, and workload in the RoboFlag simulation of UV control. Overall, the results also indicated that the Big Five trait model may be too broad to be effectively applied to predicting team performance, stress, and workload in the RoboFlag scenario.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>TABLE OF CONTENTS</th>
<th>v</th>
</tr>
</thead>
<tbody>
<tr>
<td>LIST OF TABLES</td>
<td>vii</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>viii</td>
</tr>
</tbody>
</table>

## CHAPTER 1: Introduction
Unmanned Vehicles as a New Technology in the Military ........................................ 1
   Unmanned Vehicle Description ................................................................. 1
   General problem of Unmanned Vehicle Control ......................................... 2
RoboFlag Computer Simulator ........................................................................... 2
   The RoboFlag Task .................................................................................... 2
   Teamwork Issues ........................................................................................ 5
   Team and Teamwork ................................................................................. 6
   Team Workload and Stress ....................................................................... 6
   Team Member Personality ........................................................................ 9
Study Goals and Hypothesis ........................................................................... 10

## CHAPTER 2: Method
Participants .................................................................................................. 12
Design ........................................................................................................ 12
Materials and Apparatus ............................................................................. 12
   RoboFlag ................................................................................................ 12
   Coordinate Response Measure .................................................................. 15
   Dundee Stress State Questionnaire (DSSQ) .......................................... 18
   Perceived Mental Workload (NASA-TLX) ............................................. 18
   Mini Markers Questionnaire (3M40) .................................................... 19
Procedure ...................................................................................................... 19

## CHAPTER 3: Results
Data Analysis Overview ............................................................................. 22
Performance Measures ............................................................................. 22
   RoboFlag ................................................................................................ 22
   Coordinate Response Measure ............................................................... 26
Subjective Measures .................................................................................. 28
   Dundee Stress State Questionnaire ....................................................... 28
   Perceived Mental Workload ................................................................. 31
   Mini Markers Correlations ................................................................. 32
Gender Differences Analysis ...................................................................... 37

## CHAPTER 4: Discussion
Summary of Key Findings .......................................................................... 39
Influences on Performance and Subjective Stress ...................................................... 40
Performance Measures .................................................................................................. 40
Subjective Stress .............................................................................................................. 41
Subjective Workload ...................................................................................................... 42
Personality Measures .................................................................................................... 43
Gender Differences ......................................................................................................... 43
Conclusions and Implications .......................................................................................... 44

REFERENCES .................................................................................................................. 48

APPENDIX A: Pre-task Dundee Stress State Questionnaire ............................................. 53
APPENDIX B: Post-task Dundee Stress State Questionnaire ............................................. 55
APPENDIX C: NASA-TLX Questionnaire ...................................................................... 57
APPENDIX D: Mini Markers Questionnaire .................................................................... 59
APPENDIX E: Subjective Workload Scales Graphs .......................................................... 60
LIST OF TABLES

1. Correlations between Big Five personality traits and subjective stress states ...............34

2. Correlations between Big Five personality traits and subjective workload .................36
**LIST OF FIGURES**

1. Unmanned Aerial Vehicle (left), Unmanned Ground Vehicle (center), and Unmanned Underwater Vehicle (right) ................................................................. 1
2. RoboFlag Player screen ......................................................................................... 3
3. RoboFlag Arbiter screen .......................................................................................... 4
4. RoboFlag screenshot of Player screen .................................................................... 14
5. RoboFlag screenshot of Player screen .................................................................... 14
6. Coordinate Response Measure (CRM) screenshot .................................................. 15
7. RoboFlag and CRM screenshot of player screen ..................................................... 16
8. RoboFlag and CRM screenshot of arbiter screen .................................................... 17
9. RoboFlag game team condition set-up ................................................................. 17
10. Overall players’ performance across the task phases. Error bars are standard errors. 23
11. Players’ offense performance across the task phases. Error bars are standard errors. 24
12. Players’ capture the flag task performance. Error bars are standard errors. .......... 25
13. Players’ number of flags brought home. Error bars are standard errors............. 25
14. Players’ defense performance. Error bars are standard errors.......................... 26
15. Players’ CRM reaction time. Error bars are standard errors.............................. 27
16. Players’ CRM correct responses ratio. Error bars are standard errors............... 28
17. Players’ levels of engagement across the task phases. Error bars are standard errors. 29
18. Players’ levels of distress across the task phases. Error bars are standard errors..... 30
19. Players’ levels of worry across the task phases. Error bars are standard errors. .... 30
20. Change scores for the DSSQ factors. Error bars are standard errors............... 31
21. Players’ overall workload across the task phases. Error bars are standard errors. 32
CHAPTER 1

Introduction

Unmanned Vehicles as a New Technology in the Military

*Unmanned Vehicle Description.* One of the promising directions in military operations is the use of unmanned vehicles (UVs), which is becoming increasingly wide-spread (Parasuraman, Galster, & Miller, 2003). In the near future, a large number of UVs might be used for military purposes. For example, the *RAND* institution proposed the project Low Cost Autonomous Attack System (*LOCAAS*), which could use up to 1000 UVs (Vick, Moore, Pirnie, & Stillion, 2001). A UV is a robotic vehicle that can move about and navigate through terrain with varying degrees of autonomy. Therefore, UVs primarily require supervisory as compared to manual control from an operator (Parasuraman, Galster, & Miller, 2003). UVs have several advantages over manned vehicles including system endurance, minimal human risk, and possibilities for new mission strategies. Figure 1 shows the major types of UVs including Unmanned Aerial Vehicles (UAVs), Unmanned Ground Vehicles (UGVs), and Unmanned Underwater Vehicles (UUVs). Their technical characteristics (e.g. radar range, speed, maneuverability) differ to accommodate to the environments in which they operate. However, regardless of the UV type, supervisory control is a challenging task for humans.

![Figure 1. Unmanned Aerial Vehicle (left), Unmanned Ground Vehicle (center), and Unmanned Underwater Vehicle (right).](image)
General Problem of Multiple Unmanned Vehicle Control. In recent studies by Ruff, Narayanan and Draper (2002), it was shown that a single operator cannot effectively control more than four UVs, which emphasizes the importance of teamwork when more than four UVs are employed. Effective teamwork would reduce cost and improve efficiency of the operations involving UVs; failure to perform adequately could lead to incomplete task performance, UV damage, and loss of human life.

RoboFlag Computer Simulator

The RoboFlag Task. It is important to investigate various questions related to teamwork in multiple UV control and its dynamics in a realistic environment. One of the promising directions in studying teamwork of multiple UV control is the employment of computer simulators, which may bring some of the complexity of the real world environment into controlled laboratory settings. The RoboFlag computer program has been shown to be a valid simulator of the multiple UV control task. RoboFlag was developed by researchers at Cornell University as a test bed for research on artificial intelligence, cognitive engineering, adaptive communication systems, teamwork, and other human-computer interaction applied fields (D’Andrea & Babish, 2003; Campbell et al., 2005). The RoboFlag program is similar to the game Capture the Flag. Sample screenshots are shown in Figures 2 and 3. Two teams of players (Red Team and Blue Team) control multiple robots. The robots of the Red Team are controlled by the computer and are represented by red squares (3b in Figure 2), while the robots of the Blue Team are represented by blue circles (3a in Figure 2). Each player’s robot has a yellow region around it which graphically represents the robot’s field of view, similar to radar. Once an enemy robot is in that area it becomes visible. Otherwise the players are not able to see enemy robots.
The team goal is to capture as many of the opponent’s flags (5a or 5b) as possible and bring them back to their home (1a or 1b) territory while protecting their own flag. The RoboFlag software facilitates different game scenarios. For example, it allows changes in robot parameters (e.g., speed, sensor angle and radius, fuel limit, etc.) and in the number of robots and their field coordinates. The RoboFlag software also includes the Arbiter screen (Figure 3) function which allows two players to play as a team. The second player observes the Arbiter screen, which discloses both teammate and enemy robot locations. This complementary information allows more efficient joint team action when compared to the single player condition. In addition,
RoboFlag allows the use of several computers connected to a network for a multiplayer setup, making it a valuable tool in research studies on teamwork (D’Andrea & Babish, 2003).

Several studies have been conducted using the RoboFlag simulator. Parasuraman et al. (2003) investigated the effects of opponent robot role (i.e. offensive, defensive, or mixed) and environmental uncertainty manipulated by three levels of the robots’ visual range (low, medium, or high) on players’ performance and choice of supervisory control strategies. They found that opponent robots’ posture and visual range significantly affected the choice of control strategies and performance. Shankar, Jin, Su, Adams, and Bodenheimer (2004) evaluated the RoboFlag version 2.0 interface and two improved versions and their effect on operators’ situational awareness. The results of this study show the potential for using the RoboFlag simulator as tool
for developing an efficient interface for multiple UV control. Squire, Trafton, and Parasuraman (2006) investigated the problem of switching between tasks in multiple UV environments. The results of the study suggested that not only switching among the tasks affects the performance but also switching between different strategies. The study also showed the advantage of using automated functions built into the RoboFlag software which improve performance. Finally, Funke, Galster, Nelson, and Dukes (2006) investigated the possibility of using instant messaging (IM: Computer software that permits messages to be sent and received immediately via the internet) as an alternative to face-to-face communication. The results of the study indicate that IM could be as effective as face-to-face collaboration, which shows the value of IM as a communicative tool in UV control.

**Teamwork Issues.** Although the studies just described have shown that the RoboFlag game may be an effective tool in simulating UV control, they have essentially focused upon environmental and interface issues and have not addressed the question of teamwork itself in the exercise of supervisory control. Accordingly, the focus of the present study was on the teamwork dimension. More specifically, UV control is a challenging task posing various difficulties in creating effective teams including team member selection, team training, increased stress, and workload demands. The supervisory duties of UV operators include weapon release, strategy selection, and risk evaluation. These tasks can be significantly complicated by having a large number of UVs to control and by teamwork organization issues. In addition, UVs are heavily automated, which redefines the nature of the task for team members controlling them. Automation impacts individual and team performance by replacing physical with cognitive demands, which is accompanied by increased cognitive workload, stress, and reduced situational
awareness (Parasuraman, Galster, & Miller, 2003). Finally, personality traits of the team members might influence team performance (Driskell, Goodwin, Salas, & O’Shea, 2006). Therefore, in order to make team UV control effective, it is necessary to look at teamwork, stress, workload, and team member personality. The dimensions of team stress and workload and team member personality were the primary issues of interest in this investigation.

**Team and Teamwork.** In military settings, almost any mission is performed in a team context, allowing completion of a task that cannot be handled effectively by an individual alone (Bell, 2007). “Team” is defined as a structured collection of two or more people who have common goals that are accomplished by interacting interdependently and dynamically (Cannon-Bowers & Salas, 1997). The benefits of performing as a team depend on team members exercising their complementary skills. Within a military team, organized command and control is an important part of a complex information processing system that includes decision making activity. The commander makes decisions about the actions to execute. However, the decisions of the commander are heavily dependent on the information supplied by the other team members. Teamwork is the product of a team, whose performance has a complex, dynamic and multilevel nature (Rasmussen & Jeppesen, 2006).

**Team Workload and Stress.** The workload of a task could potentially affect the success of the team so that even an effective team can fail to perform during spikes of demand or constant demand overload. Hart and Staveland (1988) define workload as the perceived relation between the amount of mental processing capability and the amount of processing required by the task. Even workload distribution is important because failure or overload of one team member can affect the entire team. Individual workload level monitoring carries the information about the
components that are needed to be improved upon in organizational or training aspects. The issue of workload and teamwork is directly relevant to the UV control environment.

High workload demand in multiple UV control elevates stress in team members (Cummings & Guerlain, 2007). According to Kowalski-Trakofler, Vaught and Scharf (2003), stress affects team members by narrowing the focus of attention, impacting decision making, and impairing judgment. Cox (1978) identified three definitions of stress including stimulus, response, and transactional definitions. The stimulus definition refers to the external event that produces a certain stress response in individual (e.g. death of a relative, marriage, illness etc.). The disadvantage of the stimulus definition is that the reaction to a life event, for example, the death of a relative, varies from individual to individual, making the magnitude and nature of stress vary as well. The response definition of stress refers to certain physiological (e.g., arousal of the autonomic nervous system) or psychological (e.g., anxiety) indicators of stress (Matthews, 2000). The response-based definition also shares a similar problem with the stimulus-based definition of failing to account for individual differences, which causes poor intercorrelations of different stress measures making it difficult to obtain any single, definitive evaluation of stress in the individual. The transactional model of stress (Lazarus & Folkman, 1984), as the main cognitive model of stress, accounts for variability in stress responses by attempting to explain stress in terms of the interaction between an external event and the individual’s internal processing of the demands imposed by the event (Matthews, 2000). An external event causes an individual to appraise it by comparing its demands to his/her own resources for coping. The perceived lack of personal resources to deal with the demands causes stress, which may be expressed in a variety of different responses. The transactional model is dynamic, in that stress
outcomes vary across time depending on the demand/resource balance. Cognitive appraisals lead to specific actions or behavioral coping strategies, which allow an individual to compensate for excessive environmental demands, as well potentially maladaptive emotion-focused and avoidance strategies.

The transactional model of stress has also guided research on the assessment of stress responses in performance settings. Matthews et al. (1999) proposed a three-dimensional model of stress on the basis that different types of transactions between the operator and the task environment elicit different subjective responses. The Dundee Stress State Questionnaire (DSSQ) is a measure of stress based on this model. It assesses three broad subjective state dimensions labeled Task Engagement, Distress, and Worry. Task Engagement refers to the individual’s energy, task motivation, and concentration. Distress refers to negative moods and lack of perceived control. Worry primarily refers to negative self evaluations and self-referent distracting thoughts. Studies of various tasks used to investigate stress response showed that the DSSQ component responses varied greatly and were task-dependent. For example, high-workload vigilance tasks reduce task engagement (Matthews, et al., 2002). By contrast, working memory tasks, which impose high time pressure but are less monotonous than vigilance tasks, increase Engagement and Distress, but lower Worry. Overall, it was shown that subjective responses cannot be just reduced to arousal dimensions since the individual’s cognitions of task demands play an important role in determining various qualitatively different patterns of stress response (Matthews, 2002).

Team Member Personality. According to Paris, Salas, and Cannon-Bowers (2000), successful performance of a team depends not only on knowledge, skills and attitudes, but also
on the traits of individual team members that facilitate team interaction and functioning (learning ability, initiative, adaptability, risk taking, etc.). Personality traits are considered to be stable over extended periods of time, making personality an important factor in team member selection. Through many decades there were many personality traits investigated in the context of various personality frameworks but a mutual agreement on trait structure was not achieved (Vollrath, 2001). Recently, it has been argued that the Big Five personality traits provide a consistent framework that integrates multiple personality models and constructs (Goldberg, 1992). The ‘Big Five’ personality model defines five major personality characteristics including Openness, Extraversion, Neuroticism, Conscientiousness, and Agreeableness (Goldberg, 1992). Openness reflects the facets of Fantasy, Aesthetics, Feelings, Actions, Ideas, and Values. Extraversion associates with the facets of Warmth, Gregariousness, Assertiveness, Activity, Excitement Seeking, and Positive Emotions. Neuroticism associates with the facets of Anxiety, Angry-Hostility, Depression, Self-Consciousness, Impulsiveness, and Vulnerability. Conscientiousness reflects the facets of Competence, Order, Dutifulness, Achievement Striving, Self-Discipline, and Deliberation. Finally, Agreeableness associates with the facets of Trust, Straightforwardness, Altruism, Compliance, Modesty, and Tender-Mindedness (Matthews, Deary & Whiteman, 2003).

Paris et al. (2000) found that Conscientiousness, Agreeableness, and Extraversion are positively correlated with team performance, while Neuroticism is negatively correlated with superior team performance. Similar findings were reported by Peeters, Van Tujil, Rutte, and Reymen (2006) who found that Agreeableness and Conscientiousness are positive predictors of
team performance. However, they found that Extraversion of team members did not appear to influence team performance.

Personality has important implications for the transactional theory of stress because it influences appraisal of the stressful situation and the choice of coping mechanisms. For example, the Neuroticism trait contributes to stress and relates to greater stress exposure, negative situation appraisals, lower appraisals of own resources, and therefore less effective coping (Vollrath, 2001). Neuroticism reliably predicts negative affect and tense arousal in performance settings. High Neuroticism people tend to experience more negative moods, which might adversely affect prolonged coping with a stressor (Matthews, Deary & Whiteman, 2003). Other traits are less reliably related to stress, but may moderate stress response in some circumstances. Extraversion might increase the chances of stress exposure, because extraverts tend to be impulsive. Such tendencies may be compensated by more effective appraisal, coping strategies, and positive mood. Conscientiousness contributes to the individual’s longevity due to positive cumulative effects of systematic coping strategies and avoidance of risky behaviors. Thus, there are various pathways through which traits might influence stress in team contexts.

**Study Goals and Hypotheses**

In the present study, the RoboFlag computer simulator was used to investigate how the workload of the RoboFlag synthetic task environment affects performance, workload, and stress in players who perform the task in two-person teams as compared to those who perform the task individually. The task was configured to be demanding and potentially stressful; the player’s robots were out-numbered by enemy robots. Each robot had only a limited field of view, making
it difficult to keep track of the enemy. In the team condition, one team member (the arbiter) viewed additional information about the enemy robots, which he or she was able to communicate to the other team member (the active player), who directly controlled the robots. In addition, workload was experimentally manipulated by including a multi-task condition in which players had to respond to auditory messages during game-play. It was predicted that performance effectiveness in the team condition would exceed performance in the single player condition, especially when participants were exposed to the need for multi-tasking. It was also expected that the participants in the single player condition would experience higher workload, due to their relative lack of information about the enemy robots, compared with the team players. Thus, single players were expected to display elevated stress in comparison to the team players as reflected by lower task engagement and greater levels of worry and distress as measured by the DSSQ.

A second goal for the present study was to determine if RoboFlag performance related positively to Conscientiousness, Agreeableness, and Extraversion and negatively to Neuroticism in the team condition, as suggested by Paris et al. (2000). It was also expected that Neuroticism would relate to stress responses, indexed by the DSSQ, especially in more demanding task conditions.

Although gender was not a major focus of the study, data were analyzed for gender differences. RoboFlag requires spatial cognition, and, on some spatial tests, males tend to perform better than females (Voyer, Voyer & Bryden, 1995). Conceivably, performance differences might be associated with gender differences in stress elicited by the task.
CHAPTER 2

Method

Participants

One hundred fifty undergraduate students from the University of Cincinnati (75 men, 75 women) participated in the study to fulfill a course requirement. Participants ranged in age from 18 to 24 years with a mean of 19.55 years ($SD = 1.5$). They had normal or corrected-to-normal vision and were fluent in English. All participants were treated ethically according to standards established by the American Psychological Association (APA, 2002).

Design

Fifty participants (equated for sex) were assigned at random to play the RoboFlag game individually. The remaining one hundred participants (also equated for sex) were assigned at random to play the game in two-person teams. Within each team, one member actively controlled the robots while the other served as an arbiter or advisor (see below). In both the single player and team conditions, players were presented with the following sequence of task-phases: RoboFlag game by itself, RoboFlag with a secondary task, and RoboFlag game by itself. The secondary task is described below. The final phase of single-task performance provided a control for possible practice effects.

Materials and Apparatus

RoboFlag. RoboFlag 2.1 computer software was used to simulate the multiple UV control environment, with each robot representing a UV. The participant’s goal was to capture the Red Team’s flag and bring it back to their home territory, while protecting their own flag. The participants controlled three robots of the Blue Team playing against the computer, which
controlled six robots of the Red Team (Figure 4). In order to control the robot, the player first needed to select the robot via a mouse left-click, and then move the pointer of the mouse to the destination point finalized by the mouse right-click. The robots moved on the shortest line possible avoiding any obstacles. Each player’s robot had a yellow region around it which graphically represented the robot’s field of view. Once the enemy robot was in that area it became visible as a red square and could be intentionally tagged. Otherwise, the players were not able to see the opposing robots. Individual and active team players scored every time the flag was captured and once more if it was brought safely to the home zone. Also, they scored every time they tagged the opponent’s robot when in their own territory. Once the robot was tagged it became disabled and went back to the home zone by the shortest path route. In the home zone, it was repaired and sent automatically back to the place where it was originally tagged. The scores were not disclosed to players. The game screen was presented on the player’s screen. The RoboFlag software also supported the Arbiter screen function which allowed the second player in the team condition to observe the game screen which disclosed both team mate and enemy robot locations (Figure 5).
Figure 4. RoboFlag screenshot of Player screen

Figure 5. RoboFlag screenshot of Arbiter screen
Coordinate Response Measure. The Coordinate Response Measure (CRM) software (Bolia, Nelson, Ericson, & Simpson, 2000) was used to simulate a military radio communication assignment in which frequency channels were required to be switched manually upon receipt of a coded message to prevent message interception by the enemy. The participants had to manually select one of the 32 buttons presented on the computer screen, which represented 32 radio frequencies. Each button was color and number coded (Figure 6). Participants were assigned unique code names (e.g. “Vector”, “Eagle” etc.) and were required to perform the appropriate selection when their code name was called. The message template was “Ready <Participant’s Code Name> go to <color of the button> <number of the button> now”. The CRM task was scheduled on a random basis with average presentation frequency of four events per minute. The gender of the software voice was the opposite of the participant’s gender to reduce interference with speech.

![Figure 6. Coordinate Response Measure (CRM) screenshot](image)
Each PC running these programs featured a 2.2 GHz PIV Dual Core processor, a 19-inch monitor with the resolution of 1440×900 pixels (single monitor set-up for the single player or team player and dual monitor for team arbiter), dual-head 512Mb video card and headset.

In the single player set-up only one computer with a single monitor was used. The single player monitor showed the RoboFlag game screen in the RoboFlag only conditions and the RoboFlag game screen with embedded CRM task screen in the multi-task condition (Figure 7). A headset was used to deliver the CRM task commands. In the team player set-up, the active team player had a computer configuration identical to the single player while the team arbiter had two monitors with one showing the RoboFlag arbiter screen and another showing the CRM task screen in the multi-task condition (Figure 8). Both team player and team arbiter wore headsets to receive CRM task commands.

Figure 7. RoboFlag and CRM screenshot of Player screen
Team members used verbal communication for the primary task. Team players were positioned at a 90° angle which provided uninterrupted communication while preventing the team members from looking at each other’s screens. Figure 9 represents the actual set-up of the team player condition.
The RoboFlag software recorded the primary task performance data including number of flags captured, number of flags brought home, and number of opponent tags. The CRM software recorded the secondary task performance data including the number of selections made, number of correct selections, and reaction time in making selections.

*Dundee Stress State Questionnaire (DSSQ).* The Dundee Stress State Questionnaire (DSSQ: Matthews et al., 2002) is a 96-item measure of subjective stress which was factor-analyzed and reduced to 10 correlated first-order dimensions. These 10 scales are Energetic Arousal (alertness-sluggishness), Tense Arousal (nervousness-relaxation), Hedonic Tone (happiness), Intrinsic Task Motivation, Self-Focused Attention (self-awareness), Concentration, Self-Esteem, Confidence and Control, Task-Relevant Cognitive Interference, and Task-Irrelevant Cognitive Interference. The dimensions were further grouped into higher-order factors consisting of Task Engagement, Distress, and Worry on the basis of a further factor analysis of the first-order scales (Matthews et al., 2002). Factor scores for the three higher-order dimensions may be estimated by using regression equations obtained from the normative sample. The DSSQ can be administered before and after task performance to gauge changes in subjective states induced by the task. Pre-test DSSQ assesses the Engagement, Distress, and Worry experienced before the task and the post-task DSSQ measures those dimensions after performing the task. In this study, short 32-item pre-task and post-task DSSQ versions were used, which have been shown to be a valid alternative to the full-size DSSQ (Matthews, Emo, & Funke, 2005). Copies of the short versions of the DSSQ can be found in Appendices A and B.

*Perceived Mental Workload (NASA-TLX) (Appendix C).* The NASA-Task Load Index (NASA-TLX; Hart & Staveland, 1988) is a multi-dimensional rating procedure that derives an overall workload score based on a weighted average of ratings on six subscales. The total of the
subscale ratings is called global workload. It is calculated from six 100-point subscales that include Mental Demand, Physical Demand, Temporal Demand, Performance, Effort and Frustration. The TLX is a standard and widely used measure of subjective workload (Wickens & Hollands, 2000). An unweighted scoring system was used to compute global workload in this study.

*Mini Markers Questionnaire (Appendix D).* The original test developed by Goldberg (1992) to assess the Big Five personality traits consists of 100 items, and is often unsuitable when the time to administer a test is limited. Accordingly, the Mini-Modular Markers Test (3M40; Saucier, 2003), a shorter version of the original test, was utilized in this study. This instrument consists of only 40 items, which is convenient for fast screening assessments (Saucier, 2003). It requires the participants to rate on a 10-point Likert scale their personality in relation to a series of adjectives with an upper score of 10 (very accurate descriptor) and a lower score of 1 (very inaccurate). A copy of the Mini-Modular Markers Test is presented in Appendix C.

**Procedure**

Prior to the beginning of the experimental task participants signed the consent forms and were randomly assigned to either single player (SP) or team player conditions. In the team condition participants were further assigned to either team player (TP) or team arbiter (TA) roles. Each participant filled out the Mini Markers personality test to identify personality traits, followed by a pre-task DSSQ which assesses mood and feelings prior to the task. Next, participants received a ten minute practice session to ensure that the necessary task skills were gained. The practice session consisted of several stages. In the first stage participants (SP or TP
and TA) were given an explanation of the RoboFlag general game rules and the participants’ roles in the task. In the team condition the participants were asked to sit side by side and the screen were turned in the same direction in order to familiarize team players with both screens. In the second stage the players (SP and TP) were shown how to control the robots and to capture and bring home the flags. After that SP or TP were asked to practice these skills for two minutes. In the third stage the players (SP and TP) were shown how to tag the enemy robots followed by a two-minute practice period. In the team condition during stages two and three the TA was asked to observe the action of the TP. In the fourth stage participants were asked to play a five-minute practice game with the task of capturing as many flags as possible and bringing them back to the home zone while protecting their own flag. In the single player condition, participants just played the game. In the team condition participants then were asked to play the practice game as a team at their own stations. Team members were allowed to communicate verbally without looking at each other’s screen. The TA observed the screen during the game and provided the TP with information about the location of the enemy robots and possible game strategies.

After the RoboFlag game skills were practiced the participants practiced the secondary task. Each participant was assigned a call name (i.e. ‘Arrow’ for the SP and TP or ‘Tiger’ for the TA) and was asked to put the headphones on. Participants were to make an appropriate selection by pushing the button of appropriate color and number after they heard the command. For example, when SP heard the command “Ready Arrow go to blue five now”, he/she had to locate and select the button of the specified color and number in the CRM window. The duration of the secondary task practice was one minute.
Upon the completion of the practice phase participants performed the three task conditions with six minute duration for each task. After each game condition, participants were asked to rate their subjective workload and stress using the computer-based DSSQ and the NASA-TLX questionnaires. After the experimental trials participants were asked about their video game experience expressed in number of hours of video games played per week. The duration of the entire experiment was approximately 90 minutes.
CHAPTER 3

Results

Data Analysis Overview

The data analysis included the following parts. The RoboFlag performance data (number of flags captured, number of flags brought home, and number of opponent tags) were used to calculate players’ overall, offense, and defense performance scores, which were individually analyzed using $2 \times 3$ mixed-model analysis of variance (ANOVA) procedures. The CRM performance data (number of responses made, number of correct number responses, and number of correct color responses) were used to calculate players’ correct responses ratios. Players’ correct response ratios and players’ reaction times were individually analyzed using $3$ one-way ANOVAs. The subjective stress (DSSQ) factors of engagement, distress, and worry were individually analyzed using $3 \times 3$ mixed-model ANOVAs. The overall, unweighted subjective workload score was calculated and individually analyzed using $3 \times 3$ mixed-model ANOVAs. The personality data were correlated with the performance (both RoboFlag and CRM), subjective stress and subjective workload data. Finally, the gender difference analysis included separate analyses of performance data (both RoboFlag and CRM), subjective stress and subjective workload data using ANOVA procedure.

Performance Measures

RoboFlag
Overall RoboFlag Players’ Performance. The overall performance was defined as the difference between the player and computer performances in number of flags captured, brought home and enemy robots tagged. The following formula was used: $\text{Overall Performance} = \Delta \text{ in number of flags captured by the player and computer} + \Delta \text{ in number of flags brought home by the player and computer} + \Delta \text{ in number of enemy robot tags by the player and computer}$. A $2 \times 3$ mixed-model ANOVA revealed a significant main effect for the task phase, $F(2,196) = 6.252, p < .01$, partial $\eta^2 = .06$, with the Huynh-Feldt adjustment applied for violation of the sphericity assumption. No significant main effect for player type or interaction was found. The analysis of contrasts for the task phase revealed no significant differences between the task phase levels, although a general trend towards performance improvement over time was evident. The overall players’ performance results are shown in Figure 10.

![Figure 10. Overall players’ performance across the task types. Error bars are standard errors.](image)

RoboFlag Players’ Offense Performance. The offense performance index was defined as the player’s ability to capture the enemy flags and bring them back to the home zone while avoiding being tagged by the computer. Computer tags prevent the player from both getting the flag and
bringing it to the home zone. The overall offense performance score was computed using the following formula: \( \text{Offense Performance} = \text{Number of flags captured by the player} + \text{Number of flags brought home by the player} + \text{Number of player robot tags by the computer} \). A \( 2 \times 3 \) mixed-model ANOVA revealed a significant main effect for task phase, \( F(2,196) = 22.22, p < .01, \text{partial } \eta^2 = .18 \), no significant main effect for the player type, and no interaction. The analysis of contrasts for the task phase revealed significant \( (p < .01) \) differences between all task phase levels. The offense performance in the X task phase \( (M = -2.82, SD = 13.00) \) was lower than in the D task phase \( (M = 1.58, SD = 13.70) \), \( F(1, 98) = 15.90, p < .01, \text{partial } \eta^2 = .14 \). The offense performance in the D task phase \( (M = 1.58, SD = 13.70) \) was lower than in the Y task phase \( (M = 5.24, SD = 13.78) \), \( F(1, 98) = 9.58, p < .01, \text{partial } \eta^2 = .09 \). The offense players’ performance results are shown in Figure 11.

![Graph showing offense performance across task types](image)

*Figure 11. Player’s offense performance across the task types. Error bars are standard errors.*

The offense performance broken down into the number of flags captured and number of flags brought home subcategories are represented in Figure 12 and Figure 13.
RoboFlag Players’ Defense Performance. The defense performance index was defined as a players’ ability to tag the enemy robots in order to prevent them from getting the flag and bringing it back to the home zone subtracting the number of flags captured and brought home by the computer. The defense performance index was computed using the following formula:

\[
\text{Defense Performance} = \text{Number of computer robot tags by the player} - \text{Number of flags captured by the computer} - \text{Number of flags brought home by the computer}.
\]

A 2(player type) × 3(task phase) mixed-model ANOVA revealed no significant main effects and no interaction. The defense performance results are shown in Figure 14.
Coordinate Response Measure

**CRM Players’ Reaction Time.** The CRM reaction time is the speed of the appropriate selection made by the players measured in milliseconds (ms). A one-way ANOVA showed a significant effect for the player type, $F(2,147) = 17.70$, $p < .01$, partial $\eta^2 = .16$. Post-hoc analysis using $t$-tests with the Bonferroni criterion for significance testing showed that reaction time for the single players ($M = 3852.48$, $SD = 639.15$) was significantly lower ($t(98) = -5.52$, $p < .01$) than reaction times for the active team players ($M = 5248.28$, $SD = 1670.99$) and the team arbiters ($M = 5159.46$, $SD = 1404.74$), $t(98) = -5.99$, $p < .01$. The players’ reaction times in the three team conditions are shown in Figure 15.
CRM Correct Responses Ratio. This ratio was calculated by dividing the number of correct responses by the total number of responses made. A one-way ANOVA showed a significant effect for the player type, $F(2,147) = 15.33, p < .01$, partial $\eta^2 = .15$. Post-hoc analysis using $t$-tests with the Bonferroni criterion for significance testing showed that the correct responses ratios for single players ($M = .97, SD = .04$) was significantly higher ($t(98) = -5.47, p = .01$) than the correct responses ratios for active team players ($M = .87, SD = .13$) or team arbiters ($M = .87, SD = .12$), $t(98) = -5.14, p < .01$. The players’ correct responses ratio results are shown in Figure 16.
Subjective Measures

Dundee Stress State Questionnaire

DSSQ Reliability. No significant differences in pre-task engagement, distress, and worry were revealed among the groups, indicating that their subjective states were similar prior to performance. The internal consistency reliability was assessed for the DSSQ scales and revealed acceptable Cronbach α coefficients for Engagement (α = .84), Distress (α = .84), and Worry (α = .81).

Engagement. A 3 (player type) × 3 (task phase) mixed-model ANOVA revealed a significant main effect for the player type, $F(2,147) = 4.13, p < .05$, partial $\eta^2 = .05$, and a significant interaction, $F(4,294) = 4.29, p < .01$, partial $\eta^2 = .05$. No significant main effect for task phase was found. Task engagement did not vary significantly across groups in the X task phase of the experiment. Post-hoc $t$-test analysis using Bonferroni criterion for significance revealed significantly higher engagement for team players ($M = 24.78, SD = 5.97$) than for single players.
(\(M = 22.00, SD = 6.15\)) and arbiters (\(M = 21.90, SD = 7.48\)) during the D task phase of the experiment (\(p < .05\)). There was significantly higher engagement for team players (\(M = 25.16, SD = 6.18\)) than single players (\(M = 20.12, SD = 7.03, p < .01\)) and arbiters (\(M = 22.48, SD = 7.64, p < .05\)) during the Y-task phase of the experiment. The engagement level was significantly lower for the single player in the Y task phase (\(M = 20.12, SD = 7.03\)) when compared to the X (\(M = 22.52, SD = 6.19\)) and the D (\(M = 22.00, SD = 6.15\)) phases. The results are shown in Figure 17.

Distress. A 3 (player type) \(\times\) 3 (task phase) mixed-model ANOVA revealed a significant main effect for task phase, \(F(2,294) = 11.33, p < .01,\) partial \(\eta^2 = .07,\) no significant main effect for player type, and no interaction. The analysis of contrasts for task phase revealed that the level of distress for the D task phase (\(M = 14.43, SD = 7.06\)) was higher than in the Y task phase (\(M = 12.48, SD = 6.84\)), \(F(1,147) = 22.68, p < .01,\) partial \(\eta^2 = .13.\) Thus, distress appeared to decline in all groups in the final phase of single-task testing. The results are shown in Figure 18.
A 3 (player type) × 3 (task phase) mixed-model ANOVA revealed a significant main effect for task phase, $F(2,294) = 14.37, p < .01$, partial $\eta^2 = .09$, no significant main effect for player type, and no interaction. The analysis of contrasts for task phase revealed that the level of worry for the X task phase ($M = 7.65, SD = 5.63$) was higher than in the D task phase ($M = 6.63, SD = 5.48$), $F(1,147) = 15.47, p < .01$, partial $\eta^2 = .09$. The results are shown in Figure 19.

Figure 18. Levels of distress across task types. Error bars are standard errors.

Figure 19. Levels of worry across task types. Error bars are standard errors.
**DSSQ change scores.** Change scores for the DSSQ factors are shown in Figure 20, to provide a graphical representation of changes induced by the entire task experience for the three player types. Change scores were calculated as final post-task state score (after task phase Y) – initial, pre-task state score. The change scores for engagement revealed an increase in engagement for all player types but they were considerably larger for the team player compared to the other two player roles. The distress change scores showed that distress rose for all player types when compared to the pre-task levels of distress. Finally, worry dropped for all player types when compared to the pre-task levels.

![Graph showing change scores](image)

*Figure 20.* Change scores for the DSSQ factors. Error bars are standard errors.

**Perceived Mental Workload**

**Overall workload.** A 3 (player type) × 3 (task phase) mixed-model ANOVA revealed a significant main effect for task phase, $F(2,294) = 25.15, p < .01$, partial $\eta^2 = .15$, a significant main effect for the player type, $F(1,147) = 6.59, p < .01$, partial $\eta^2 = .08$, and a significant
interaction, $F(4,294) = 3.12, p < .05$, partial $\eta^2 = .04$. The post-hoc analysis using the Bonferroni adjustment revealed significantly lower overall workload ($p < .01$) for the single player ($M = 48.90, SD = 14.39$) than for team players ($M = 56.88, SD = 10.56$) and arbiters ($M = 57.28, SD = 12.48$) during the X-task phase of the experiment. The overall workload did not vary significantly in the D task phase of the experiment. There was significantly lower overall workload ($p < .01$) for single players ($M = 47.77, SD = 13.79$) than for team players ($M = 57.62, SD = 12.28$) and for team arbiters ($M = 57.10, SD = 13.85$) during the Y-task phase of the experiment. The players’ overall workload is shown in Figure 21.

![Graph showing overall workload across task phases]

Figure 21. Overall workload across the task phases. Error bars are standard errors.

Mini Markers Correlations

3M40 and Performance (RoboFlag and CRM tasks). No significant correlations were revealed between the personality traits and RoboFlag task performance. Also, no significant correlations were revealed between the personality traits and CRM task performance.
**3M40 and DSSQ.** The correlational analysis of relations between the MINI Markers questionnaire and DSSQ data revealed modest but significant relationships between the team players’ neuroticism and distress for team players only. In this condition, the correlations were significant for all three experimental task phases, ranging from .28 to .34. In addition, there was a weak positive relationship between team arbiter’s openness and engagement in the X and Y task phases. These and other correlations are presented in Table 1.
Table 1

*Correlations Between the Big Five Personality Traits and Subjective Stress States*

<table>
<thead>
<tr>
<th>DSSQ</th>
<th>Task Phase</th>
<th>Agreeableness</th>
<th>Openness</th>
<th>Neuroticism</th>
<th>Extraversion</th>
<th>Conscientiousness</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>SP</td>
<td>TP</td>
<td>TA</td>
<td>SP</td>
<td>TP</td>
</tr>
<tr>
<td>Eng-t</td>
<td>PRE</td>
<td>.13</td>
<td>.37**</td>
<td>.07</td>
<td>.15</td>
<td>-.04</td>
</tr>
<tr>
<td></td>
<td>X</td>
<td>.15</td>
<td>.06</td>
<td>.04</td>
<td>.19</td>
<td>.02</td>
</tr>
<tr>
<td></td>
<td>D</td>
<td>.12</td>
<td>.12</td>
<td>.03</td>
<td>.18</td>
<td>-.12</td>
</tr>
<tr>
<td></td>
<td>Y</td>
<td>.15</td>
<td>.11</td>
<td>-.06</td>
<td>.00</td>
<td>.02</td>
</tr>
<tr>
<td>Distress</td>
<td>PRE</td>
<td>.09</td>
<td>-.15</td>
<td>-.08</td>
<td>.01</td>
<td>.10</td>
</tr>
<tr>
<td></td>
<td>X</td>
<td>-.08</td>
<td>-.18</td>
<td>.06</td>
<td>-.11</td>
<td>.01</td>
</tr>
<tr>
<td></td>
<td>D</td>
<td>.03</td>
<td>-.23</td>
<td>-.02</td>
<td>-.18</td>
<td>-.05</td>
</tr>
<tr>
<td></td>
<td>Y</td>
<td>.12</td>
<td>-.14</td>
<td>-.01</td>
<td>.11</td>
<td>.18</td>
</tr>
<tr>
<td>Worry</td>
<td>PRE</td>
<td>.13</td>
<td>.15</td>
<td>-.02</td>
<td>.01</td>
<td>.27</td>
</tr>
<tr>
<td></td>
<td>X</td>
<td>.20</td>
<td>.22</td>
<td>.14</td>
<td>.03</td>
<td>.27</td>
</tr>
<tr>
<td></td>
<td>D</td>
<td>.17</td>
<td>.10</td>
<td>.14</td>
<td>.13</td>
<td>.12</td>
</tr>
<tr>
<td></td>
<td>Y</td>
<td>.14</td>
<td>.05</td>
<td>.08</td>
<td>.15</td>
<td>.08</td>
</tr>
</tbody>
</table>

*Note.* **- Correlation is significant at the 0.01 level (2-tailed).
*- Correlation is significant at the 0.05 level (2-tailed).
3M40 and NASA-TLX Workload. The correlational analysis revealed significant correlations between the team player’s neuroticism and TLX frustration in all three experimental task phases ranging from .32 to .46. There were also moderate correlations between the team players’ effort and openness in the D and Y task phases as well as for neuroticism and temporal demand. Team arbiter openness weakly correlated with effort in the X and D task phases. These and other correlations are presented in Table 2. In addition, there were weak but significant correlations between team players’ neuroticism and overall workload in the X and D task phases ($r_X = .34, p < .05; r_D = .32, p < .05$).
Table 2
Correlations Between the Big Five Personality Traits and Subjective Workload Scales

<table>
<thead>
<tr>
<th>Task Phase</th>
<th>Agreeableness</th>
<th>Openness</th>
<th>Neuroticism</th>
<th>Extraversion</th>
<th>Conscientiousness</th>
</tr>
</thead>
<tbody>
<tr>
<td>TLX</td>
<td>Agreeableness</td>
<td>Openness</td>
<td>Neuroticism</td>
<td>Extraversion</td>
<td>Conscientiousness</td>
</tr>
<tr>
<td>Mental X</td>
<td>.18</td>
<td>-.01</td>
<td>.13</td>
<td>.18</td>
<td>.14</td>
</tr>
<tr>
<td>D</td>
<td>-.03</td>
<td>-.15</td>
<td>.00</td>
<td>.10</td>
<td>.01</td>
</tr>
<tr>
<td>Y</td>
<td>.03</td>
<td>-.03</td>
<td>.20</td>
<td>-.13</td>
<td>.09</td>
</tr>
</tbody>
</table>

| Physic X   | .06           | .03      | .02         | -.25         | .16               | .09               | .15               | -.13              | .16               | -.11              | -.23              | .02               | .25               | .11               | .13               |
| D          | .04           | -.03     | .00         | -.18         | .16               | .04               | .10               | -.08              | .11               | -.17              | -.20              | .06               | .25               | .13               | -.08              |
| Y          | .04           | .09      | -.06        | -.24         | .20               | .04               | .06               | -.12              | .24               | -.07              | -.15              | -.09             | .18               | .15               | -.07              |

| Temp X     | .16           | .24      | .32**       | -.18         | .11               | .26               | -.01              | .21               | -.16              | -.26              | .17               | .10               | .30**              | .02               | .23               |
| D          | -.24          | -.11     | -.04        | -.41**       | .07               | .11               | .15               | .30**              | .00               | -.16              | .19               | -.06              | -.12              | -.08              | .06               |
| Y          | -.12          | .08      | .01         | -.23         | .16               | -.01              | .11               | .33**              | -.02              | -.22              | .13               | -.23              | .05               | .04               | .15               |

| Perform X  | -.05          | .12      | .08         | -.04         | -.05              | -.07              | -.01              | .32**              | -.19              | -.26              | .29**             | -.10              | -.03              | -.18              | -.12              |
| D          | -.09          | .14      | .01         | .00          | .12               | -.01              | .06               | .19               | -.11              | .00               | .16              | -.12              | -.05              | -.29**             | .00               |
| Y          | -.23          | .33**    | .04         | .06          | -.06              | -.20              | .24               | .14               | .08               | -.06              | .12               | -.11              | .03               | -.07              | -.12              |

| Effort X   | .00           | .12      | .17         | -.12         | .18               | .32**             | .17               | -.08              | -.13              | .03               | .16              | .07               | .20               | .14               | .16               |
| D          | .09           | .00      | .11         | -.02         | .33**             | .29**             | .15               | .21               | -.19              | .13               | .01              | .03               | .18               | -.03              | .18               |
| Y          | .00           | .27      | .11         | -.17         | .41**             | .18               | .21               | -.06              | -.21              | -.04              | .11              | -.08              | .12               | .09               | .32**             |

| Frustr X   | .22           | .04      | .15         | -.34*        | .18               | .04               | .02               | .46**              | .01               | .21               | -.10             | .13               | .13               | .09               | .01               |
| D          | -.04          | .02      | .02         | -.27         | .14               | .00               | .25               | .38**              | .16               | -.16              | -.15             | .14               | .00               | -.03              | -.01              |
| Y          | .13           | .14      | .02         | -.11         | .14               | -.07              | .10               | .32**              | .20               | -.16              | -.10             | .00               | .16               | .17               | .14               |

Note. ** Correlation is significant at the 0.01 level (2-tailed).
* Correlation is significant at the 0.05 level (2-tailed).
Gender Differences Analysis

RoboFlag Performance and CRM Performance. A 2 (gender) × 3 (task phase) mixed-model ANOVA of the RoboFlag game performance of the team player revealed a significant main effect for gender, $F(1,48) = 7.27, p < .05, \eta^2 = .13$, meaning that men performed better ($M_X = -66.31, SD_X = 23.33; M_D = -65.92, SD_D = 21.17; M_Y = -55.92, SD_Y = 22.87$) than women ($M_X = -80, SD_X = 17.27; M_D = -77.62, SD_D = 24.84; M_Y = -70.29, SD_Y = 21.18$) in the X, D, and Y task phases. No significant main effects or interaction were found for the single player condition in the RoboFlag game performance. No significant gender difference in CRM task performance was found.

DSSQ: Engagement. A 2 (gender) × 3 (task phase) mixed-model ANOVA revealed a significant interaction for the team arbiters, $F(2,96) = 4.47, p < .05, \eta^2 = .08$, meaning that men ($M_X = 24.52, SD_X = 6.32; M_D = 22.22, SD_D = 8.21; M_Y = 22.7, SD_Y = 8.17$) had higher engagement at the beginning of the experimental trials than women ($M_X = 20.85, SD_X = 6.53; M_D = 21.63, SD_D = 6.95; M_Y = 22.30, SD_Y = 7.32$), but in the last task phase engagement became similar. No significant main effects for gender were found for the single player or team player.

DSSQ: Distress. A 2 (gender) × 3 (task phase) mixed-model ANOVA revealed a significant main effect for gender in the single players, $F(1,48) = 4.06, p < .05, \eta^2 = .08$, meaning that men ($M_X = 11.88, SD_X = 6.31; M_D = 12.58, SD_D = 6.75; M_Y = 10.58, SD_Y = 5.33$) had lower distress than women ($M_X = 16.5, SD_X = 7.72; M_D = 15.96, SD_D = 7.62; M_Y = 13.08, SD_Y = 7.13$) in the X, D, and Y task phases. No significant main effects for gender were found for the team players and team arbiters.
DSSQ: Worry. No significant main effects for gender were found for all player types.

NASA-TLX: Overall Workload. 2 (gender) × 3 (task phase) mixed-model ANOVAs revealed a significant main effect for gender in the single players and team players, \([F(1, 48) = 5.91, p < .05, \eta^2 = .11; F(1, 48) = 8.66, p < .01, \eta^2 = .15]\). In the single player condition men (\(M_X = 44.90, SD_X = 14.01; M_D = 51.87, SD_D = 13.38; M_Y = 43.68, SD_Y = 14.28\)) had lower workload than women (\(M_X = 52.60, SD_X = 13.99; M_D = 60.77, SD_D = 12.35; M_Y = 51.54, SD_Y = 12.42\)) in the X, D, and Y task phases. In the team player condition men (\(M_X = 53.42, SD_X = 10.59; M_D = 57.31, SD_D = 13.70; M_Y = 52.75, SD_Y = 12.00\)) had lower workload than women (\(M_X = 60.62, SD_X = 9.35; M_D = 64.48, SD_D = 10.80; M_Y = 62.88, SD_Y = 10.44\)) in the X, D, and Y task phases. No significant main effects for gender were found for the team arbiter.
CHAPTER 4
Discussion

Summary of Key Findings

The results of this study provide mixed support for the original hypotheses. This study was designed, in part, to investigate how the workload of the RoboFlag synthetic task environment affects performance, perceived stress and workload in participants who perform the task in teams, as compared to those who perform the task alone. It was predicted that performance effectiveness in the team condition would exceed that in the single player condition, especially when participants were exposed to the need for multitasking. It was also expected that the single player condition would have elevated stress level in comparison to the stress level in the team condition, as reflected by less task engagement and greater levels of worry and distress. The perceived workload was expected to be higher for the single player vs. team players. A second goal for the study was to determine whether personality traits influence performance. Specifically, it was predicted that Conscientiousness, Agreeableness, and Extraversion would be positively related to team performance, while Neuroticism would be negatively related to team performance but positively related to stress. Overall, the hypothesis that the team would perform better than a single player, especially in the dual task phase, was not confirmed in this study. The results mostly confirm the predictions regarding the effects of player type on engagement level. However, the results do not support the hypothesis of the single player having higher level of distress and worry when compared to the team players. The results did not support the hypothesis that perceived workload of the single players would be significantly higher than in teams. The predictions that Conscientiousness, Agreeableness, and Extraversion would positively relate to
the team performance were not supported. However, it appears that there is a positive relationship between team player neuroticism and distress, regardless of the task phase.

**Influences on Performance and Subjective Stress**

*Performance Measures.* The RoboFlag performance data did not reveal any significant difference in the overall performance of single player vs. team. It appears that in both single player and team conditions, players improved their overall performance due to practice. However, team player performance was positively accelerated over time, while the single player performance was negatively accelerated, suggesting that with more trials, team performance could potentially exceed that of single player performance. Both single player and team offense performance significantly improved from task phase X to task phase Y. The defense performance stayed unchanged regardless of the player or task phase. The CRM task performance revealed a significantly lower performance for the team, reflected in slower reaction time and lower number of correct responses ratio. Overall, the hypothesis that the team would perform better than a single player, especially in the dual task phase, was not confirmed in this study. While single player and team conditions performed similarly on the RoboFlag task, single players performed better on the CRM task, making their overall performance better than that in the team condition. These results, however, do not mean that performing in the group is less beneficial, since it is possible that a team did not have enough time to become cohesive. Paris et al. (2000) emphasized the importance of sufficient team training in order to make team member behaviors efficient. In the current study, while participants received individual skill training, they did not receive any team communication and behavior coordination training, which could have potentially negated the benefits of performing in the team. Salas et al. (2000) proposed several
criteria for successful team training. Team members need to be shown that the skills they acquire during the training have a practical impact on performance. Secondly, the skills need to be challenging enough to keep team members’ interest and require more than just simple repetition to learn. Finally, skills need to be practiced from time to time, and team members need to understand that infrequently used skills are as important as the more commonly used ones.

**Subjective Stress.** The subjective stress data revealed that the engagement level was higher for the team player when compared to the levels of engagement for the single player and team arbiter. It was also shown that the team (i.e. team player and team arbiter) was able to maintain a constant level of engagement, while single player engagement dropped significantly by the last task phase. These results mostly confirm the predictions regarding the engagement level, with the exception that the team arbiter engagement level was lower than that of team player engagement level and similar to that of the single player at the beginning stages. These findings highlight the benefits of performing in a team due to its ability keep team members engaged in the task, which gives better chances to maintain a steady level of engagement, which in turn may have an impact on overall performance. They also suggest that training as part of a team may serve to maintain engagement over time than training alone.

The level of distress dropped significantly for both single player and team player toward the end of the experimental trials, perhaps because practice builds confidence. These results do not support the hypothesis of the single player having a higher level of distress when compared to the team player. A possible explanation for this finding is that for team players and single players, the factors contributing to overall distress could have been qualitatively different. According to Kontogiannis and Kossiavelou (1999), a team’s lack of cohesion, communication

41
interferences, intra-group conflict or role ambiguity can elevate stress in team members, which is particularly relevant considering the team training shortcomings of this study. On the other hand, the transactional model (Lazarus & Folkman, 1984) states that stress is not experienced unless the situation is appraised as stressful. In this study, single players had limited awareness about their own performance, and may therefore have under-appraised the demands of the situation. Team arbiters may have appraised the situation to be more highly demanding due to the teamwork task demand coupled with the responsibility of providing information to the team player. This explanation is consistent with the findings that arbiters experienced elevated distress as well as relatively consistent level of worry throughout the study (although not significantly different from the players’ levels of worry).

Subjective Workload. The results revealed a significantly higher overall workload level for the teams in single task phases. The elevated team workload results in this study could indicate insufficient coordination between team members (i.e. backup, mutual performance monitoring, information flow, etc.) which could have contributed to the overall workload (Paris et al., 2000). Again, the result points to the importance of practice for effective teamwork. Early in practice, there may be an additional workload cost for teams, associated with the need to develop strategies for communication. The secondary task increased the overall workload for the single player while having little effect on the team player and team arbiter. The most likely explanation is that team players partly ignored the secondary task, which is shown by the significantly lower performance when compared to single player performance. Therefore, the perceived workload of the secondary task contributed less to the overall workload for the team players and team arbiters.
Personality Measures. The predictions that Conscientiousness, Agreeableness, and Extraversion would positively relate to team performance were not supported. However, it appears that there is a positive relationship between team player neuroticism and distress, regardless of the task phase. Moreover, a similar positive relationship was found between team player neuroticism and frustration. The analysis used in this study did not address whether the personality traits of the team members need to be combined into one team trait score. According to Driskell et al. (2006) there are four approaches to analyzing team composition. An additive approach refers to the averaging of a certain trait among the team members. A compensatory approach relates to the heterogeneity or variability of traits among team members, suggesting that it may be important to include team members with various personality trait accentuations. A disjunctive approach refers the extremes of the team member personalities and their impact on team performance. Overall, it appears that the Big Five, while shown to be a universal way to capture important features of personality, appears to be too broad to be effectively applied to team performance. Driskell et al. (2006) suggested that more sub-factors are required in order to link personality to team performance. In addition, it is necessary to develop a theoretical foundation of team work in order to have better understanding about specific personality characteristics to apply to team performance studies.

Gender Differences. In the RoboFlag task men performed better than women in the team player condition, but no significant performance difference was found in the single player condition. This finding is somewhat surprising due to the expected higher communicational and facilitating abilities in women (Metcalfe & Linstead, 2003), which could improve team performance. However, it is possible that men in the team player condition were able to use team
arbiter instructions more effectively by extracting more information from them due to men’s better performance on spatial visualization tasks (Hamilton, 1995). That is, women possibly have had difficulty in communicating spatial information in this task paradigm and processing spatial information as players. No significant difference was found in the CRM task. At the beginning of the experimental trials, men in the role of team arbiters tended to have higher engagement than women, but this difference leveled out towards the end of the experiment. This finding is likely due to men’s generally higher interest in computer games. However, the RoboFlag game is not graphically as rich as many modern games, which could explain the decline in engagement in men. Women displayed higher distress in the single player condition, when compared to team arbiter and team player showing possible women’s preference of working in teams (Metcalfe & Linstead, 2003). No significant gender difference in worry was found in this study.

**Conclusion and Implications**

In summary, it appears that performing in a team is not necessarily beneficial in the short run. However, the relative performance acceleration comparison suggests that team performance could potentially be better in the long run, especially when assisted by sufficient team training, although longer training sessions may be needed to demonstrate such an effect. Different roles in the game (i.e. single player, team player, and team arbiter) produce differential subjective stress patterns, which further confirm the usefulness of the transactional stress approach and DSSQ dimensions specifically. Driskell et al. (2006) pointed out that the desirable qualities of the team members are often described in trait terms (e.g. dependable, flexible, cooperative, etc). This study showed only fragmented relationships between the Big Five personality traits and
team performance, stress, and workload, suggesting that using the Big Five personality
terminology may not be as effective when applied to team work. However, high neuroticism
appears to be undesirable especially for team members placed in the player role with direct
control over operations. The demands of game play coupled with the social demands of
interacting with the arbiter may be especially difficult for the high neuroticism individual.

This study has dealt with only a few of the teamwork factors that are relevant to
operational settings. One way to approach issues of teamwork organization in the UV control
environment more generally is to employ the theories of social psychology. However, even
within this specific branch, there is no uniform agreement in terms of the specific components
that create an effective team. With this situation in mind, Salas, Sims and Burke (2005)
conducted a meta-analysis that identified five core components and three mechanisms necessary
for a good team. The five core components include leadership, mutual performance monitoring,
back-up behavior, adaptability, and team orientation. Leadership was identified as an important
component of teamwork since leaders facilitate interdependent actions, provide important
information on team’s standing, and establish behavioral and performance goals. Mutual
performance monitoring refers to the members’ responsibility of monitoring other team member
performance while continuing their own task. Backup behavior refers to providing assistance to
other team members in a form of coaching, working together on the teammate’s task, or taking
over the task if the team-mate is overloaded (Marks, Zaccaro, & Mathieu, 2000). Team
adaptability refers to the ability to adjust to the working environment in order to make teamwork
most efficient. Team orientation refers to the idea of an attitudinal measure of team member
tendency to prefer working as a team.
However, these five team characteristics require driving mechanisms to operate properly. In the same review paper, Salas and his collaborators identified three coordinating mechanisms which include shared mental models, closed-loop communication, and mutual trust. Shared mental models are the structures existing in team cognition, which include knowledge about team roles, interaction patterns of the team members, and the member’s own role within the team. Shared mental models are particularly important when a team is faced with high cognitive demand, which limits information processing, decision making, and action time frame (Salas et al., 2005). When such a condition exists, the shared mental models take the place of direct communication, allowing the team to adjust to a dynamically changing task environment (Stout, Cannon-Bowers, & Salas, 1996). Communication refers to an exchange of information between two or more individuals. In UV control, communication among two team members could provide them with the opportunity of strategy development, as well as exchanging task-related information, and could potentially improve situational awareness of team members. Teamwork is heavily based on mutual trust among the team members. The lack of trust draws on a significant amount of energy sources for verification of work performed by other team members, and may also separate the team members into independently working individuals and evoke negative attitude towards each other (Simons & Peterson, 2000). Consideration of these factors is potentially important when studying teamwork of UV control environment.

It would be interesting to further explore the subjective stress and workload responses in individual vs. team RoboFlag task performance by upgrading the following components of the experimental setup. Firstly, longer practice and experimental trials could further improve teams’ readiness for the task, making the teamwork environment closer to the real team work...
conditions. Secondly, it is necessary to employ new measures assessing team-specific personality traits, behavioral teamwork performance, and situational awareness. Finally, non-invasive physiological stress measurement of physiological stress indicators (e.g. heart rate, galvanic skin conduction, or mouse pressure strength) could provide a more fine-grained account of how stress responses change dynamically during teamwork assignments (Liao, Zhang, Zhu, Ji, & Gray, 2006).
REFERENCES


Appendix A

DSSQ-3 STATE QUESTIONNAIRE

PRE-TASK QUESTIONNAIRE

Instructions. This questionnaire is concerned with your feelings and thoughts at the moment. Please answer every question, even if you find it difficult. Answer, as honestly as you can, what is true of you. Please do not choose a reply just because it seems like the 'right thing to say'. Your answers will be kept entirely confidential. Also, be sure to answer according to how you feel AT THE MOMENT. Don't just put down how you usually feel. You should try and work quite quickly: there is no need to think very hard about the answers. The first answer you think of is usually the best.

Before you start, please provide some general information about yourself.

Age............. (years)                                         Sex.   M  F   (Circle one)
Occupation.....................                                If student, state your major..............
Date today.....................                                Time of day now..............

For each statement, circle an answer from 0 to 4, so as to indicate how accurately it describes your feelings AT THE MOMENT.

Definitely false = 0, Somewhat false = 1,
Neither true nor false = 2, Somewhat true = 3, Definitely true = 4

1. I feel concerned about the impression I am making.  0 1 2 3 4
2. I feel relaxed.  0 1 2 3 4
3. The content of the task will be dull.  0 1 2 3 4
4. I am thinking about how other people might judge my performance.  0 1 2 3 4
5. I am determined to succeed on the task.  0 1 2 3 4
6. I feel tense.  0 1 2 3 4
7. I am worried about what other people think of me.  0 1 2 3 4
8. I am thinking about how I would feel if I were told how I performed  0 1 2 3 4
9. Generally, I feel in control of things.  0 1 2 3 4
10. I am reflecting about myself.  0 1 2 3 4
11. My attention will be directed towards the task.  0 1 2 3 4
12. I am thinking deeply about myself.  0 1 2 3 4
<p>| | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>13. I feel energetic.</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>14. I am thinking about things that happened to me in the past</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>15. I am thinking about how other people might perform on this task</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>16. I am thinking about something that happened earlier today.</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>17. I expect that the task will be too difficult for me.</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>18. I will find it hard to keep my concentration on the task.</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>19. I am thinking about personal concerns and interests.</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>20. I feel confident about my performance.</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>21. I am examining my motives.</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>22. I can handle any difficulties I may encounter</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>23. I am thinking about how I have dealt with similar tasks in the past</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>24. I am reflecting on my reasons for doing the task</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>25. I am motivated to try hard at the task.</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>26. I am thinking about things important to me.</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>27. I feel uneasy.</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>28. I feel tired.</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>29. I feel that I cannot deal with the situation effectively.</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>30. I feel bored.</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
</tbody>
</table>
Appendix B

DSSQ-3 STATE QUESTIONNAIRE

POST-TASK QUESTIONNAIRE

Instructions. This questionnaire is concerned with your feelings and thoughts while you were performing the task. Please answer every question, even if you find it difficult. Answer, as honestly as you can, what is true of you. Please do not choose a reply just because it seems like the 'right thing to say'. Your answers will be kept entirely confidential. Also, be sure to answer according to how you felt WHILE PERFORMING THE TASK. Don't just put down how you usually feel. You should try and work quite quickly: there is no need to think very hard about the answers. The first answer you think of is usually the best.

For each statement, circle an answer from 0 to 4, so as to indicate how accurately it describes your feelings WHILE PERFORMING THE TASK.

Definitely false = 0, Somewhat false = 1,
Neither true nor false = 2, Somewhat true = 3, Definitely true = 4

1. I felt concerned about the impression I am making. 0 1 2 3 4
2. I felt relaxed. 0 1 2 3 4
3. The content of the task was dull. 0 1 2 3 4
4. I thought about how other people might judge my performance 0 1 2 3 4
5. I was determined to succeed on the task. 0 1 2 3 4
6. I felt tense. 0 1 2 3 4
7. I was worried about what other people think of me. 0 1 2 3 4
8. I thought about how I would felt if I were told how I performed 0 1 2 3 4
9. Generally, I felt in control of things. 0 1 2 3 4
10. I reflected about myself. 0 1 2 3 4
11. My attention was directed towards the task. 0 1 2 3 4
12. I thought deeply about myself. 0 1 2 3 4
13. I felt energetic. 0 1 2 3 4
14. I thought about things that happened to me in the past 0 1 2 3 4
15. I thought about how other people might perform on this task 0 1 2 3 4
<p>| | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>16. I thought about something that happened earlier today.</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>17. I found the task was too difficult for me.</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>18. I found it hard to keep my concentration on the task.</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>19. I thought about personal concerns and interests.</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>20. I felt confident about my performance.</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>21. I examined my motives.</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>22. I felt like I could handle any difficulties I encountered</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>23. I thought about how I have dealt with similar tasks in the past</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>24. I reflected on my reasons for doing the task</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>25. I was motivated to try hard at the task.</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>26. I thought about things important to me.</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>27. I felt uneasy.</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>28. I felt tired.</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>29. I felt that I could not deal with the situation effectively.</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>30. I felt bored.</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
</tbody>
</table>
Appendix C

NASA-TLX

We are interested in the "workload" you experienced during the experiment. Workload is difficult to define but can be seen as made up of different factors (e.g. physical or mental components). A set of six rating scales has been developed to evaluate the workload experienced by a team during different tasks. Please read the descriptions of the scales carefully. If you have a question about any of the scales, please ask the experimenter about it. It is extremely important that they be clear to you. After reading each description, please mark each scale with a horizontal line at the point which expresses your team’s experiences during the task most accurately.

<table>
<thead>
<tr>
<th>RATING SCALE DEFINITIONS</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Title</strong></td>
</tr>
<tr>
<td>MENTAL DEMAND</td>
</tr>
<tr>
<td>PHYSICAL DEMAND</td>
</tr>
<tr>
<td>TEMPORAL DEMAND</td>
</tr>
<tr>
<td>EFFORT</td>
</tr>
<tr>
<td>PERFORMANCE</td>
</tr>
<tr>
<td>------------</td>
</tr>
<tr>
<td>FRUSTRATION</td>
</tr>
</tbody>
</table>
Appendix D
3M40

How Accurately Can You Describe Yourself?

Please use this list of common human traits to describe yourself as accurately as possible.

Describe yourself as you see yourself at the present time, not as you wish to be in the future.

Describe yourself as you are generally or typically, as compared with other persons you know of the same sex and of roughly your same age.

Before each trait, please write a number indicating how accurately that trait describes you, using the following rating scale.

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
</tr>
</thead>
<tbody>
<tr>
<td>Extremely</td>
<td>Very</td>
<td>Moderately</td>
<td>Slightly</td>
<td>Neither</td>
<td>Slightly</td>
<td>Moderately</td>
<td>Very</td>
<td>Extremely</td>
</tr>
<tr>
<td>Inaccurate</td>
<td>Inaccurate</td>
<td>Inaccurate</td>
<td>Inaccurate</td>
<td>Accurate</td>
<td>Accurate</td>
<td>Accurate</td>
<td>Accurate</td>
<td>Accurate</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Bashful</th>
<th>Energetic</th>
<th>Moody</th>
<th>Systematic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bold</td>
<td>Envious</td>
<td>Organized</td>
<td>Talkative</td>
</tr>
<tr>
<td>Careless</td>
<td>Extraverted</td>
<td>Philosophical</td>
<td>Temperamental</td>
</tr>
<tr>
<td>Cold</td>
<td>Fretful</td>
<td>Practical</td>
<td>Touchy</td>
</tr>
<tr>
<td>Complex</td>
<td>Harsh</td>
<td>Quiet</td>
<td>Uncreative</td>
</tr>
<tr>
<td>Cooperative</td>
<td>Imaginative</td>
<td>Relaxed</td>
<td>Unenvious</td>
</tr>
<tr>
<td>Creative</td>
<td>Inefficient</td>
<td>Rude</td>
<td>Unintellectual</td>
</tr>
<tr>
<td>Deep</td>
<td>Intellectual</td>
<td>Shy</td>
<td>Unsympathetic</td>
</tr>
<tr>
<td>Disorganized</td>
<td>Jealous</td>
<td>Sloppy</td>
<td>Warm</td>
</tr>
<tr>
<td>Efficient</td>
<td>Kind</td>
<td>Sympathetic</td>
<td>Withdrawn</td>
</tr>
</tbody>
</table>
Appendix E

![Mental workload graph]

**Task Phase**
Mental workload for X, D, and Y task phases. Error bars are standard errors.

![Physical workload graph]

**Task Phase**
Physical workload for X, D, and Y task phases. Error bars are standard errors.

![Temporal workload graph]

**Task Phase**
Temporal workload for X, D, and Y task phases. Error bars are standard errors.
Task Phase
Performance workload for X, D, and Y task phases. Error bars are standard errors.

Task Phase
Effort workload for X, D, and Y task phases. Error bars are standard errors.

Task Phase
Frustration workload for X, D, and Y task phases. Error bars are standard errors.