Australian Government
Department of Defence
Defence Science and
Technology Organisation

# Exercise Black Skies 2008: Enhancing Live Training Through Virtual Preparation

# Part Two: An Evaluation of Tools and Techniques

*Eleanore Tracey, Sam Hasenbosch, Julian Vince, Daniel Pope, Aaron Stott, Christopher Best, Christopher Shanahan and Melanie Finch*

**Air Operations Division**
Defence Science and Technology Organisation

DSTO-TR-2305

## ABSTRACT

Exercise Black Skies 2008 (EBS08) was a simulation exercise conducted in the weeks prior to the live air combat training exercise, Pitch Black 2008 (PB08). During EBS08, a Royal Australian Air Force (RAAF) air battle management (ABM) team conducted a series of vignettes designed to prepare them for their tactical command and control role in PB08. A broad goal of EBS08 was to provide an environment within which a number of new simulation tools and training techniques could be evaluated and developed for future implementation within the Royal Australian Air Force (RAAF). The tools and techniques evaluated were: the Air Defence Ground Environment Simulator (ADGESIM), Toteboard, Air Warfare Assessment and Review tool, After Action Review tools and the Team-Dimensional Training framework. These evaluations, which form the basis or this report, were informed by the observations of human factors scientists and interviews with the RAAF participants. Participants provided extensive feedback on each tool, highlighting shortcomings and providing valuable suggestions for improvement. This feedback will guide the further development of these tools for implementation in future research exercises and in RAAF training programs. Overall, the participants found EBS08 extremely valuable and were satisfied with the quality of training they received. The simulation tools and training techniques assisted in promoting the quality of this experience and prompted participants to consider methods by which the current RAAF approach to training could be improved.

## RELEASE LIMITATION

*Published by*

*Air Operations Division*
*DSTO Defence Science and Technology Organisation*
*506 Lorimer St*
*Fishermans Bend, Victoria 3207 Australia*

*Telephone: (03) 9626 7000*
*Fax: (03) 9626 7999*


*AR-014-564*
*June 2009*

# Exercise Black Skies 2008: Enhancing Live Training Through Virtual Preparation

# Part Two: An Evaluation of Tools and Techniques

## Executive Summary

Exercise Black Skies 2008 (EBS08) was a simulation exercise conducted under the Defence Science and Technology Organisation's (DSTO's) Support to Air Force Readiness and Renewal task (AIR 07/232) and Project Arrangement PA-AF-0025 "Distributed Mission Training Effectiveness Research" with Air Force Research Laboratories in the United States. EBS08 was envisaged as a preparatory exercise for a Royal Australian Air Force (RAAF) air battle management (ABM) team who were to be involved in Exercise Pitch Black 2008 (PB08) and was conducted in the weeks prior to the commencement of PB08. Pitch Black is a biennial, live air combat training exercise, hosted by the RAAF and involving participants from a number of foreign military forces. EBS08 was designed to resemble PB08 as closely as possible in terms of scenario vignettes, orders of battle, airspace, procedures and tactics. The role of ABM teams in both exercises was to provide command and control of the Red Force air assets. At EBS08, a White Force, including RAAF simulator operators (SIMOPs) and a White Force mission director, were tasked with generating and controlling the synthetic Red and Blue Forces in a manner consistent with the scale and complexity of the PB08 scenario.

This report is the second in a two-part publication series addressing the research outcomes of EBS08. EBS08 had two broad goals, the first of which is discussed in Part One of the series (Shanahan, et al., in preparation) and the second of which will be discussed in the present report. The first goal of the exercise was to examine the benefits of providing the RAAF ABM team with a synthetic training environment within which to prepare for their tactical command and control role in PB08. Shanahan, et al. therefore discusses the effectiveness of EBS08 as a preparatory exercise. The second goal was to provide an environment within which a number of new simulation tools, decision aids, performance measurement tools and training techniques could be evaluated and developed for future implementation within the RAAF. The tools and techniques trialled in EBS08 were the Air Defence Ground Environment Simulator (*ADGESIM*) suite of simulation tools, the Toteboard, the Air Warfare Assessment and Review (*AWAR*) tool, a number of After Action Review (AAR) tools and the Team-Dimensional Training (*TDT*) framework. The present report provides an evaluation of these tools and techniques in terms of their potential for use in a RAAF training context and suggests methods by which they may be improved. This evaluation was based on information gleaned from structured discussions with the ABM team and SIMOPs and the general observations of DSTO human factors scientists.

The *ADGESIM* suite of tools consists of the *DISVOX* radio simulator, the Pilot Simulator Interface (*PSI*), the *SensorLINK* simulated sensor gateway, the *SimMan* simulator manager, the Tactical After Action Review for Distributed Interactive Simulation (*TAARDIS*) and the *WorldVIEW* three-dimensional visualisation tool. These tools are currently in use at the Surveillance and Control Training Unit (SACTU) in Williamtown, New South Wales, and the School of Air Warfare (SAW) in Sale, Victoria. Prior to EBS08, the *ADGESIM* tools were modified to reduce SIMOP workload in large scenarios and to improve the consistency of behaviour representation. The version of *ADGESIM* trialled at EBS08 was considered a significant improvement over the version in use at SACTU and SAW. The RAAF participants were impressed with the potential of the new system to run effectively with fewer operators than the previous version.

The *Toteboard* is a customised electronic spreadsheet that was developed at SACTU to provide ABM teams with situation awareness on mission-specific aspects, including aircraft fuel states, alert states, weapon loadouts, assigned airspace and radio frequency, tanker offload levels, endurance estimates, and aircraft turn-around timings. The ABM team felt that the *Toteboard* provided all information required during mission execution. It also provided the additional benefit of facilitating coordination between the tactical control centre and the White Force control centre.

The goal of *TDT* is to improve team performance in four areas, namely information exchange, communication, supporting behaviour and initiative/leadership. The RAAF participants reported that these teamwork constructs are implicit in current RAAF training and performance assessment, and that addressing them explicitly during debrief is unlikely to provide additional training benefit. During EBS08, *TDT* was only partially implemented and therefore conclusions cannot be drawn regarding the utility of the approach. Had *TDT* been implemented in its entirety however, the reluctance of RAAF participants to accept and implement the approach may have impacted negatively on its effectiveness. This highlights the importance of gaining the RAAF participants’ acceptance of any teamwork tool incorporated in future exercises, either through a thorough explanation of the benefits of the chosen approach or by modifying an approach to bring it into line with the culture of the RAAF.

The *AWAR* tool, created at DSTO, was designed to promote the consistency and objectivity of assessment ratings and to provide an assessor with focus areas around which to structure debriefs. The *AWAR* tool, incorporating the ABM Defensive Counter Air team goal hierarchy offered a marked improvement on current RAAF methodology and tools, in terms of the breadth, depth and consistency of ratings. Participants felt that the team goal hierarchy would be useful for designing basic course scenarios at SACTU and SAW. The ABM team found the visual layout of the debrief user interface intuitive. The scores and comments relating to each sub-goal were extremely useful, as they provided the team with an understanding of their current performance, showed how much room existed for improvement, and provided instructions for improving.

The AAR tools were to be used by the assessor to visually and aurally support the debriefing comments he made during the AARs. The RAAF participants found the AAR tools very effective in minimising the amount of preparation required between the end of a mission and the beginning of the task performance debrief. The task assessor also felt that the AAR tools improved the time efficiency of the debrief process itself by facilitating the brevity and sharpness of learning points. The playback of video and audio was considered an extremely useful method of illustrating important learning points.

Overall, the participants found EBS08 to be an extremely valuable exercise and were satisfied with the quality of the preparation they received for PB08. The simulation tools and training techniques evaluated in the exercise assisted in promoting the quality of this experience and prompted participants to consider methods by which the current approach of the RAAF to training could be improved. Participants provided extensive feedback on each tool, highlighting shortcomings and providing valuable suggestions for improvement. This feedback will guide the further development of these tools for implementation in future research exercises and RAAF training programs.

# Authors


**Ms. Eleanore Tracey**
Air Operations Division


*Eleanore Tracey is a human factors specialist in the Air Operations Division of the Defence Science and Technology Organisation (DSTO). She joined DSTO after completing an honours degree in psychology at Deakin University. Her work at DSTO has covered a range of issues relating to human performance modelling, team training and human-machine interface design for helicopters, fast jets and surveillance aircraft. This research has involved computer simulation modelling, small scale lab-based experimentation, high fidelity simulation and observational data collection during live military exercises.*

---


**Mr. Sam Hasenbosch**
Air Operations Division


*Sam Hasenbosch was a Research Engineer at the DSTO, Air Operations Division in Melbourne. His major research interest was synthetic training systems for Air Battle Management. Mr Hasenbosch is a former RAAF Air Combat Officer and earned a Bachelor of Science in Computer Science and Management from the University of New South Wales. He is now employed as a Senior Research Engineer at National ICT Australia, Queensland Research Laboratory. His current research interest is modelling Air Traffic Controllers to assess the human factors associated with future Air Traffic Management concepts.*

---


**Mr. Julian Vince**
Air Operations Division


*Julian Vince is a psychologist working in the Air Operations Division of DSTO as a human factors specialist. He has a Master's degree in Industrial/ Organisational psychology from Deakin University. Julian's work at DSTO has centred on the cognitive aspects of human-machine interface, group process, training and the development of expertise. This work has had a fast-jet and surveillance aircraft focus.*

---

## Mr. Daniel Pope

YTEC Pty Ltd

*Daniel Pope is a software developer contracted to the Air Operations Division of the Defence Science and Technology Organisation (DST0). He holds a B.InfTech (hons) specialising in artificial intelligence and computer science, and has worked for the last five years in distributed simulation systems research and development for Air Battle Management training. His previous roles include Air Traffic Management systems, and lecturing A.I. at Griffith University. His research interests include philosophy of mind with respect to the computational metaphor, and the use of agent based systems for training and simulation.*

---

## FLTLT Aaron Stott

Surveillance and Control Training Unit, RAAF

*FLTLT Aaron Stott is an Air Battle Manager (ABM) instructor at Surveillance and Control Training Unit, RAAF Williamtown. Before joining DSTO, he completed a Computer Science degree at Australian Defence Force Academy (ADFA), and gained over six years experience as an ABM. His role at DSTO was to act as an Air Combat Officer (ACO) Liaison Officer. In this role he provided specialist input into ABM simulation development and training and small scale laboratory exercises.*

---

## Dr. Christopher Best

Air Operations Division

*Christopher Best is a research scientist within Air Operations Division's Crew Environments and Training branch (CET). He holds a B.A.(Hons) in Psychology (awarded in 1997) and a PhD in Psychology (awarded in 2001). His research and professional interests include human perception and cognition, human-machine interaction, team performance and team training. Dr Best is Science Team Leader for the Human Sciences component of AIR 07/232 (Support to Air Force Readiness and Renewal), Australian leader on two project arrangements with USAF Research Laboratories, an Australian representative on TTCP HUM Technical Panel 2 (Training Technologies), and an Honorary Fellow of the School of Psychology, Deakin University.*

---

**Dr. Christopher Shanahan**
Air Operations Division

*Dr. Christopher Shanahan joined the Crew Environments and Training (CET) branch of Air Operations Division in 2006. He is a forensic psychologist by training, and prior to joining DSTO he worked as a clinician in the NSW Department of Corrective Services' Violent Offender Treatment Program, and served as Clinical Coordinator of the postgraduate program in forensic Psychology at the University of New South Wales (where he obtained his PhD). His research and practice interests include emotion and cognition, psychological assessment techniques, and the science of training.*

---

**Ms. Melanie Finch**
Air Operations Division

*Melanie Finch works in Crew Environments and Training within the Air Operations Division. Melanie has a B.Sc. (Hons) in psychology (awarded in 2005) and joined DSTO in 2006. Melanie's interests include human perception, team training and measurement, and cognitive biases in decision making.*

---

# Contents

# Glossary

| | |
|---|---|
| **3CRU** | 3 Controlling and Reporting Unit |
| **41WG** | 41 Wing |
| **AAR** | After Action Review |
| **ABD** | Air Battle Director |
| **ABM** | Air Battle Management / Manager |
| **ACG** | Air Combat Group |
| *ADGESIM* | Air Defence Ground Environment Simulator |
| **ADF** | Australian Defence Force |
| **AIR 07/232** | DSTO's Support to Air Force Readiness and Renewal Task |
| **AS JET** | Australian Joint Essential Task |
| *AWAR* | Air Warfare Assessment and Review |
| **C** | Competent |
| **C2** | Command and Control |
| **DCA** | Defensive Counter Air |
| **DEAD** | Destruction of Enemy Air Defences |
| **DIS** | Distributed Interactive Simulation |
| **DSTO** | Defence Science and Technology Organisation |
| **EBS08** | Exercise Black Skies 2008 |
| **ECADEX** | East Coast Air Defence Exercise |
| **MSCT** | Multiple Source Correlator Tracker |
| **OAS** | Offensive Air Support |
| **OCA** | Offensive Counter Air |
| **ORBAT** | Orders of Battle |
| **IFF** | Identity Friend or Foe |
| **JPR** | Joint Personnel Recovery |
| **NYC** | Not Yet Competent |
| **PB08** | Pitch Black 2008 |
| **PC** | Personal Computer |
| **PDU** | Protocol Data Units |
| **PSI** | Pilot Simulation Interface |
| **RAAF** | Royal Australian Air Force |
| **ROE** | Rules Of Engagement |
| **RSC** | Red SIMOP Coordinator |
| **SA** | Situation Awareness |
| **SACTU** | Surveillance and Control Training Unit |
| **SAW** | School of Aviation Warfare |
| **SIMOP** | Simulator Operator |
| **SME** | Subject Matter Expert |
| **SRG** | Surveillance and Response Group |
| *TAARDIS* | Tactical After Action Review for Distributed Interactive Simulation |
| *TDT* | Team Dimensional Training |
| **US** | United States |
| *VBA* | Visual Basic for Applications |
| **VNC** | Video Network Client |
| **WD** | Weapons Director |
| **WFMD** | White Force Mission Director |

# 1. Introduction

Exercise Black Skies 2008 (EBS08) was a simulation exercise conducted under the Defence Science and Technology Organisation's (DSTO) Support to Air Force Readiness and Renewal task (AIR 07/232) and Project Arrangement PA-AF-0025 "Distributed Mission Training Effectiveness Research" with Air Force Research Laboratories in the United States. EBS08 was envisaged as a preparatory exercise for a Royal Australian Air Force (RAAF) air battle management (ABM) team who were to be involved in Exercise Pitch Black 2008 (PB08) and was conducted in the weeks prior to the commencement of PB08. Pitch Black is a biennial, live air combat training exercise, hosted by the RAAF and involving participants from a number of foreign military forces. EBS08 was designed to resemble PB08 as closely as possible in terms of scenario vignettes, orders of battle (ORBAT), airspace, procedures and tactics. EBS08 can be classified as an 'exercise workup', as defined in the continuum of simulated training events in Appendix A. The role of ABM teams in both exercises was to provide command and control of the Red Force air assets. At EBS08, a White Force, including RAAF simulator operators (SIMOPs) and a White Force mission director (WFMD) on contract from Milskil Integrated Defence Solutions (Cooks Hill, NSW), were tasked with generating and controlling the synthetic Red and Blue Forces in a manner consistent with the scale and complexity of the PB08 scenarios. The WFMD had been involved in the development of the PB08 scenario and was to play a key White Force role in PB08.

EBS08 had two broad goals. These goals are discussed along with relevant research outcomes in a two-part publication series. The first goal, discussed in Shanahan, et al. (in preparation), was to examine the benefits of providing a RAAF ABM team from 41 Wing (41WG) Surveillance and Response Group (SRG) with a synthetic training environment within which to prepare for PB08. The second goal, discussed in the present report, was to provide an environment within which a number of new simulation tools, decision aids, performance measurement tools and training techniques could be evaluated in terms of their utility within RAAF training programs and the methods by which they could be improved. The tools and techniques evaluated in the exercise, and thus discussed in this report, are the Air Defence Ground Environment Simulator (*ADGESIM*) suite of simulation tools, the *Toteboard*, the Air Warfare Assessment and Review (*AWAR*) tool, a number of After Action Review (AAR) tools and the Team-Dimensional Training (*TDT*) framework. The environment produced through the use of these tools and techniques represented an evolved approach to training over methods currently implemented by the RAAF.

The body of this report is composed of two sections. The first section considers aspects relating to the planning and conduct of EBS08. It provides a description of the personnel involved in the exercise and the roles they played, and of scenario management and exercise workflow. The exercise facility, infrastructure and tools are then described in detail, outlining the methods by which they were implemented during EBS08 and the potential improvements they offer over approaches currently employed by the RAAF. The second section discusses the utility of each tool or method, potential modifications and potential applications in RAAF training programs. These discussions are based on information gleaned from interviews with the RAAF participants during the exercise and the observations of DSTO human factors scientists. The discussion of each tool concludes with a list of recommendations. A consolidated list of these recommendations is available in Appendix B.

# 2. Methods and tools

## 2.1 Exercise planning and execution

The following subsections will provide an overview of the scenarios employed at EBS08, the participant roles and responsibilities the roles of exercise staff. The schedule for the week is outlined, along with the system malfunctions that occurred and their effect on the exercise schedule. The procedure is then described, including data collection and debriefing procedures.

### 2.1.1 EBS08 Scenarios

A major goal of EBS08 was to prepare a 41WG ABM team for their Red Force tactical command and control (C2) role in PB08. The scenario vignettes used in EBS08 were identical to those proposed for PB08, in an effort to promote the transfer of training benefits from the simulated exercise to the live exercise. The exercise scenario developed for PB08, and thus used also in EBS08, was designed to provide the Blue Force military participants, particularly those from Air Combat Group (ACG), with a broad range of training experiences. The development and execution management of the specific vignettes (i.e., portions of the exercise scenario) was contracted to Milskil Integrated Defence Solutions, a specialist Defence training company.

Seven vignettes were developed for PB08, four of which were used in EBS08. The vignettes varied in terms of the training objective, ORBAT, rules of engagement (ROE), airspace, threats and targets. The collective nature and sequence of the vignettes embodied an escalation in tension between the opposing forces, and the progression of military actions over the three weeks of PB08. The scenario required Blue Force participants to conduct a variety of missions, including offensive counter air (OCA), offensive air support (OAS), deep strike, destruction of enemy air defences (DEAD) and joint personnel recovery (JPR). The Red Force participants responded by employing defensive counter air (DCA) tactics. Two 41WG ABM teams worked in shifts to provide tactical C2 support to the live multinational Red Force that participated in PB08. One of these ABM teams also participated in EBS08, commanding a synthetic Red Force controlled by RAAF SIMOPs using the *ADGESIM* suite of tools.

The vignettes chosen for days one to four of EBS08 respectively were Vignette A (a Blue Force OCA mission), Vignette B (involving Blue Force OCA and strike), Vignette D (a Blue Force JPR mission) and Vignette E (a Blue Force deep strike mission). Vignette B was repeated on day five of the exercise for the purpose of comparing ABM team performance at the start and end of the activity. The Milskil WFMD had been involved in the development of the PB08 vignettes and was to play a key White Force role in PB08. This promoted the consistency of scenario implementation between EBS08 and PB08. The Milskil member had previously been employed by the RAAF as a fighter pilot and was able to provide expert guidance on the Blue Force tactics expected in PB08.

### 2.1.2 Participants

The EBS08 ABM team consisted of five RAAF participants from 41WG SRG: an Air Battle Director (ABD), a Weapons Director (WD)[1] and three ABMs in direct control roles. A WD assistant and additional ABM are sometimes required when workload is high, however, the RAAF was unable to provide these due to manning constraints. Although the roles played by ABMs varied to suit each vignette, team membership was consistent throughout EBS08, allowing team members to develop an understanding of how their team mates worked and how best to support them. The time between exercises was minimised to promote the transfer of training benefits from EBS08 to PB08. EBS08 was completed two weeks prior to the commencement of PB08, allowing time for the conduct of other important preparation activities.

In virtual exercises of this kind, the experience levels of White Force role players and the ratio of role players to active entities are significant determinants of training fidelity (Best, et al., 2007). During EBS08, the White Force consisted of five RAAF SIMOPs from 41WG SRG, two assigned to Blue roles and three assigned to Red roles, a Red SIMOP Coordinator (RSC), played by a DSTO staff member with extensive ABM experience, and a WFMD, played by a Milskil staff member with operational experience as a RAAF fighter pilot and who had been involved in the development of the PB08 vignettes. The EBS08 scenario was complex, with SIMOPs controlling between 40 and 60 Red and Blue aircraft. Three of the SIMOPs had between two and five years experience and two SIMOPs were not yet qualified and were not experienced in the control of large forces. Although fewer SIMOPs were involved than was optimal, no more personnel, either trained or in training, were available. As a result, the workload of SIMOPs was expected to be high.

### 2.1.3 Exercise staff and preparation

Human factors scientists designed data collection protocols and coordinated the planning of EBS08. Scenario design and participant liaison was supported by a RAAF ABM instructor who was posted to DSTO in the months preceding EBS08. The physical components of the synthetic training environment were constructed by engineers and technicians engaged in human-in-the-loop simulation research. The software components were configured and supported by YTEK Pty. Ltd. (Fishermans Bend, Vic) engineers, contracted by DSTO for the development of *ADGESIM*. All exercise staff worked within the Crew Environments and Training branch, Air Operations Division, based at DSTO, Melbourne.

RAAF participation in the exercise was confirmed three weeks prior to the exercise, and the assembly of the physical environment completed two weeks prior. The remaining two weeks were spent configuring and testing the information technology infrastructure, testing recent modifications and software builds, and developing fixes as problems arose.

---

[1] At the time the exercise was conducted, the title of this position was Weapons Director. This title will therefore be used throughout this report. Importantly, the title of the position has since been changed to Tactical Director.

### 2.1.4 Exercise schedule

The planned schedule for EBS08 is depicted in Figure 1. At the beginning of the week, participants were briefed on the goals and purposes of the exercise, the exercise scenario, the structure of the synthetic training environment and the schedule of events. The ABM team was briefed on how their task performance and team processes would be measured and discussed after each vignette. An initial measurement session was conducted to record participant demographics, experience levels and their expectations of the exercise. All participants were afforded time to familiarise themselves with specific systems. Each subsequent day of the exercise consisted of the following key events in order of execution:

1. A scenario update, including information on expected threats, ORBAT, ROE, airspace, and enemy intent related to the forthcoming vignette.
2. Preparation time, in which the ABM team could revise their plans and the White Force team could discuss the behaviours to be produced.
3. Mission execution.
4. A measurement session, requiring ABM team members and SIMOPs to complete a short workload questionnaire. In addition, ABM team members completed questionnaires addressing team efficacy, team cohesion and team processes.
5. An after action review (AAR) session, to discuss ABM team task and teamwork performance, supported by scenario replay.
6. An exercise feedback session, where all RAAF participants took part in discussions relating to the specific tools and techniques under evaluation at EBS08, and provided exercise general feedback. On Day 4, data was also collected to determine the extent to which participant's expectations of the exercise were met.

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | 8:00 | 9:00 | 10:00 | 11:00 | 12:00 | 1:00 | 2:00 | 3:00 | 4:00 | 5:00 | |
| Monday, May 19, '08 | Introduction Brief | Scenario Brief | Measurement Brief | Measurement Session | *TDT* Exercise | System Famil. | Lunch | Prep Time | Mission 1 (*Vignette A*) | Measurement Session | AAR |
| Tuesday, May 20, '08 | Day 2 Update | Prep Time | Mission 2 (*Vignette B*) | Measurement Session | Lunch | AAR | Exercise Feedback (AAR, *AWAR*, *Toteboard* & General) | | | | |
| Wednesday, May 21, '08 | Day 3 Update | Prep Time | Mission 3 (*Vignette D*) | Measurement Session | Lunch | AAR | Exercise feedback (*ADGESIM*, *TDT* & General) | | | | |
| Thursday, May 22, '08 | Day 4 Update | Prep Time | Mission 4 (*Vignette E*) | Measurement Session | Lunch | AAR | Measurement & General Feedback | | | | |
| Friday, May 23, '08 | Day 5 Update | Prep Time | Mission 5 (*Vignette B*) | Measurement Session | AAR | | | | | | |

*Figure 1: The schedule of events for days 1 to 5 of EBS08*

#### 2.1.4.1 Unscheduled events

Mission 1 was disrupted and Mission 2 was partially degraded due to instability in the simulation software. In order to reduce the risk to the remaining vignettes and allow further software testing, an improvised 'air warfare game' was conducted in place of the AAR session on Day 2. This game required each ABM to pair with a SIMOP and control two fighter pairs. The aim of the game was to be the first formation to fly through the centre of an arbitrary hexagonal airspace and back to their starting corner, without being destroyed by an opposing formation. The participants reported that this session was a useful alternative, as it allowed the ABMs to practice basic fighter controls skills, and it granted the SIMOPs additional practice with the human-machine interface of the latest simulation tools. The software proved to be stable throughout the game and the remaining vignettes.

### 2.1.5 Procedure

DSTO human factors scientists were responsible for managing: adherence to the exercise schedule, the delivery of briefings and associated materials, and access to the facility by staff, participants and visitors. The following paragraphs describe other exercise roles performed by participants and the exercise management team.

#### 2.1.5.1 Participant roles and responsibilities

The ABM team was required to command Red Force air assets to defend airspace and key points against the larger and technically superior Blue Force. Although the ABM team was not expected to prevent the loss of all key points or assets, the effectiveness of their defence was assessed by the RAAF ABM instructor (from a task work perspective) and the ABD (from a teamwork perspective). Both assessors monitored the same situation display and communication channels available to the ABM team, and utilised the assessment tools described later in this report. In an effort to promote the consistency of measurement across EBS08 and PB08, the same ABM instructor was tasked with assessing ABM team task performance at the two exercises.

The Red SIMOPs were responsible for manipulating the synthetic fighter assets under the control of the ABM team, and emulating the communications of those aircraft pilots. Due to the large pool of simulated aircraft to be controlled (generally between four and eight aircraft per SIMOP), a DSTO staff member, formerly employed as a RAAF ABM, fulfilled the role of RSC. This involved 'scrambling' the fighter formations requested by the ABM team, and handing them over to the SIMOP whose apparent workload was lowest at that time. Similarly, the RSC would take control of aircraft which were 'kill-removed', returning to base or requiring air-to-air refuelling. This permitted SIMOPs to focus their efforts on mission phases requiring their domain knowledge and experience, rather than benign phases.

The Blue SIMOPs were required to manipulate the Blue Force air assets under the direction of the WFMD. Due to the large number of assets (generally between 12 and 20 aircraft per SIMOP), the SIMOPs were required to manipulate formations as chained groups, that is, with the lead group being used to control all trailing groups. In terms of workload, the management of large numbers of aircraft was only possible because these SIMOPs were not required to communicate with external parties. The WFMD directed the Blue SIMOPs to arrange different aircraft types (with associated roles) into a certain order and position them at a specific location and time. In this way the WFMD was able to create a scenario representative of the Blue Force tactics expected in PB08.

#### 2.1.5.2 AAR session

After each vignette, the ABM instructor and ABD debriefed the ABM team on their task work and teamwork performance respectively, facilitating discussions of key strengths and weaknesses. The task assessor replayed relevant portions of the mission to support ratings and comments against performance measures. These replays represented scenario ground truth, that is, activity in the flight environment as it actually occurred, and user-specific views, that is, activity in the flight environment as it appeared on a specific user's screen. Ground truth differed from user-specific views as a result of quality and range of the simulated radar. Participant communications were also recorded during execution and could be replayed in synchrony with video replays. The ABD then discussed the quality of team processes. At all stages the evaluations were open to comment and discussion by all ABM team members. Please refer to Subsection 2.3.5 for a more detailed discussion of the approach employed and tools utilised during the AAR.

#### 2.1.5.3 Data collection

Human factors scientists collected data regarding the suitability of the training environment and the value of the exercise. Qualitative and quantitative data was entered by participants into electronic forms at the beginning of the exercise and after every mission. This information included participant demographics, qualifications and experience, exercise expectations, post-exercise evaluations of whether expectations were met, workload, team efficacy and team cohesion. This data, along with data obtained from the teamwork and task work assessments, will contribute to discussions on the effectiveness of EBS08 in preparing the ABM team for PB08 (Shanahan, et al., in preparation).

Data was also collected to assist in the evaluations of a number of tools and techniques, namely the *ADGESIM* suite of tools, the *Toteboard*, the *AWAR* tool, the *TDT* framework and the AAR tools. These tools and techniques will be described in detail in Subsection 2.3. Qualitative data was obtained through semi-structured interviews with RAAF personnel regarding the suitability of the tools and techniques for use in a RAAF training environment. Information obtained through these interviews forms the basis for discussion in Section 3 of this report.

## 2.2 Exercise infrastructure

The following subsections provide a description of the layout of the exercise facility and the components of the information technology infrastructure.

### 2.2.1 General facility layout

The general layout of the simulation facility is depicted in Figure 2. As shown, the facility was split into four areas: the server room, the White Force control centre, the tactical control centre and the observation room. The server room was separated from the other components by solid walls, while the other areas were separated by two metre high partitions. The server room contained the *ADGESIM* simulation and file servers. The SIMOPs and WFMD were situated in the White Force control centre, which was arranged by Force. The ABMs, WD, ABD and RAAF instructor were situated in the tactical control centre. The WD was positioned such that he could monitor the three subordinate ABMs. The observation room was constructed to allow the exercise management team and visitors to view activity inside the tactical control centre unobtrusively via live feed projections and a plasma display.

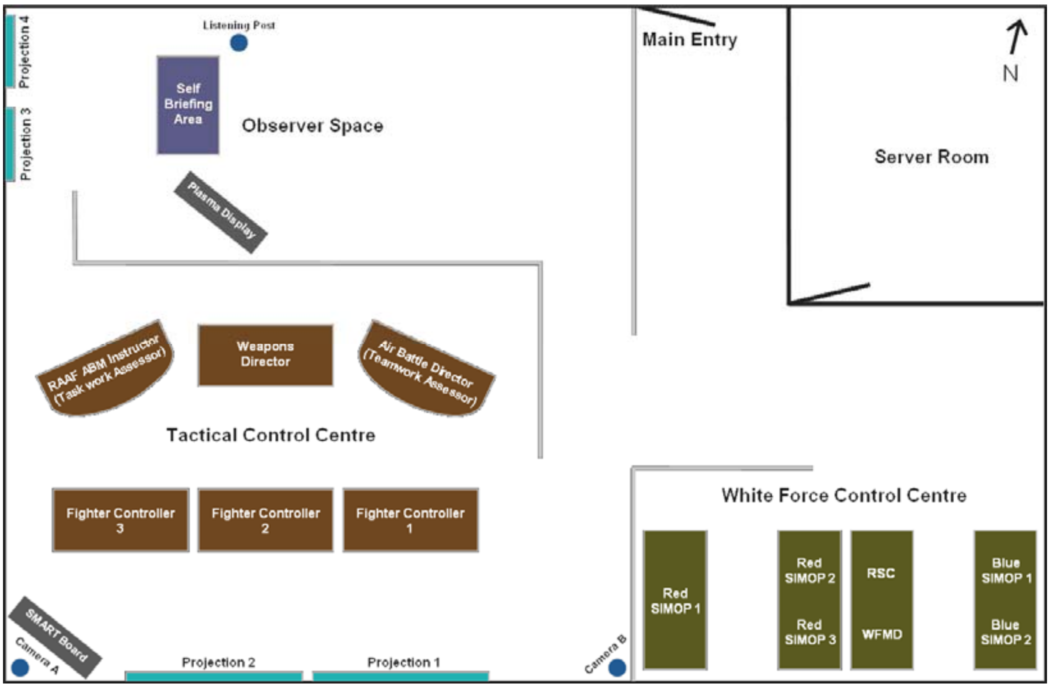



*Figure 2: A top-down representation of the simulation facility, showing the control cells and positioning of personnel*

### 2.2.2 Information technology infrastructure

#### 2.2.2.1 Common systems

All exercise participants were provided with two separate personal computers (PCs) on a common network: one as the communications console, the other as the main workstation. The communications console consisted of touch screen monitors, push-to-talk pedal and a headset. An *ADGESIM* application called *DISVOX* was operated on this console. This provided basic radio and intercom simulation and a graphical user interface similar to the operational systems used by the RAAF. A single communications configuration was developed which included a unique operator role for each physical position. The software utilises the Distributed Interactive Simulation (DIS) protocol which allows participant communications to be logged with any DIS logging system. The main workstation consisted dual monitors (each with a resolution of 1280x1024). The software being operated depended on the participant role.

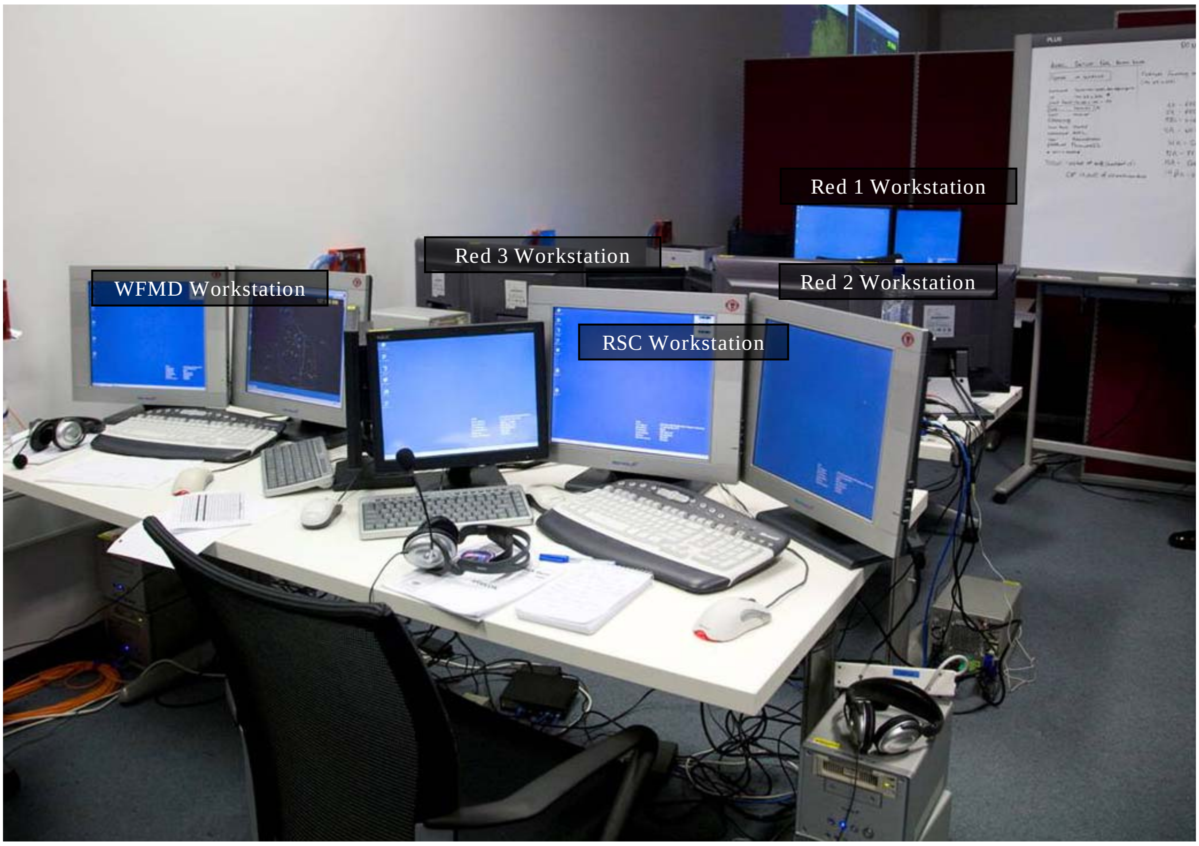



*Figure 3: The White Force control centre, showing the positioning of white force personnel and the layout of workspaces*

#### 2.2.2.2 Scenario generation systems

The White Force control centre is represented in Figure 3. All members of the White Force were provided with the Pilot Simulation Interface (*PSI*) software, which allows entities to be created, controlled and transferred between team members. Please refer to Subsection 2.3.1 for a more detailed description of *PSI*. Each instance of *PSI* was connected to a server operating the VR-Forces software, which models all of the aircraft in the simulation environment and broadcasts this via the DIS protocol. The WFMD also used a free video network client (VNC), called UltraVNC, to monitor the situation display used by the WD. This provided the WFMD awareness of when Blue forces were detected and what rules of engagement might be applied, at any point in time. The RSC also used a shared instance of the *Toteboard* application to determine when to create entities that had been scrambled by the ABM team.

#### 2.2.2.3 Command and control systems

All members of the ABM team and the RAAF instructor assessing task work performance were provided with the Solipsys tactical situation display software. The software was configured with airspace and waypoint data, and a RAAF-developed user profile which defines aspects such as data block fields, identify friend or foe (IFF) decoding and a multitude of appearance settings. Each instance of Solipsys was connected to the multiple source correlator / tracker (MSCT), which in turn was stimulated by an *ADGESIM* server application called *SensorLink*. *SensorLink* was used to model the PB08 radar types and sites, and to translate the detection of entities received via the DIS protocol to a plot protocol compatible with the MSCT. This accurately replicates the operational C2 system used by the RAAF, with the exception that *SensorLink* replaces the real sensor network. Each member was also provided with a shared instance of the *Toteboard* application. This was used to collect and disseminate status information concerning aircraft assigned to team members. During mission execution, projections 1 and 2 in Figure 4 replicated the Solipsys and *Toteboard* displays of the WD.


*Figure 4: The tactical control centre, showing the positioning of members of the ABM team, the layout of operator workstations and the visual tools used during AAR*

#### 2.2.2.4 AAR infrastructure

The AAR infrastructure consisted of two parts: the collection system, which was used during mission execution, and the review system, which was used during the AAR. Both systems utilised components of the *ADGESIM* Tactical After Action Review for Distributed Interactive Simulation (*TAARDIS*) suite. The collection system was composed of a DIS logging and bookmarking application that operated on the task assessor's workstation, and individual screen capture applications that operated on each ABM team member workstation. The task assessor was also provided with the *AWAR* tool for prompted performance measurement and the *Toteboard* application for situation awareness (SA) during mission execution. Due to the number and nature of these applications and the requirement for concurrent access, the task assessor's workstation was fitted with four monitors. The review system included a rear projection SMART board (see Figure 4) to display the ratings and comments made by the task assessor using the *AWAR* tool. The task assessor's comments were supported by wall projections and audio replays of relevant mission segments that were controlled by an AAR operator from the task assessor's workstation.

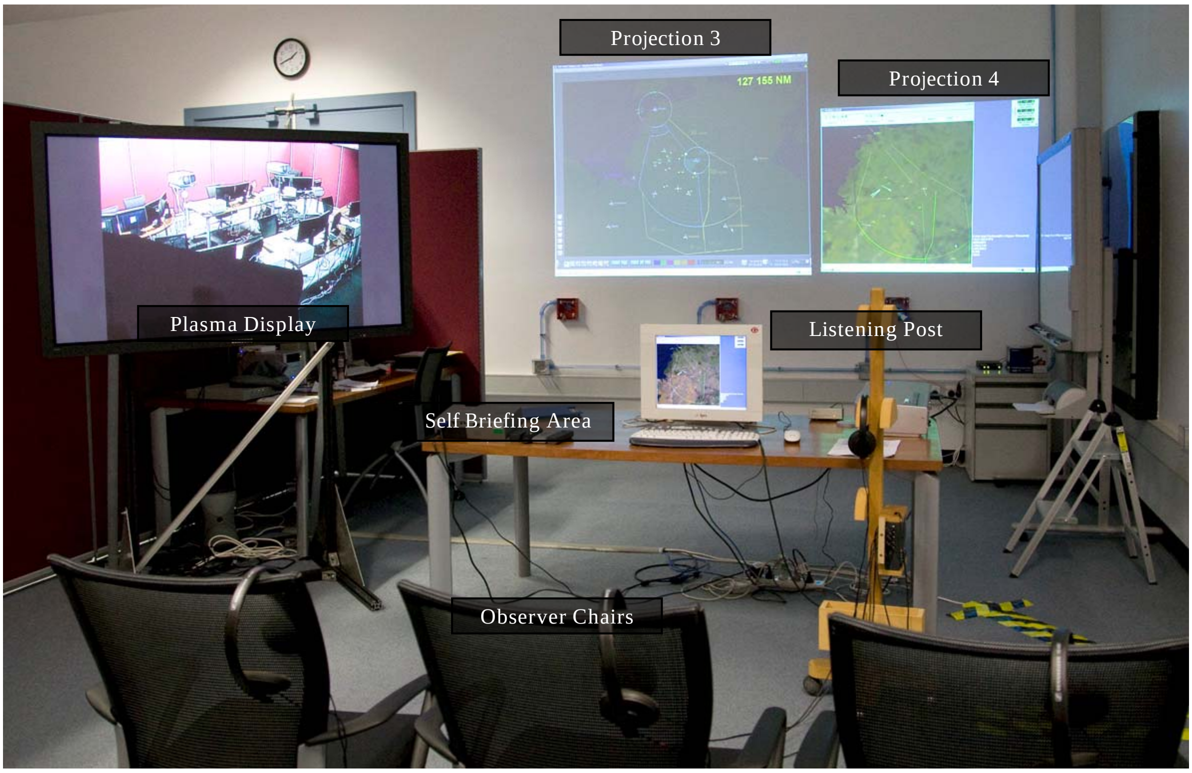



*Figure 5: The observation room, showing the tools available to observers to aid in their understanding of participant activity*

### 2.2.3 Observation room

The observation room was constructed to provide the exercise management team and visitors with an area from which they could view the exercise without disturbing the RAAF participants. Figure 5 shows the arrangement of the observation room. Observers were provided with a number of tools to assist in their understanding of mission events and team interactions during the missions and AAR. As observers were most interested in the activity within the ABM team, a real-time image of the team was displayed on the plasma screen located within the observation room. The input to this display was from one of two cameras, Camera A and Camera B, located within the tactical control centre (refer to Figure 2). In Figure 5, the input to the plasma display is from Camera A. Input to the display was modified to provide the best view of activity within the tactical control centre. In general, Camera A provided the best view during missions and Camera B provided the best view of AARs because it showed the task assessor referring to various visual aids on the projected images and SMART Board.

Two images were projected onto the wall in the observation room. Projection 3 (see Figure 5) was the WD tactical situation display, which was updated in real time and showed the focus of the WD's attention. Projection 4 showed ground truth data and was also updated in real time. A listening post was available, providing five headphones for observers to listen to communication within and between the tactical control centre and White Force control centre. Observers could select the communications channels of interest using the communications system interface on the computer provided in the self briefing area. This computer could also display ground truth and could be used to zoom or pan to areas of interest on the ground truth display.

## 2.3 Tools and techniques

A major goal of EBS08 was to provide an environment within which a range of tools and techniques could be evaluated and developed for future use in RAAF training exercises. The following paragraphs describe the tools that were under evaluation in EBS08 in terms of their use and the improvements they may offer over current RAAF operations.

### 2.3.1 The *ADGESIM* simulation suite

The *ADGESIM* simulation system is comprised of the *DISVOX* radio simulator (described in Subsection 2.2.1), *PSI*, the *SensorLINK* simulated sensor gateway (described in Subsection 2.2.2.3), the SimMan simulator manager and the Tactical After Action Review for Distributed Interactive Simulation (*TAARDIS*; described in Subsection 2.3.5). The nature of the *PSI* software influenced SIMOP behaviour and consequently the fidelity of the training provided to the ABMs. For this reason, the properties of *PSI* are explained in further detail in Subsections 2.3.1.1 and 2.3.1.2.

#### 2.3.1.1 The *PSI* user interface

The *PSI* user interface is represented in Figure 6. Within *PSI*, the simulation scenario is displayed on a 3D world map with all objects drawn at ground level. Map overlays depicting the coastline, airspace, waypoints, routes and ground based radar coverage may be selectively displayed. Entities are drawn as small circles with history trails and velocity leaders, all of which may be colour coded to represent the force, domain or identity of the entity. Absolute and relative cursor position is displayed, and a bearing and range line tool is provided for measuring the relative position of, and distance between, entities. Alpha-numeric data block fields are drawn adjacent to entities, describing actual or derived state parameters. These fields are categorised as identity, position, targeting or missile related. A table summarising the IFF modes and codes, weapon and fuel state, velocity, altitude and current task of entities assigned to each instance of *PSI* is also displayed at the top of the panel.

Sensor coverage is depicted as one or more transparent zones around each entity on the world map, coloured by sensor type (as shown in Figure 6). Each entity has a visual sensor which is always on and, if fitted, radar and / or infrared sensors which may be exclusively enabled. Entities within a coverage zone of another entity are considered to have been detected by that entity. The supported types and modes of each sensor as well as the associated range, angle of azimuth and elevation constraints may be set for each entity type. For entities in formation, the sensor type in use can be synchronised across formation members, and the coverage zones are drawn as a merged volume. Alternatively the coverage zones of distinct formations are drawn as overlapping volumes. The state of 'weapon locks' between entities, in the form of pairing lines, warnings, weapon selection, weapon employment range and imminent loss of 'lock' indications, are also displayed on the map.

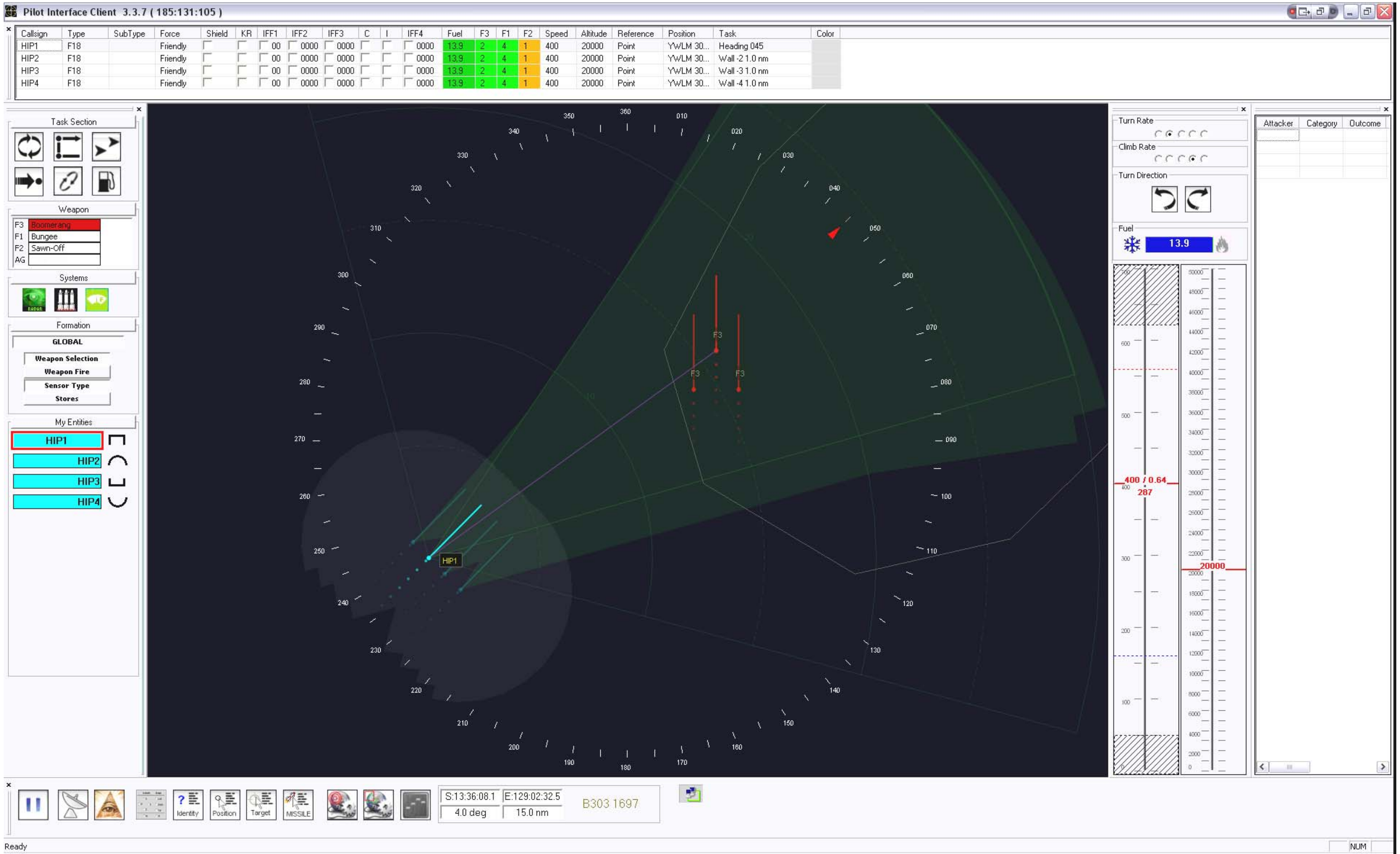



*Figure 6: The PSI user interface, depicting a formation of friendly entities intercepting an enemy formation. Shaded areas show sensor coverage approximations and weapons engagement zones.*

*PSI* provides several mechanisms by which SIMOPs may control entities. A user may quickly switch control between entities by selecting an entity on the map or cycling through the entity list with key commands. A SIMOP may manipulate the speed, altitude, and heading of an entity or formation via simple point and click controls on or adjacent to the map. The number of clicks determines the rate of change in the dimension specified. For instance, a single click speed change uses the 'military power' throttle condition and a double click uses the 'after burner' throttle condition. The magnitude of the command relative to the current setting may influence the manner in which the entity attains the desired state. For example, a formation will perform an immediate 'in place' manoeuvre, for a ~180 degree turn or a 'delay' manoeuvre for a ~90 degree turn. Entities may be tasked to fly a holding pattern, or follow a 3D route or another entity. Formations may be synchronised in terms of their sensor modes, weapon types and time of engagement. An entity may target another entity within its sensor coverage by right clicking on the target, and observe the state of weapon acquisition via distinct missile entity symbology.

A typical pattern of control in a simulation session would involve the following phases. Firstly, a SIMOP would configure the *PSI* display, in terms of overlays, map range and centre, data block fields and entity colour coding. Then a SIMOP would create entities required for the scenario from a script defining the starting states and tasks, or assign entities that had already been created. Subsequently, the SIMOP is able to select an entity or formation and modify their flight parameters, weapons system state or unattended tasks as described above and as required by the scenario. Finally a SIMOP may either destroy or release entities to other SIMOPs if no longer required for their direct control.

#### 2.3.1.2 *PSI* enhancements for EBS08

Each entity is accompanied by a data block listing its various characteristics and parameters. *PSI* considers ownership and context when selecting information to display in each data block. Whilst all entities are created on a central VR-Forces server, they may only be *assigned* to (that is, owned by) a single SIMOP at a time. Information concerning the sensors, weapons and fuel state of an entity are known only to the owner of that entity. The relative position and aspect of two entities, which may be required for tasks such as weapon targeting, may not be known unless one entity is *attached* (that is, the SIMOP controlling that entity has selected that aircraft as the *frame of reference*). Information is also filtered based on context. The first level of context relates to whether entity perspective is applied, that is, whether all of the entities in a scenario are displayed or just those that are assigned or detected by assigned entities. The second level relates to the type of sensor volume an entity is located within. For example, some position fields switch on and off as the range between entities increases or decreases (to display the units and frame of reference required by doctrine), and some identity fields are disabled because a given sensor type is not capable of determining a given parameter.

The context-based rules used to determine which information is displayed in data blocks in the latest version of *ADGESIM* are considerably different to those implemented in the version of *PSI* employed at SACTU and 3 Controlling and Reporting Unit (3CRU). The aim was to improve the consistency of SIMOP behaviour by restricting available information to only that which could reasonably be known by a pilot in that context. This change was expected to reduce SIMOP workload. Whilst the modifications achieved this aim, the difference in information requirements of SIMOPs and pilots was not considered. The EBS08 SIMOPs would have benefited from a higher level of SA to counteract the additional responsibility of coordinating the execution of the training scenario with other White Force members, a responsibility that would not be required of the pilots the SIMOPs represented. The overall effect of the modification to the *PSI* software was to create a higher level of SIMOP workload than would have been experienced with the previous version of *ADGESIM*.

### *2.3.2 Toteboard*

The *Toteboard* (example screenshot shown in Figure 7) is an electronic tool used to provide the ABM team with information on mission parameters, such as aircraft fuel states, alert states, weapon loadouts, assigned airspace and radio frequency, tanker offload levels, endurance estimates, and aircraft turnaround timings. The *Toteboard* is a customised Microsoft Excel spreadsheet that performs embedded calculations using Visual Basic for Applications (*VBA*) scripting. It was originally developed at SACTU for use in training events conducted in the SACTU simulator, but has since been utilised by the RAAF in a number of live military exercises. The *Toteboard* is a shared spreadsheet that allows every member of an ABM team to update information. As responsibilities are divided across members of an ABM team, this permits information on specific formations to be updated by the team member controlling those formations, circumventing double handling and reducing the risk of error.

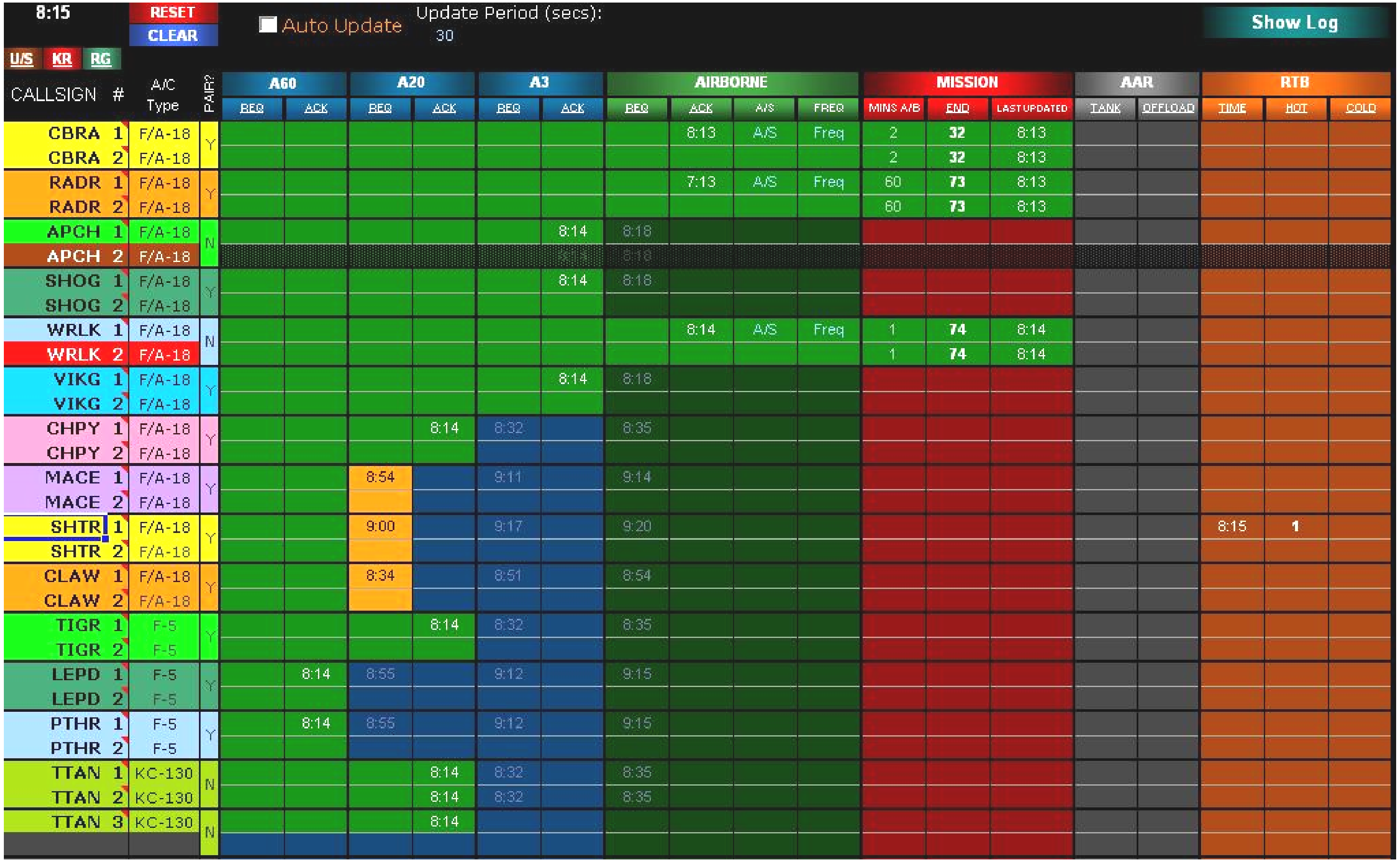
### 2.3.3 The *TDT* framework

*TDT* was developed by the US Navy in an effort to improve teamwork processes (Smith-Jentsch, Zeisig, Acton & McPherson, 1998). The goal of *TDT* is to improve team performance in four areas: information exchange, communication, supporting behaviour and initiative/leadership. Table 1 provides examples of good performance under each of these categories. The teamwork categories provide a framework for the structure of the *TDT* process.

*Table 1: The four areas of teamwork considered in the TDT framework*

| **Information exchange:** | **Communication delivery:** |
|---|---|
| - Utilising all sources of information,<br>- Passing information without being asked,<br>- Providing big picture situation updates. | - Using correct phraseology,<br>- Providing complete reports,<br>- Avoiding excess chatter,<br>- Communicating all information clearly. |
| **Supporting behaviour:** | **Initiative/leadership:** |
| - Monitoring and correcting team errors,<br>- Providing and requesting backup or assistance to balance workload. | - Providing guidance or suggestions to team mates,<br>- Stating clear and appropriate priorities. |

The *TDT* procedure was not fully implemented in EBS08, with the categories serving only to provide a common vocabulary for assessment, feedback and discussion of team processes within the context of the exercise. The ABM team was introduced to the *TDT* categories at the beginning of the exercise. They were provided with a written summary of an accident involving two US Army Blackhawk helicopters being shot down (included in Appendix D) and were asked to discuss the incident, identifying positive and negative examples of behaviours related to the four *TDT* teamwork categories. The ABD rated the teamwork performance of the ABM team under each of the categories during each mission, providing a teamwork debrief to the team after the completion of each mission.




| CALLSIGN | # | A/C Type | PAIR? | A60 REQ | A60 ACK | A20 REQ | A20 ACK | A3 REQ | A3 ACK | AIRBORNE REQ | AIRBORNE ACK | AIRBORNE A/S | AIRBORNE FREQ | MISSION MINS A/B | MISSION END | MISSION LAST UPDATED | AAR TANK | AAR OFFLOAD | RTB TIME | RTB HOT | RTB COLD |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| CBRA | 1 | F/A-18 | Y | | | | | | | | 8:13 | A/S | Freq | 2 | 32 | 8:13 | | | | | |
| CBRA | 2 | F/A-18 | | | | | | | | | | | | 2 | 32 | 8:13 | | | | | |
| RADR | 1 | F/A-18 | Y | | | | | | | | 7:13 | A/S | Freq | 60 | 73 | 8:13 | | | | | |
| RADR | 2 | F/A-18 | | | | | | | | | | | | 60 | 73 | 8:13 | | | | | |
| APCH | 1 | F/A-18 | N | | | | | | 8:14 | 8:18 | | | | | | | | | | | |
| APCH | 2 | F/A-18 | | | | | | | 8:14 | 8:18 | | | | | | | | | | | |
| SHOG | 1 | F/A-18 | Y | | | | | | 8:14 | 8:18 | | | | | | | | | | | |
| SHOG | 2 | F/A-18 | | | | | | | | | | | | | | | | | | | |
| WRLK | 1 | F/A-18 | N | | | | | | | | 8:14 | A/S | Freq | 1 | 74 | 8:14 | | | | | |
| WRLK | 2 | F/A-18 | | | | | | | | | | | | 1 | 74 | 8:14 | | | | | |
| VIKG | 1 | F/A-18 | Y | | | | | | 8:14 | 8:18 | | | | | | | | | | | |
| VIKG | 2 | F/A-18 | | | | | | | | | | | | | | | | | | | |
| CHPY | 1 | F/A-18 | Y | | | | 8:14 | 8:32 | | 8:35 | | | | | | | | | | | |
| CHPY | 2 | F/A-18 | | | | | | | | | | | | | | | | | | | |
| MACE | 1 | F/A-18 | Y | | | 8:54 | | 9:11 | | 9:14 | | | | | | | | | | | |
| MACE | 2 | F/A-18 | | | | | | | | | | | | | | | | | | | |
| SHTR | 1 | F/A-18 | Y | | | 9:00 | | 9:17 | | 9:20 | | | | | | | | | 8:15 | 1 | |
| SHTR | 2 | F/A-18 | | | | | | | | | | | | | | | | | | | |
| CLAW | 1 | F/A-18 | Y | | | 8:34 | | 8:51 | | 8:54 | | | | | | | | | | | |
| CLAW | 2 | F/A-18 | | | | | | | | | | | | | | | | | | | |
| TIGR | 1 | F-5 | Y | | | | 8:14 | 8:32 | | 8:35 | | | | | | | | | | | |
| TIGR | 2 | F-5 | | | | | | | | | | | | | | | | | | | |
| LEPD | 1 | F-5 | Y | | 8:14 | 8:55 | | 9:12 | | 9:15 | | | | | | | | | | | |
| LEPD | 2 | F-5 | | | | | | | | | | | | | | | | | | | |
| PTHR | 1 | F-5 | Y | | 8:14 | 8:55 | | 9:12 | | 9:15 | | | | | | | | | | | |
| PTHR | 2 | F-5 | | | | | | | | | | | | | | | | | | | |
| TTAN | 1 | KC-130 | N | | | | 8:14 | 8:32 | | 8:35 | | | | | | | | | | | |
| TTAN | 2 | KC-130 | | | | | 8:14 | 8:32 | | 8:35 | | | | | | | | | | | |
| TTAN | 3 | KC-130 | N | | | | 8:14 | | | | | | | | | | | | | | |

*Figure 7: The Toteboard used in EBS08. The rows in the spreadsheet correspond to specific entities and the columns to their corresponding parameters. Note that the colour-coding of entities by callsign matches the colour these entities will appear in PSI.*

### 2.3.4 The *AWAR* tool

The *AWAR* tool provides a means by which to capture team performance evaluations and provide structured debriefs. It was created under AIR 07/232 by customising a set of Excel spreadsheets using *VBA* scripting. During past exercises in this research program, the Mentor tool, developed by Calytrix Technologies (Fyshwick, ACT), has been used for this purpose (e.g., Best & Burchat, 2006; Best, et al., 2007). However, a tailored tool was required on this occasion to meet the specific requirements of EBS08. The *AWAR* tool can be populated with an extensive list of fine measures that can be used repeatedly across trials, allowing greater breadth, depth and consistency of ratings. The categories and measures also serve as focus areas around which an assessor may structure task performance debriefs.

For EBS08, the *AWAR* tool was populated with the ABM DCA team goal hierarchy, which was derived from the Australian Joint Essential Tasks (AS JETs). The AS JETs (McCarthy et al., 2003) are a comprehensive list of all tasks essential to the preparation, planning and conduct of Australian Defence Force (ADF) operations. The ABM DCA team goal hierarchy (Hasenbosch & Best, 2007) is an extrapolation of a subset of the AS JETs, including only those goals that relate to the work of ABM teams conducting DCA missions. The hierarchy consists of six team goal categories as well as subordinate goals, subgoals, and criteria. The relationship between levels of the hierarchy is such that the lower level goals and subgoals explain *how* the higher level goals and goal categories are achieved and the higher level goals and goal categories explain *why* the lower level goals and goal categories are important to achieve. The criteria provide a description of what constitutes good performance on each subgoal. The hierarchy represents the work of the ABM team in DCA missions in a way that facilitates understanding and reporting of how that work contributes to the higher-level organisational goals of the ADF.

#### *2.3.4.1 Performance assessment*

The training assessment approach currently employed in the RAAF involves subject matter experts (SMEs) rating trainee performance as competent (C) or not yet competent (NYC) on a small number of broad categories, with extensive notes detailing the positive and negative aspects of performance. Guidelines are provided for each category to increase the consistency of ratings, however, ratings still rely heavily on the SME's experience and may vary between assessors due to different preferences and different experiences of what works and doesn't work. There is no requirement to rate all measures, and assessors are free to focus on the areas they see as most important. Due to this flexibility, ratings may even vary for the same assessor over time. As a result, it can be difficult to make valid performance comparisons between students of a single course or to detect the performance improvement of a student or cohort over time. Although the operators agree that the consistency and objectivity of ratings should be improved, they value the flexibility of the current system and the avenue it provides for experts to pass on their experiences to students.

A screenshot of the *AWAR* tool, populated with the ABM DCA team goal hierarchy, is shown in Figure 8. As shown, the six team goal categories can be expanded to reveal the constituent team goals and team sub-goals. Each sub-goal is accompanied by a criterion for rating team performance. Assessors rate performance by selecting a score from a drop down menu. They can also provide explanatory remarks. In EBS08, the assessor selected a score from the following five options: 1 = Terrible, 2 = Poor, 3 = Adequate, 4 = Good, and 5 = Excellent. Sub-goals that were not rated retained the default score of 0. Figure 8 depicts performance on *Sub-Goal 5.1.4: Manage AAR operations* being rated as '*Terrible*'. The assessor could also label overall team goals as strengths or weaknesses by depressing the '*strength*' or '*weakness*' button corresponding to the relevant goal. Some sub-goals relating to safety, for example *Sub-Goal 5.2.1: Apply Current ROE*, were labelled as critical, and a rating of 1 or 2 would result in the team failing the overall mission. This 'critical fail' function could be switched off if appropriate.

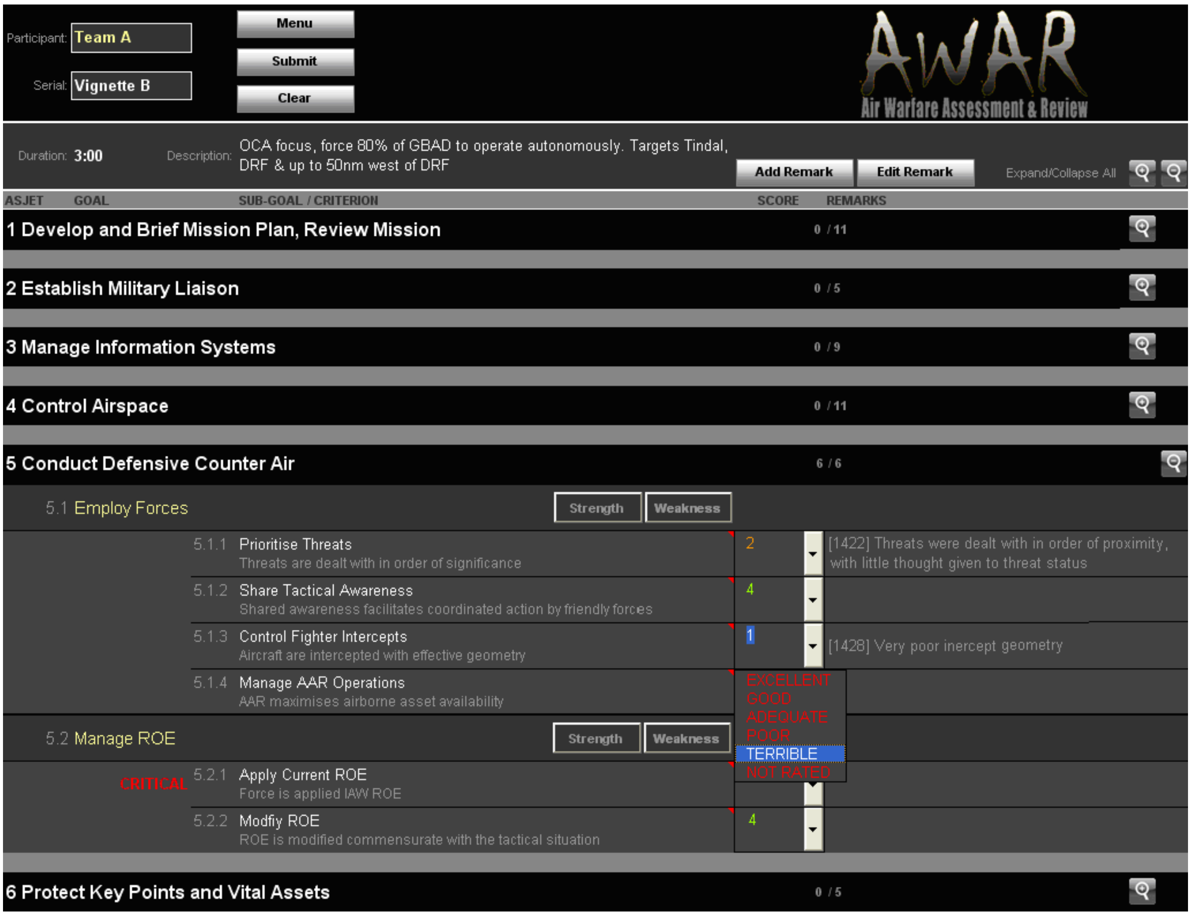



*Figure 8: The AWAR assessment user interface. Goal category 5 has been expanded to reveal the constituent goals and sub-goals. The figure also shows the drop down rating menu. Note that the comments included in this figure are fictional and are not those made by the assessor during EBS08.*

The DCA team goal hierarchy was incorporated into the *AWAR* tool with the intention of increasing the consistency of performance assessments and the specificity of feedback over the assessment approach currently employed by the RAAF. Inclusion of the team goal hierarchy prompted the assessor to rate performance on a wide range of specific sub-goals, rather than broad multifaceted goals, and performance criteria provided guidance on what constituted good performance on each sub-goal, promoting the consistency of performance assessment across missions. The same goal hierarchy was used for the duration of EBS08 and PB08, and therefore, ABM teams were rated on the same list of measures on all occasions. This permitted the comparison of performance across missions, across exercises and between teams. Note that due to time restrictions, the task assessor was able to rate performance on most, but not all, sub-goals. Use of the team goal hierarchy served to highlight the relationship between specific team behaviours and high level ADF goals, thereby raising the team's awareness of the importance of each sub-goal. It also provided a framework for providing the team with performance feedback.

#### *2.3.4.2 Performance review*

A screenshot of the *AWAR* debrief user interface is shown in Figure 9. In this figure, *Team Goal 3.2: Manage Equipment* has been expanded to reveal the underlying ratings of team performance. The scores awarded by the task assessor on each sub-goal were aggregated to produce the average for each team goal and goal category, for example, the average of all scores underlying *Team Goal 3.2: Manage Equipment* is equal to 3.4. Unrated sub-goals were omitted from this calculation. Pass / fail criteria can be modified to suit the goals of an exercise. In EBS08, sub-goals on which performance was rated as 1 or 2 were designated fails, as were goals and goal categories where underlying scores averaged to less than 2.5. In addition, any goal or goal category with an underlying critical fail was designated a fail also. Scores for sub-goals, goals and goal categories were colour coded, with high scores appearing in green, moderate scores appearing in yellow, low scores appearing in orange, and scores resulting in a fail appearing in red. Sub-goals that were not rated were omitted from the performance summary.

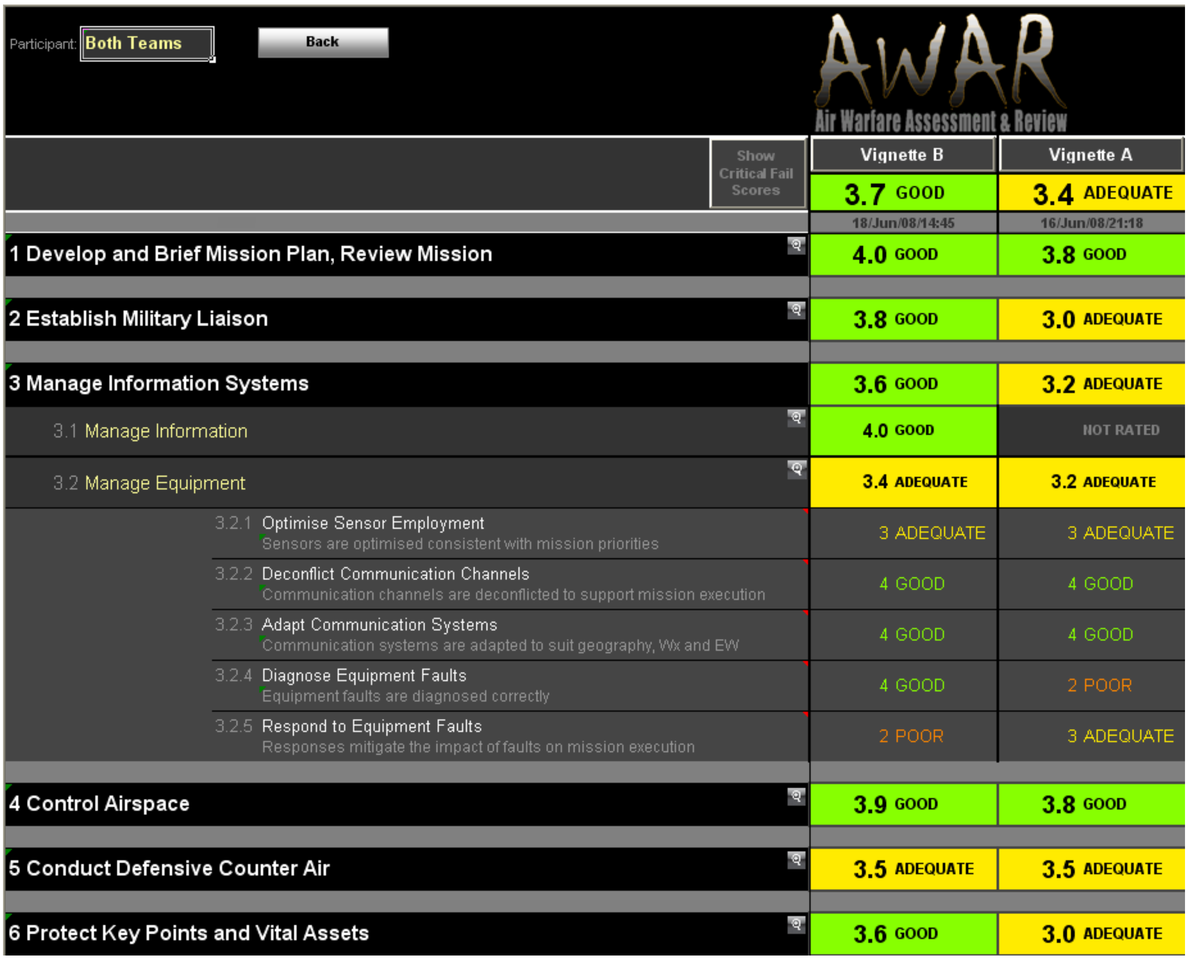



*Figure 9: The AWAR debrief user interface, showing the ABM team's performance on Vignettes A and B. Goal 3.2 has been expanded to reveal performance ratings on all underlying sub-goals.*

The populated *AWAR* tool provides access to the electronic scores entered by an assessor immediately after the completion of a mission. This data can be used to guide the task performance debrief, with colour coding drawing the attention to instances of exceptionally good or exceptionally poor performance and critical fails. Data can be displayed alongside data from previous missions for comparison. In Figure 9, team performance on Vignette A is displayed alongside team performance on Vignette B. Figure 9 also shows that comments can be viewed by floating the mouse over the associated score. Graphs can be produced to compare the performance of a single team between missions, or of two or more teams completing the same mission. The prompt availability and electronic format of scores contrasts with the approach currently employed by the RAAF, where scores and comments are hand written on paper forms. While these notes guide the debrief, no scores displayed to the team. Scores and comments are typed into an electronic form and printed for team members after the completion of the mission. However, this provision of written feedback may take weeks, depending on the workload of the assessor, by which time team member's memories of specific aspects of the mission will have degraded and feedback will be less effective. The *AWAR* tool allows access to this data immediately and circumvents the double handling that currently occurs.

### 2.3.5 AAR approach and tools

At the conclusion of each EBS08 mission, the ABM instructor and ABD debriefed participants on their task work and teamwork performance. During this activity, carried out as an interactive AAR session using the tools in the *ADGESIM* suite, assessors highlighted specific areas for improvement, made comparisons between the mission just completed and previous missions, and outlined specific instances where breakdowns in doctrine, communications, and tactics occurred.

The approach taken to AAR in EBS08 was to juxtapose two different visual representations of the scenario under review. The first representation, known as 'user truth', was comprised of elements of the scenario as they were seen by individual team members, and the resultant actions taken in response to these elements. This representation is distinctly different from the second, 'ground truth', which describes what actually transpired, insofar as capturing the idealised representation of entities and interactions at the fidelity modelled by the simulator itself. This juxtaposition allowed the assessors to compare what was happening in the simulation with what the users were seeing and their consequent reactions, giving participants the opportunity to learn lessons about their actions in a broader, exercise-wide, context.

In order to implement this method, the simulation environment was configured with a subset of tools from the *ADGESIM* suite, designed to facilitate simulation recording and AAR. This package of tools, known as the *TAARDIS* suite, allows an AAR activity of this kind to take place in real-time, with minimal downtime between the end of simulation and the beginning of AAR. The sub-applications of which *TAARDIS* is comprised are: Console, *WorldView*, *Pensieve* and *ReMarks*, with each playing a specific role in the AAR activity.

The main component of *TAARDIS*, the Console, primarily allows real-time recording and playback of scenario network data, in the form of DIS Protocol Data Units (PDU). These PDUs encapsulate all entity activities, weapon interactions, and electromagnetic transmissions (including IFF, weapons-lock, and radio traffic), as well as additional simulation data used in modelling and scenario representation. This stream of data describes exactly what data was used in the generation of the scenario, providing a 'ground truth' for the exercise. *TAARDIS* Console allows for playback of this data in real-time, or at speeds slower or faster than real-time. This data can be used to re-stimulate simulation components in order to regenerate the scenario, or can be viewed directly using the *TAARDIS WorldView* tool, as was the case in EBS08. Console also provides for remote control and data collection for other *TAARDIS* components. This means that a single assessor has control over data collection and replay of an entire scenario.

Replay data from *TAARDIS* Console is interpreted by *TAARDIS WorldView*, and displayed symbolically on a three-dimensional world map, in a manner consistent with other *ADGESIM* simulation tools. Additional second order information, such as targeting and weapons deployment information, is also displayed, allowing an assessor to view scenario events as they actually happened.

The corollary to this, that is, what the user actually sees (i.e. 'user truth'), is captured by *TAARDIS Pensieve*. This tool is remotely controlled by Console, and is used to capture the screen activities, audio input and audio output of a single ABM team member's console. When deployed across all work stations, it allows the capture of 'user truth' for the entire team. For AAR, *Pensieve* switches from recording to playback mode, and shows the entire screen recording as a video playback. This playback is controlled remotely by Console, ensuring that 'ground truth' and 'user truth' remain in synchrony, allowing viewers to make comparisons between the two.

The final element of the suite, *ReMarks*, allows an assessor to insert timestamped, descriptive bookmarks into the Console recording during mission execution. These bookmarks highlight events of interest for later review, and during an AAR session assessors can rapidly skip to these bookmarks to emphasise a particular lesson or learning outcome.

Given this synchronous recording, bookmarking, and replay capability, assessors were able to accurately represent both user and ground truth in a unified, synchronous manner. During EBS08, AAR generally involved only the WD's video, displayed on Projection 2, with complementary ground truth being shown on Projection 1 (see Figure 4). This side-by-side representation of data was valuable during the activity as it allowed ABM team members to compare their perception of the exercise (as shown on their console), with the events that actually occurred. In addition, the task assessor was able to visually illustrate shortcomings or events of interest to the team and, by using bookmarks, navigate from one event of interest to the next easily and rapidly.

# 3. Participant feedback on tools and techniques

Overall, the feedback from EBS08 participants was very positive, with all participants believing their performance at BS08 would benefit from the additional training they received at EBS08 and several remarking that activities of this kind should be used to support preparation for all large scale live exercises involving 41WG in the future. Participants also provided feedback on each of the tools and techniques used in EBS08. Overall, feedback was positive, but some shortcomings were noted. The following subsections summarise the information provided by participants and the observations made by human factors scientists regarding the strengths and limitations of the tools and techniques evaluated. Each subsection concludes with a list of recommendations derived from this information. The complete list of recommendations is included in Appendix B.

## 3.1 The *ADGESIM* simulation system

The following sections outline the comments made by the participants and the observations of human factors scientists in regard to the *ADGESIM* simulation system. The feedback of the SIMOP team focussed mainly on preparation, capacity and coordination issues, and the feedback of the ABM team addressed the fidelity of the training experience they received. The feedback of exercise staff generally related to the performance of the system and the number of operators required for successful implementation.

### 3.1.1 SIMOP feedback on the *ADGESIM* simulation system

Members of the SIMOP team reported that the one-hour familiarisation session on the new version of *PSI* and the communications system, *DISVOX*, was not sufficient given the differences between these tools and the baseline *ADGESIM* capability with which they were more familiar. In addition, some aspects of the context based information display concept embodied in the latest version of *PSI* were not completely validated prior to EBS08. Due to the relatively brief training provided on Day 1 of EBS, the SIMOPs were not aware of all available software functions, with some even mistaking unfamiliar features for faults or vice versa. Even by the last day of the exercise, the operators were not using these systems to their full potential due to lack of familiarity and their performance was affected by minor software faults or deficiencies. The SIMOP team requested that a period of two days be set aside prior to the exercise for the provision of training on unfamiliar systems in future. They believed that this would improve the quality of training provided to the ABM team through a reduction in operator error and inefficient behaviour.

The workload reported by the SIMOPs during the EBS08 scenario vignettes was high, as shown in Figure 10. Figure 10 also shows a considerable difference between the workload reported by SIMOPs and that reported by the ABM team, suggesting that the SIMOPs found it difficult to create a sufficiently challenging environment for the ABM team. Under high workload conditions the SIMOPs were observed to shed a number of tasks, including maintaining awareness of entity altitude and speed, announcing the future intent of their entities and acknowledging requests from team members. Specific cognitive aspects of these tasks, and the relative importance of the tasks themselves, would have contributed to the likelihood that they were shed ahead of other tasks. These unfulfilled tasks reduced the realism of the training scenario. The SIMOP team attributed a large proportion of their workload to the number of entities they were required to control. In a SACTU training context, the number of entities assigned to an individual operator is not likely to exceed 8, yet at EBS08 this number was at times greater than 20. The SIMOPs believed that, with further improvement and automation, *PSI* could support the coordinated control of a large number of entities. Additional workload was incurred when a single operator was assigned multiple formations with varying warfare roles, such as ‘strikers’ versus ‘sweepers’. These roles are distinguished in terms of aircraft performance constraints as well as timing and risk related objectives; and the Blue SIMOP team found it difficult to switch between the mindsets required to perform these different roles.

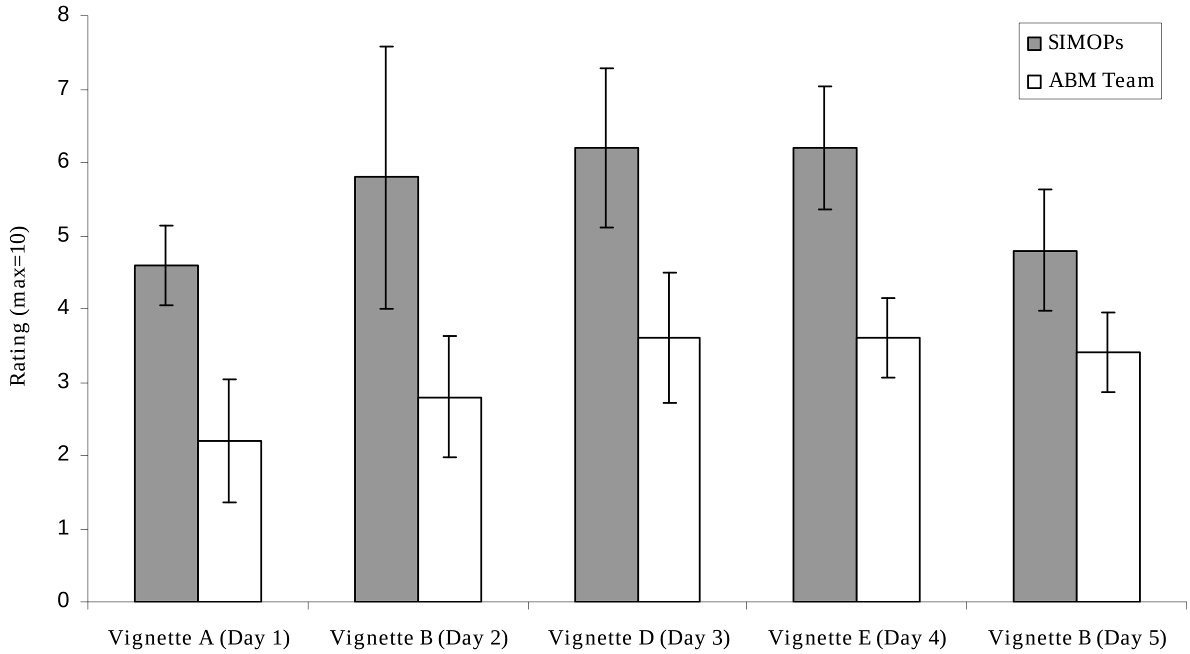



*Figure 10: The self-reported workload of the SIMOPs and ABM team during the vignettes they experienced over the five days of EBS08*

The Red SIMOP team reported an extremely high internal communications load. They believed that this was due to a lack of awareness within the team of which tasks were performed by each member. Additional communication was required to verify entity ownership and the radio frequencies monitored by each SIMOP. The RSC was sometimes unsure of which SIMOP was best placed to receive a scrambled or 'regenerated' formation. As a result, workload was often not balanced between team members. Over the course of the exercise, coordination of the Red SIMOP team improved via discussion of a takeover / handover plan after each vignette.

### 3.1.2 ABM feedback on training fidelity

Members of the ABM team reported that there were some instances during the exercise where they were required to modify their behaviour to account for limitations in the fidelity of the exercise. These limitations fell within two main categories. The first category was the perceived fidelity of the synthetic C2 'picture' provided by the *ADGESIM* applications. This representation is a combination of the qualities of sensor outputs and the behaviour of detected entities. The latter is confounded by both the performance of aircraft models and the application of tactics by SIMOPs. Second, the perceived fidelity of communications in the simulation system, in terms of the intensity and complexity of information exchange behaviours and the constraints of the medium.

The ABM team reported that sensor coverage and aircraft models generally produced realistic detections. However, at the beginning of exercise, a *SensorLink* fault caused the IFF code of a specific mode that was assigned to a given aircraft to be reported as part of a radar contact for another aircraft. Under exercise conditions, specific IFF modes and codes are used to identify aircraft and indicate to all participants when aircraft are 'out of exercise'. The fault was noticed when the status of an aircraft was incorrectly decoded as ‘out of exercise’, confusing the ABM team. Fortunately, the fault occurred only at the beginning and was rectified for the remainder of the exercise and therefore caused little disruption.

While the EBS08 C2 picture was complex relative to many other simulation exercises, the ABM team felt it was somewhat less complex than live exercises. This was due primarily to constraints imposed by the size of the White Force. Under conditions of high workload SIMOPs attempted to reduce the complexity of their control task, by limiting the use of altitude changes to entire formations. This resulted in a scenario presentation that was more two-dimensional than usual. In live, particularly multinational, exercises, formations may often remain within a single safe altitude block to reduce the risk of collision. However, in general the ABM team perceived that the SIMOPs did not have the same instinctive appreciation for doctrine and hence three-dimensional geometry as pilots. In smaller scale simulations SIMOPs can play pilot roles convincingly, yet the high workload experienced in EBS08 reduced the ability of SIMOPs to ‘fly’ realistically. The ABM team did not believe that it would be feasible to train SIMOPs to produce the desired behaviours under such high workload.

The ABM team reported a lack of realism in the pattern of information exchange with the simulated pilots. In the live domain, both the ABM team and pilots push and pull information. During EBS08 the SIMOPs only pushed a subset of the typical information, and occasionally information was not pulled as expected. For example, SIMOPs were unable to replicate the breadth of ’fill in’ communications, such as those from pilots sharing information with other members of the formation and the ABM. Similarly, there were fewer instances of communications being ‘stepped on’, that is, multiple parties attempting to transmit on a radio channel at once. SIMOPs are unable to ‘step on’ their own transmissions, as a single SIMOP may only communicate for one ‘pilot’ of the group they represent at a time. The reduction of ‘fill in’ communication may have been exacerbated by high workload. The challenge of maintaining and communicating accurate SA, despite the constraints of a real communications system, is a substantial source of ABM workload and hence scenario intensity in a live environment. The ABM team believed that the absence of some of these constraints decreased training fidelity. Specifically, it forced the ABM team to pull more information from the SIMOPs and exercise greater control over their actions than is typical in the live domain.

The ABMs could hear communications between White Force members and often waited until the SIMOP ‘flying’ the aircraft they needed to redirect was available to communicate. This was possible because the ABM team was physically separated from the White Force only by thin partitions that did not reach the ceiling and will be resolved in future exercises if the teams are located in separate rooms. In a live environment, ABMs may delay communications to a pilot in their zone if the pilot is overheard speaking with another pilot via radio. However, the additional avenue of insight into pilot communications loads that arose at EBS08 was artificial and impacted on the realism of the simulation.

The ABM team also stated that the simulated radio coverage provided in EBS08 was inconsistent with the simulated radar coverage and did not degrade in any realistic way. In the live domain, and if ground based radio and radar systems are co-located, aircraft that cannot be ‘seen’ are not generally ‘heard’. Furthermore, the loss of radio coverage may be gradual such that either the ground station or the aircraft may still receive a weakened signal, even if their own transmissions are not received. A loss of coverage may prompt other aircraft to relay messages between ABMs and pilots who do not have direct communication. Such behaviour did not occur in EBS08, as perfect radio communication was possible at all times.

### 3.1.3 Exercise staff feedback on the *ADGESIM* simulation system

*PSI* and its related server components suffered several major and minor faults through the exercise. These were severe on the first day of the exercise, but were reduced over the course of the week. Some of the faults were related to new functionality added just prior to exercise commencement. Future exercises should utilise mature software versions and prohibit last minute enhancements. Adequate time must be set aside to perform scaled, scenario-specific testing to ensure that all technical issues are resolved before the exercise begins. Preparation activities should also include the development of a human-machine interface and team coordination procedures training package that is relevant to the activity. This should be provided to the SIMOPs in clear and concise manner in advance of the exercise.

Regardless of enhancements to preparation or software interfaces, the available White Force personnel were insufficient to represent the number and nature of entities required by the EBS08 / PB08 scenario without some compromises to behavioural fidelity. Ideally at least four Red Force and three Blue Force SIMOPs, as well as one experienced mission director assigned to each side would have been employed as the White Force. Unfortunately, all of the RAAF personnel assigned as SIMOPs at the time were involved in the exercise. In the short term, additional SIMOPs are required to meet the training needs of similar preparation activities. These could be either recruited by the RAAF, or contracted through civilian service providers such as Milskil. In the long term, additional automation is required within computer generated forces packages such as *ADGESIM* to reduce SIMOP workload. Automation should address mechanisms by which the ratio of SIMOPs to exercise participants can be reduced, such as voice recognition and synthetic speech technologies to allow agents to control entities.

### 3.1.4 Recommendations:

1. All simulator-related technical briefs should be clear and contain the appropriate level of detail.
2. SIMOPs should receive adequate training in the use of all software and simulation systems before the commencement of an exercise. The SIMOP team suggested that two days of focused training in the week prior to EBS08 would have permitted them to become sufficiently familiar with the new versions of *PSI* and *DISVOX*.
3. Adequate time should be reserved in the weeks preceding an exercise for a thorough and systematic testing of *ADGESIM* systems. All technical issues must be resolved prior to the exercise with no additions or modifications made after testing.
4. SIMOP workload will be considerably reduced if entities are assigned by role rather than location. White Force scenario managers should endeavour to achieve this wherever possible.

5. Further workload reduction can be achieved by increasing the number of SIMOPs assigned to the White Force. However, all qualified and partially qualified RAAF SIMOPs were present at EBS08 and the pool of available SIMOPs must therefore be increased.
6. The *ADGESIM* simulation suite should be further automated with the goal of further reducing the number of SIMOPs required to produce a complex scenario. However, modifications must be sufficiently tested to ensure that they are not made at the expense of realism.
7. The ABM team was able to hear the activity of the SIMOPs and adjusted their communications accordingly. Ideally, SIMOPs and ABM teams should operate in completely separate rooms rather than partitioned sections of the same room.

## *3.2 Toteboard*

### 3.2.1 General comments

The *Toteboard* was originally developed at SACTU for use in training events conducted in the SACTU simulator, but has since been more broadly utilised. 41WG ABMs have used the *Toteboard* in all recent major exercises involving their participation, including Pitch Black, Aces North, Aces South and the East Coast Air Defence Exercise (ECADEX). All ABM team members had therefore been exposed to the *Toteboard*, either through a SACTU simulation exercise or a live exercise at 3CRU, prior to participating in EBS08.

The ABM team felt that the *Toteboard* provided all of the information they required during mission execution. An additional benefit was that the WFMD was able to view the *Toteboard* during missions. This provided an effective means for coordination between the tactical control centre and the White Force control centre, for example when launching formations. In these instances, an ABM team member would launch a formation and annotate it accordingly on the *Toteboard*. The WFMD would receive the request and action it through coordination with the SIMOPs.

The RAAF participants did, however, offer two suggestions for improvement. Firstly, they suggested that the accuracy of endurance estimates be improved. The version of the *Toteboard* used in EBS08 calculates endurance, that is, airborne time remaining before the aircraft runs out of fuel, by subtracting the number of minutes an aircraft has been airborne from the number of minutes that type of aircraft could be expected to operate on a full fuel load. The ABM team suggested that the calculation would be much more accurate if it also considered weather, airfield conditions and fuel burn rate. They also requested an estimate of *bingo* fuel, that is, the time after which the aircraft will not have sufficient fuel to return from its current location to home base. Secondly, the *Toteboard* should be protected from corruption due to competing user inputs. The *Toteboard* was a shared spreadsheet and information could therefore be updated by multiple users. However, if two users concurrently updated the same piece of information, the *Toteboard* malfunctioned, causing the information to be lost, the spreadsheet to be locked prohibiting the entry of further information or the system to crash. The *Toteboard* has since been upgraded in line with these recommendations.

### 3.2.2 Recommendations

1. The most recent version of the *Toteboard* should be used for all future ABM simulation exercises.

## 3.3 The *TDT* framework

### 3.3.1 General comments

The initial reaction of the ABM team to the *TDT* concept was positive. However, they did not find the *TDT* introductory exercise (refer Appendix C) valuable. They felt that the context and behaviours displayed by personnel described in the written summary of the Blackhawk accident were extreme and were not sufficiently aligned with the ABM team's own work environment. As a result, they reported that they did not believe the introductory *TDT* exercise improved their ability to recognise subtle examples of good and bad teamwork behaviour. This negative appraisal may have had an impact on the team's participation in *TDT* discussions and on their motivation to attend to teamwork factors during the exercise.

A requirement of the *TDT* approach is that the workload of the participating team be sufficiently high, with sufficient strain placed on communication and coordination between team members. Unfortunately, ABM team workload was lower than intended due to system instability and the SIMOPs lack of familiarity with *ADGESIM*. As a result, the EBS08 vignettes did not place a great amount of strain on teamwork behaviours and the ABD struggled to find examples of behaviour that could be improved. This further impacted on the motivation of the ABM team to engage in the *TDT* portion of the debriefs.

Smith-Jentsch, et al. (1998) suggest that teamwork performance be evaluated by four *TDT* assessors, with one allocated to each *TDT* category. This arrangement was not possible due to personnel constraints and the ABD was therefore tasked with assessing performance within all four categories. According to Smith-Jentsch, et al., this is an acceptable situation if the assessor selected has extensive experience in *TDT* assessment. Therefore, the *TDT* aspects of the EBS08 assessment and feedback may have been improved had the ABD been given greater opportunity to become familiar with the approach prior to the commencement of the exercise.

The role of the facilitator during a *TDT* debrief is to guide the team through a process of self-evaluation and correction. At EBS08, this role was not fully implemented, with the ABD often providing direct feedback, rather than facilitation. The ABM team did report that there were instances where the *TDT* debrief stimulated discussions not raised in the task performance debrief. However, the largely uni-directional nature of feedback may have reduced the motivation of the team members and is likely to have reduced benefits to teamwork. If *TDT* is to be used to structure assessment and feedback in future exercises, assessors should be provided with training that improves their understanding of the framework itself and of the importance of the facilitatory style on which it is based.

The ABM team reported that all four of the *TDT* dimensions were familiar constructs and that they were reflected in the performance assessment approach currently employed within the RAAF but were more implicitly than explicitly trained. The ABD felt that the issues raised in his *TDT* debriefs would have been implicitly covered in a standard debrief, regardless of whether he had been introduced to the *TDT* procedure, and the ABM team felt that an integrated debrief would be less disjointed and more effective. The ABM team did not feel that the *TDT* framework complemented the culture within the Australian military, which was described as having less emphasis on rigid doctrine and structured training than the US military. This reflected the overall sentiment of participants that the *TDT* process, as it was implemented at EBS08, offered little means for the improvement of teamwork processes. The ABM team felt that the *TDT* process would be more appropriate for less experienced personnel who had not yet grasped the teamwork concepts implicit in training. They believed it would be of particular use if its implementation resulted in the introduction of a common parlance with which to discuss issues of team behaviour.

There is considerable evidence to suggest that training focusing on team processes can lead to improved team performance (Kozlowski & Ilgen, 2006). There is also considerable support for the *TDT* framework as an effective means by which to train teams to work together better, at least within the context of the US Navy (Smith-Jentsch, et al., 1998). At EBS08, the implementation of *TDT* was only partial and therefore conclusions regarding the utility of the approach cannot be drawn. What was clear, however, was the resistance of the RAAF ABM team to the approach. For approaches aimed specifically at improving teamwork processes to be effective in future exercises, they must be accepted by participating teams. This will require an investigation of potential modifications to bring these approaches into line with RAAF culture.

### 3.3.2 Recommendations

1. If a *TDT* introductory exercise is used to train a team to recognise instances of the four *TDT* elements, care should be taken when selecting the scenario description to ensure that it bears relevance to the context within which the team usually works. The description should include realistic examples of good and bad behaviour as well as extreme examples included to maintain interest. Learning to recognise these subtle examples will better equip the team to recognise such instances in their own performance.
2. The scenario workload must be high enough to genuinely test teamwork and communication skills.
3. If *TDT* is to be included in future exercises, assessment should be completed by four *TDT* assessors, with one assessing performance under each *TDT* category. If this is not possible due to personnel constraints, the assessors that are involved should undergo extensive training in *TDT* assessment prior to the commencement of the exercise.
4. Future *TDT* assessors should undergo training that improves their understanding of the importance of guiding the team through a process of self-evaluation and correction, rather than providing direct and unidirectional feedback, and their skills in encouraging team member participation.
5. The ABM team reported that the structure inherent to *TDT* did not compliment Australian military culture. Modifications aimed at improving the suitability of *TDT* within the context of the Australian military should therefore be considered.

## 3.4 The *AWAR* tool

### 3.4.1 Comparison with current RAAF assessment approach

Use of the team goal hierarchy forced the RAAF instructor assessing task performance to consider a wider range of performance aspects in greater detail than would be the case using the assessment system currently employed by the RAAF. The ABM team felt that the team goal hierarchy permitted performance to be assessed more objectively, as the sub-goals were more specific than the measures currently used. However, they felt that the sub-goals were still too broad, leaving them open to interpretation. It was suggested that performance could be more objectively assessed if the *AWAR* tool was populated with a longer list of finer-grained measures, and an assessor was asked to rate performance as C or NYC, rather than on a five-point scale. Due to the complexity of the environment within which an ABM team operates, such a list would need to be extremely extensive and detailed. Assessors are already under considerable time pressure and could not commit the additional time required to evaluate performance on a still greater number of measures. The ABM team and the task assessor felt that the ratings of '*terrible*', '*poor*', '*adequate*', '*good*' and '*excellent*' were subjective. However, these rating labels were chosen because research suggests that they attract greater agreement between assessors than a purely numerical scale (Whitley, 1996). In addition, a five-point rating scale with meaningful and quantifiable rating labels provides greater insight into how much improvement is required. If a five-point rating scale is to be used, the ABM team suggested that a detailed description of each of the five performance ratings on each measure be included. While this may promote the reliability and validity of performance measurement, the amount of SME input that would be required to clearly describe behaviours characteristic of performance for each rating of each measure would be prohibitive. It is not clear whether the best approach to assessment is for assessors to rate team performance as C or NYC on a long, fine-grained list of measures or to rate performance on a shorter list of broader measures using a five-point rating scale. It does seem, however, that a shorter list of broader measures would be more resilient if the two approaches imperfectly applied due to time constraints. Further research is needed to determine the approach most likely to produce consistent ratings between assessors and over time given the constraints within which RAAF training takes place.

The RAAF participants felt that the team goal hierarchy would be extremely useful for designing basic course scenarios at SACTU and the School of Air Warfare (SAW). The current process involves the assessors inserting events into a scenario that they believe will provide students with a valuable training experience and the opportunity to showcase their skills. The method is not structured, relying entirely on the expertise of assessors, and the events inserted and training received is therefore likely to vary between assessors and over time. In addition, scenarios are usually created under time pressure just prior to their use. Using a structured tool like the team goal hierarchy to design scenarios would promote consistency between students in terms of their exposure to events. Using the same scenario for all students at a course level would permit valid performance comparisons between students.

### 3.4.2 Performance assessment

The RAAF instructor playing the role of task assessor was not able to rate performance on all measures while missions were in progress and therefore completed performance assessments after their conclusion. The reasons given for this were two-fold. Firstly, some measures were considered to relate to holistic, rather than specific, aspects of performance and therefore could not be rated until after the conclusion of the mission. The utility of the team goal hierarchy relies on the specificity of measures, and this comment therefore suggests the need to revise some measures. Secondly, although the task assessor found the *AWAR* tool simple and intuitive, he was not sufficiently familiar with the arrangement of goals and sub-goals in the team goal hierarchy and therefore spent a considerable amount of time searching for sub-goals. Assessors at SACTU and the SAW) would be similarly unfamiliar with the team goal hierarchy, and would require sufficient training if it was to be adopted by these groups. In addition to promoting familiarity with the team goal hierarchy, such training should aim to develop skills in assessing entire teams rather than individual team members.

### 3.4.3 Performance feedback

The *AWAR* tool debrief user interface was designed so that only one goal category at a time could be expanded to reveal its constituent goals and sub-goals. This was done in order to minimise distraction from scores that were not related to the current discussion. The ABM team found this arrangement effective and intuitive, but felt that it drew unwarranted attention to the six goal category averages. These averages did not provide information or direction on how to improve and could therefore be seen, themselves, as a distraction. It was suggested that the format be modified to display all sub-goals and accompanying scores, with the assessor scrolling to those of relevance as they are mentioned in the task performance debrief. Perhaps the elements within the goal category or goal of interest could be highlighted with all others visually faded. As opinions were mixed regarding the comparative utility of these options, the solution may be to display collapsed goal categories on the debrief user interface, but to provide students with a printout of all scores and comments for reference during the task performance debrief. The colour coding of goal category averages in the debrief user interface was highly valued as it drew attention to underlying low scores and critical fails. The scores and comments relating to each sub-goal were extremely found to be useful, as they provided the team with an understanding of their current performance, showed how much room existed for improvement, and provided instructions for improving.

There are a number of advantages to storing scores and comments in an electronic format. The electronic format permits the graphical or tabular comparison of scores achieved by different teams or by the same team on different missions. This permits the evaluation of performance differences between teams that have experienced different training conditions or the evaluation of performance improvement over time for a single group. The ABM team valued the ability to display their scores at the beginning of the week alongside their scores at the end of the week to determine the sub-goals on which they improved over the course of the week. As stated, the ABM team did not find the average score achieved on each goal category useful, and therefore did not find it useful to graph these averages. It may be, however, that graphs that permit more meaningful comparisons, for example, performance differences on specific measures between teams or between instances for the same team, may be viewed more positively. In addition, information from the electronic database can be used to improve

RAAF training courses. Anonymous data could be compiled and analysed to indicate the specific tasks that RAAF students at certain qualification levels find most difficult, and training could be modified to improve performance on these tasks. Note that this possibility exists only if performance is always assessed using the same performance measures.

Entering performance assessment scores electronically from the outset circumvents the double handling that occurs when paper and pencil performance ratings are transferred into a database. The current RAAF practice is for paper and pencil assessments to be entered into a database, printed and provided to students as soon as possible after the completion of a mission. Students then refer to this report in their own time. However, during busy periods, providing written feedback can take up to a week, after which time comments and suggestions are far less effective as students may have trouble remembering the specific details of the relevant mission. If reports could be provided to students within minutes of completing a mission, and referred to during debrief, student learning is likely to improve considerably. The current version of *AWAR* does not include a function that allows the list of sub-goals and associated performance ratings to be printed. However, the importance of this feature has been recognised and it will therefore be a feature that is sought in tools of this kind for future exercises. It was suggested that assessors should have the ability to export scores and comments to a spreadsheet either under the goal categories listed in the team goal hierarchy or in the order of timestamps attached to assessor comments. The RAAF participants suggested that the performance report include three to five good aspects of their performance and three to five areas for improvement. They also suggested that it may be possible for basic course students to debrief themselves after the completion of a mission using the performance report and video and audio recordings. This would be particularly beneficial when there are very few instructors with great demands on their time, as is the case at SAW. The benefits and drawbacks of this approach, in terms of student learning, would need to be thoroughly evaluated before implementing this approach.

The ABM team suggested that if a tool like *AWAR* is to be used by an ABD, efforts should be made to increase its portability, as ABDs often rate performance while standing behind ABMs. Consideration of very lightweight tablet PC with handwriting recognition software is therefore required. If handwriting recognition software is employed, it should permit assessors to store commonly used symbols, shorthand and abbreviations. Participants also suggested that more space be provided for assessors to comment on performance and be able to draw freehand diagrams to support the comments made. The task assessor suggested that the utility of the tool could be improved through greater automation. He suggested the inclusion of a feature permitting an assessor conducting a task performance debrief to click on a timestamp next to a comment recorded in *AWAR* to cue all audio and video recordings to that point in the mission. This would eliminate the need for separate bookmarking and permit the assessor to focus on improving the quality of comments.

### 3.4.4 Recommendations

1. Assessment should involve the evaluation of performance on a pre-determined list of measures that covers all relevant aspects of performance, such at the ABM DCA team goal hierarchy.
2. Performance ratings will be more objective and consistent if measures refer to specific aspects of performance, are clearly worded and are unambiguous. Although the ABM DCA team goal hierarchy offers an improvement on the current RAAF assessment approach, further refinement is required. Further research will be required to inform such improvements.
3. Structuring training scenarios around the elements of the team goal hierarchy would promote consistency between teams in terms of their exposure to events and resultant learning.
4. Using the same scenarios for all students at a particular course level would permit reliable performance comparisons to be made between students.
5. Assessors must be sufficiently trained in the use of tools such as *AWAR* and the arrangement of elements in the incorporated hierarchy before using them to assess performance.
6. Electronic databases of student performance data may draw attention to tasks that students find particularly difficult and training may be modified to address this. This and other potential uses of electronic performance data should be considered.
7. Tools like *AWAR* allow performance assessment data to be entered straight into an electronic database, circumventing the double handling that currently occurs in the RAAF when paper and pencil performance ratings are transferred into a database.
8. A function should be incorporated into performance assessment tools that permits scores to be exported to a spreadsheet and printed in a user-specified order. This performance report should include three to five good aspects of their performance and three to five suggestions for improvement and should be given to students prior to the debrief.
9. The benefits and drawbacks of creating a process for students to use a performance report and audio and video replays of a mission to debrief themselves should be considered.
10. Lightweight tablet PCs with handwriting recognition should be considered to improve the portability of software tools such as *AWAR*. Handwriting recognition software should permit assessors to store commonly used symbols, shorthand and abbreviations.
11. The *AWAR* spreadsheet should provide more space for assessor comments and should permit freehand diagrams to support assessor comments.

## 3.5 AAR approach and tools

### 3.5.1 Collecting data for an AAR

The ABM task assessor was tasked with providing team task assessments via the *AWAR* tool and with bookmarking key scenario events via the *TAARDIS* Console application. He found it extremely difficult to conduct both of these tasks simultaneously due to workload, and suggested that the tasks be allocated to two separate task assessors. However, as the key scenario events may also represent evidence for assessments, these tasks may be more effectively and efficiently completed by one person. An alternative is for assessors to create timestamped bookmarks in *AWAR* and to export these to other applications (such as *TAARDIS*). Import and export functionality must be set up to support this capability.

The task assessor felt that monitoring the communication of specific team members would improve the quality of his team performance ratings. Unfortunately, due to a limitation of the task assessor's headset, the relative volume of the communications system was less than the general noise level of the room. On occasion he needed to leave his workstation and stand behind the team member he wished to listen to. This meant that the task assessor was not always able to enter comments into either *AWAR* or *TAARDIS* at this time, and had to update these upon his return. This was likely to have reduced the clarity with which the assessor remembered important aspects of the mission and increased the chance that some would be forgotten entirely. In order to improve flexibility, future exercises should investigate the utility of completing *AWAR* assessments and creating bookmarks using a tablet PC.

### 3.5.2 Preparing for an AAR

AAR sessions should occur as soon as possible after the completion of each vignette in order to promote the training benefit of feedback. In EBS08, some measures required the evaluation of performance across the entire vignette and ratings were therefore finalised after the completion of vignettes. This generally delayed the AAR session by a few minutes. The AAR session was also delayed by the need for technical staff to synchronise the *Pensieve* videos and Console ground truth recording, and to test the levels of the replay audio. Although the delay was not considered critically long, it was unnecessary and the tools should be further modified to automate this synchronisation. The task assessor felt that the AAR tools, particularly the bookmarking feature, significantly reduced the preparation time required prior to the provision of feedback.

### 3.5.3 Conducting an AAR

The task assessor felt that the AAR tools facilitated the identification of key points and permitted the displaying of concise examples of specific behaviours. The task assessor typically selected the ground truth and WD views for projected display during the AAR. The WD view was chosen because the view selected by the WD was typically wider than that selected by the other team members. ABM team members stated a preference for short mission reviews that identify and emphasise key learning points. They felt that the replays of the scenario ground truth, the individual console video and associated communication recordings were an effective and efficient means of illustrating learning points. The synchronisation between recordings and ability to cue to relevant bookmarks increased the efficiency of AAR sessions. However, there was a synchronisation problem associated with seeking to a bookmark. The *Pensieve* video would seek and begin playback immediately, whilst the *WorldView* map display would show the history trail of entities in their ‘pre-seek’ position and gradually update the position of entities over the following five seconds. This distracted the training audience and interrupted the flow of the AAR and must be fixed for future exercises.

The task assessor believed that the AAR tools created an effective structure for measuring performance, collecting evidence and understanding the decision making processes of the team over time. However, he stated that his leadership of the AAR was affected by the large number of tools at his disposal and his familiarity with each of them. On some occasions this led him to make use of a tool or piece of information that was not suited to the purpose. He also believed that he could have provided more targeted feedback if he had been aware of the affordances of each tool at different phases of the training cycle. This issue may be resolved in future exercises if the assessor is trained in the use of the AAR tools prior to the commencement of the exercise.

The task assessor indicated that the requirements of an AAR, in terms of the primary data source to be displayed, varied across individual and team training contexts. Participants believed that the *Pensieve* console recordings were an excellent enhancement to the training of individual controllers, but in a team training context, they would prefer to display a Solipsys recording rather than a specific console video. This would have provided an interactive view of team awareness, that is, the ability to pan and scan the map display and ignore individual user interactions with the interface. It is not yet possible to synchronise the native Solipsys recording or playback with *TAARDIS* Console ground truth logging and methods of achieving this should therefore be investigated.

### 3.5.4 Suggested AAR tool enhancements

The task assessor made some further suggestions on how the AAR tools might be enhanced. The Solipsys playback functionality could be extended to include accompanying selectable audio recordings taken from the operator stations. This should enable switching between the video and audio recordings as appropriate. This may be added to a multi-station split screen recording display, which would be useful when training a team of four or five ABM’s working together on different aspects of the same mission. The task assessor felt that *TAARDIS* and *ADGESIM* could be adapted to support this capability.

### 3.5.5 Recommendations

1. The task assessor found it difficult to concurrently rate performance using *AWAR* and create timestamped bookmarks using *TAARDIS*. The efficiency of this process would be improved if task assessors were asked to complete both of these tasks using *AWAR*.
2. Additional functionality should be incorporated into the *AWAR* tool to permit timestamped bookmarked to be exported to other applications such as *TAARDIS*.
3. The *TAARDIS* suite of tools should be further modified to automate the synchronisation of multiple video and audio recordings.

4. Assessors should be adequately trained in the functionality of all tools and procedures prior to the commencement of an exercise.
5. Technical Staff should investigate methods of synchronising native Solipsys recording or playback with *TAARDIS* Console ground truth logging.
6. Technical Staff should investigate the utility of extending the Solipsys playback functionality to include accompanying selectable audio only recordings taken from the operator stations.

# 4. Conclusions

EBS08 had two broad goals. These were, firstly, to examine the benefits providing a RAAF ABM team with a synthetic training environment within which to prepare for their tactical command and control role in PB08, and secondly, to provide an environment within which a number of new simulation tools, decision aids, performance measurement tools and training techniques could be evaluated and developed for future implementation within the RAAF. The focus of the current report has been on the latter, with the former being addressed in a companion report (Shanahan et al., in preparation). The exercise provided a suitable environment within which to evaluate an evolved environment for RAAF collective training composed of the *ADGESIM* suite of tools, the *Toteboard*, AS JETs-based performance assessment and the associated *AWAR* tool, and the *TDT* framework. The body of this report discusses these tools and techniques in terms of their potential for use in a RAAF training context and suggests methods by which they may be improved.

The version of *ADGESIM* trialled at EBS08 was considered a significant improvement over the version currently in use at SACTU and SAW. The RAAF participants were impressed with the potential of the new version of the *ADGESIM* system to run effectively with fewer people than the previous version. However, due to time constraints, the new version of *ADGESIM* was not thoroughly tested prior to the commencement of EBS08, and a number of glitches arose, some of which disrupted the exercise. In addition, the SIMOPs felt they did not have enough time at the start of the exercise to become familiar with the latest version of *ADGESIM*, and therefore did not understand all of its functionality. All efforts will be made to provide sufficient time prior to the commencement of future exercises for thorough testing of the simulation system and operator training.

The ABM team felt that the *Toteboard* provided all of the information they required during mission execution and provided the additional benefit of facilitating coordination between the tactical control centre and the White Force control centre. The RAAF participants offered suggestions for improvement, which have since been incorporated.

The ABM team reported that all four of the *TDT* dimensions were familiar constructs that were reflected in their current performance assessment but were more implicitly than explicitly trained. They felt that the *TDT* process, as experienced at EBS08, added little to the development of teamwork behaviours within the ABM team. There is considerable evidence to suggest that training that focuses on team processes can lead to improved team performance (Kozlowski & Ilgen, 2006). There is also considerable support for the *TDT* framework, as an effective means by which to train teams to work together better, at least within the context of the US Navy (Smith-Jentsch, et al., 1998). During EBS08, *TDT* was only partially implemented and therefore strong conclusions regarding the utility of the approach cannot be drawn.

The *AWAR* tool, incorporating the ABM DCA team goal hierarchy offered a marked improvement on current RAAF methodology and tools, in terms of the breadth, depth and consistency of ratings. Participants felt that the team goal hierarchy would also be useful for designing basic course scenarios at SACTU and SAW. The ABM team found the visual layout of the debrief user interface to be intuitive. The scores and comments relating to each sub-goal were considered extremely useful, as they provided the team with an understanding of their current performance, showed how much room existed for improvement, and provided instructions for improving.

The RAAF participants found the AAR tools very effective in minimising the amount of preparation required between the end of a mission and the beginning of the task performance debrief. The task assessor also felt that the AAR tools improved the time efficiency of the debrief process itself by facilitating the brevity and sharpness of learning points. The playback of video and audio was considered an extremely useful method of illustrating important learning points.

Overall, the participants found EBS08 to be an extremely valuable exercise and were satisfied with the quality of the preparation they received for PB08. The simulation tools and training techniques evaluated in the exercise assisted in promoting the quality of this experience and prompted participants to consider methods by which the current approach of the RAAF to training could be improved. Participants provided extensive feedback on each tool and while some shortcomings were highlighted, valuable suggestions for improvement were obtained. This feedback will guide the further development of these tools for implementation in future research exercises and in RAAF training programs.

# 5. References

*1994 Black Hawk shootdown incident*. (n.d.). Retrieved April 21, 2008, from http://en.wikipedia.org/wiki/Black_Hawk_Incident

Best, C. & Burchat, E. (2006). *Evaluation of the 'Mentor' assessment and feedback system for Air Battle management team training.* DSTO Technical Report. DSTO-TR-1942.

Best, C., Hasenbosch, S., Skinner, M., Crane, P., Burchat, E., Finch, M., Gehr, S., Kam, C., Shanahan, C. & Zamba, M. (2007). *Exercise Pacific Link 2: Coalition distributed mission training research.* Proceedings of the Twelfth Australian International Aerospace Congress. Melbourne, Australia.

Hasenbosch, S., & Best, C. (2007). *A hierarchical analysis of air battle management team goals in the defensive counter air mission.* DSTO Technical Note. DSTO-TN-0781.

Kozlowski, S.W.J., & Ilgen, D.R. (2006). Enhancing the effectiveness of work groups and teams. *Psychological Science in the Public Interest, 7(3),* 77-124.

McCarthy, A., Kingston, G., Johns, K., Gori, R., Main, P. & Kruzins, E. (2003). *Joint warfare capability assessment - Final report: Australian Joint Essential Tasks volume 1.* DSTO Client Report. DSTO-CR-0293.

Shanahan, C., Best, C., Finch, M., Tracey, E., Vince, J., Hasenbosch, S. & Stott, A. (in preparation). *Exercise Black Skies 2008: Enhancing live training through virtual preparation. Part One: An evaluation of training effectiveness.* DSTO Technical Report. DSTO-TR-2307.

Smith-Jentsch, K.A., Zeisig, R., Acton, B., & McPherson, J. (1998). Team Dimensional Training: A strategy for guided team self-correction. In J.A. Cannon-Bowers & E. Salas (Eds), *Making decisions under stress: Implications for individual and team training.* (pp. 271-297). Washington: APA

Whitley, B.E. Jr. (1996). *Principles of research in Behavioural Science.* California: Mayfield Publishing Company.

# Appendix A: Continuum of simulated training events

| Event type | Purpose of training activity and relationship to 'live' experiences | Timing | Mission phases | Scenario management | Simulation models | Entity control | Scenario continuity | Fidelity |
|---|---|---|---|---|---|---|---|---|
| *Doctrine Development* | For the development of future, or the validation of current, tactics and procedures in emerging operational contexts | Not in real-time | Selected phases | Experienced 'game' controller required | Highly abstracted models, limited parameters | Course entity behaviour, scripted aggregate entities | Low | Low |
| *Skills Development* | To enhance the skills of specific categorisation(s), to apply the current tactics and procedures within generic operational contexts | Real-time | Selected phases | Distributed scenario management | Basic models, many parameters | Variable entity behaviour, many free play, experienced and rehearsed role players, one role player to one entity group, few scripted entities | Low | Medium |
| *Exercise Workup* | To familiarise specific sub-groups or organisations with a specific exercise environment, and refine plans and procedures for its conduct | Real-time | Most phases | Central scenario management | Representative models, many parameters | Variable entity behaviour, many free play, experienced and rehearsed role players, one role player to many entity groups, few scripted entities | High | High |
| *Mission Rehearsal* | To improve the performance of operators executing a prepared mission plan | Real-time | All phases | Central scenario management | Highly detailed models, extensive parameters | Fine entity behaviour, experienced and rehearsed role players, one role player to one entity group | Continuous | Very high |

Increasing system complexity and fidelity

# Appendix B: Summary of recommendations

## B.1. *ADGESIM* simulation suite

1. All simulator-related technical briefs should be clear and contain the appropriate level of detail.

2. SIMOPs should receive adequate training in the use of all software and simulation systems before the commencement of an exercise. The SIMOP team suggested that two days of focused training in the week prior to EBS08 would have permitted them to become sufficiently familiar with the new versions of *PSI* and *DISVOX*.

3. Adequate time should be reserved in the weeks preceding an exercise for a thorough and systematic testing of *ADGESIM* systems. All technical issues must be resolved prior to the exercise with no additions or modifications made after testing.

4. SIMOP workload will be considerably reduced if entities are assigned by role rather than location. White Force scenario managers should endeavour to achieve this wherever possible.

5. Further workload reduction can be achieved by increasing the number of SIMOPs assigned to the White Force. However, all qualified and partially qualified RAAF SIMOPs were present at EBS08 and the pool of available SIMOPs must therefore be increased.

6. The *ADGESIM* simulation suite should be further automated with the goal of further reducing the number of SIMOPs required to produce a complex scenario. However, modifications must be sufficiently tested to ensure that they are not made at the expense of realism.

7. The ABM team was able to hear the activity of the SIMOPs and adjusted their communications accordingly. Ideally, SIMOPs and ABM teams should operate in completely separate rooms rather than partitioned sections of the same room.

## B.2. Toteboard

1. The most recent version of the *Toteboard* should be used for all future ABM simulation exercises.

## B.3. *TDT* framework

1. If a *TDT* introductory exercise is used to train a team to recognise instances of the four *TDT* elements, care should be taken when selecting the scenario description to ensure that it bears relevance to the context within which the team usually works. The description should include realistic examples of good and bad behaviour as well as extreme examples included to maintain interest. Learning to recognise these subtle examples will better equip the team to recognise such instances in their own performance.

2. The scenario workload must be high enough to genuinely test teamwork and communication skills.

3. If *TDT* is to be included in future exercises, assessment should be completed by four *TDT* assessors, with one assessing performance under each *TDT* category. If this is not possible due to personnel constraints, the assessors that are involved should undergo extensive training in *TDT* assessment prior to the commencement of the exercise.

4. Future *TDT* assessors should undergo training that improves their understanding of the importance of guiding the team through a process of self-evaluation and correction, rather than providing direct and unidirectional feedback, and their skills in encouraging team member participation.

5. The ABM team reported that the structure inherent to *TDT* did not compliment Australian military culture. Modifications aimed at improving the suitability of *TDT* within the context of the Australian military should therefore be considered.

## B.4. *AWAR* tool

1. Assessment should involve the evaluation of performance on a pre-determined list of measures that covers all relevant aspects of performance, such at the ABM DCA team goal hierarchy.

1. Performance ratings will be more objective and consistent if measures refer to specific aspects of performance, are clearly worded and are unambiguous. Although the ABM DCA team goal hierarchy offers an improvement on the current RAAF assessment approach, further refinement is required. Further research will be required to inform such improvements.

2. Structuring training scenarios around the elements of the team goal hierarchy would promote consistency between teams in terms of their exposure to events and resultant learning.

3. Using the same scenarios for all students at a particular course level would permit reliable performance comparisons to be made between students.

4. Assessors must be sufficiently trained in the use of tools such as *AWAR* and the arrangement of elements in the incorporated hierarchy before using them to assess performance.

5. Electronic databases of student performance data may draw attention to tasks that students find particularly difficult and training may be modified to address this. This and other potential uses of electronic performance data should be considered.

6. Tools like AWAR allow performance assessment data to be entered straight into an electronic database, circumventing the double handling that currently occurs in the RAAF when paper and pencil performance ratings are transferred into a database.

7. A function should be incorporated into performance assessment tools that permits scores to be exported to a spreadsheet and printed in a user-specified order. This performance report should include three to five good aspects of their performance and three to five suggestions for improvement and should be given to students prior to the debrief.

8. The benefits and drawbacks of creating a process for students to use a performance report and audio and video replays of a mission to debrief themselves should be considered.

9. Lightweight tablet PCs with handwriting recognition should be considered to improve the portability of software tools such as *AWAR*. Handwriting recognition software should permit assessors to store commonly used symbols, shorthand and abbreviations.

10. The *AWAR* spreadsheet should provide more space for assessor comments and should permit freehand diagrams to support assessor comments.

## B.5. The AAR tools

1. The task assessor found it difficult to concurrently rate performance using *AWAR* and create timestamped bookmarks using *TAARDIS*. The efficiency of this process would be improved if task assessors were asked to complete both of these tasks using *AWAR*.

2. Additional functionality should be incorporated into the *AWAR* tool to permit timestamped bookmarked to be exported to other applications such as *TAARDIS*.

3. The *TAARDIS* suite of tools should be further modified to automate the synchronisation of multiple video and audio recordings.

4. Assessors should be adequately trained in the functionality of all tools and procedures prior to the commencement of an exercise.

5. Technical Staff should investigate methods of synchronising native Solipsys recording or playback with *TAARDIS* Console ground truth logging.

6. Technical Staff should investigate the utility of extending the Solipsys playback functionality to include accompanying selectable audio only recordings taken from the operator stations.

# Appendix C: Semi-structured interview questions

## C.1. ABM instructor (task work assessor) questions:

### C.1.1 The AAR tools

- How straightforward and timely was the process of selecting, seeking, and synchronising AAR data sources?
- How much time do you need to prepare the data sources for the AAR?
    - Did you find this to be inhibitive? Annoying? Justified?
- Did you have all of the information you needed to make the intended teaching points clearly and effectively?
    - Did you require any information that wasn't available?
- Was the information you used during AARs easily accessible?
- Given no restrictions, how would you design the ideal data enhanced / supported AAR?
    - How many screens would you have? What would you display, etc.?
- Do you feel that the information available to you for use during AARs improved the ABM team's understanding of your comments?
- What is the impact on learning of providing students with objective performance data?
- If given the opportunity, would you use this array of information in SACTU simulation exercises?
- Is an AAR of this type equally suited to both individual and team training?
- Do you have any further comments on the use of additional visual data to support the AAR.

### C.1.2 The ABM DCA Team Goal Heirarchy

- Can you comment on the use of the ABM DCA hierarchy as a tool for scenario design?
- Do you feel that use of the hierarchy (or something similar) would add anything to the current RAAF approach to designing scenarios?
- Are there any drawbacks or benefits of designing scenarios such that performance on such low level tasks can be assessed?
- Do you value the current approach of granting SMEs the flexibility to inject additional training experiences to scenarios in real time?
- Would it be more useful to assess all students in a course using a common scenario?
- Can you comment on the usefulness of the hierarchy for performance assessment?
- Do you feel that the use of this hierarchy (or something similar) would add anything to the current RAAF approach to assessing performance?
- Does the team goal hierarchy promote the objectivity of performance assessment?
- Does the hierarchy cover all relevant areas?
    - Are there any tasks that are important to assess that are missing?
- Did the hierarchy prompt you to assess performance in areas that you may have otherwise forgotten?

- Does use of the hierarchy permit a valid assessment of performance?
- Does the use of *AWAR* populated with the ABM DCA hierarchy facilitate a more objective measure of performance than is currently used at SACTU?
    - Is this beneficial?
- Did use of the hierarchy make the process of assessment unnecessarily complex?
- Are there any benefits or drawbacks to being assessing a team of students as opposed to an individual?
- Can you comment on the usefulness of the hierarchy for performance feedback?
- Do you feel that the use of this hierarchy (or something similar) would add anything to the current RAAF approach to providing performance feedback?
- Did the students find it useful during the AAR to organise the discussion of their performance around the ABM task hierarchy?
- Can you comment any further on the use of the ABM DCA hierarchy for scenario design, performance assessment and performance feedback?

### C.1.3 The *AWAR* tool

- Are there any benefits or drawbacks of using the *AWAR* tool to assess team performance rather than the traditional paper and pencil assessment method employed at SACTU.
- How did you find the organisation of measures?
    - Could you find the required measures quickly and easily?
- How did you find the scoring system?
- Are there any physical limitations that relate to the use of *AWAR*?
    - Are the restrictions *AWAR* poses on mobility a problem?
- What do you feel the use of *AWAR* would add or detract from assessment if it were to replace current assessment techniques at SACTU?
- Does *AWAR* provide a good structure for the AAR?
- Was it useful to have performance data available in an electronic format immediately after the missions?
- How did the ABM team find the organisation of the debrief user interface?
    - Was it intuitive?
- Did the students find the aggregation and colour coding of the debrief user interface useful?
- Did the ABM team find the graph of average performance on the ASJETs useful for understanding their performance and areas for improvement?
- If given the opportunity would you be happy to replace the assessment methods currently employed at SACTU with the *AWAR* technique?
- Can you make any further comments on the use of *AWAR* for performance assessment and feedback?

## C.2. ABM team

### C.2.1 The AAR tools

- Does an AAR that is supported by these different sources of data add anything to the debrief approach currently employed by the RAAF?
    - What additional information can be obtained from these data sources?

- Are there any disadvantages to utilising these different data sources in debriefs when compared to the approach currently employed by the RAAF?
    - Are there any features of your normal debrief techniques that are *better* than the setup we have here at Black Skies?
    - Were there any instances where the data display was more distracting than informative?
- Did the data displayed in the AAR increase your understanding of the assessor's feedback about your performance?
    - How?
- Was there any information not displayed that might have provided useful feedback on your performance?
    - Would you like to have seen statistics such as shot logs?
- If there were no restrictions, how would you design the ideal data display to support AAR?
    - How many screens would you have?
    - What information would you display?
- How easy was it to interpret the data presented in the AAR, for example the source ID, size, clarity, relative positioning of information?
- Do you have any more comments on the utility of an AAR supported by these different sources of data?

### C.2.2 The ABM DCA Team Goal Heirarchy

- In what ways does using a team goal hierarchy differ from the way in which your performance is usually assessed?
    - Does it add to the current RAAF approach? How?
    - Does it detract from the current approach? How?
- Do you think that employing a team goal hierarchy is a useful way to structure a team performance debrief?
    - What are the advantages of organising the AAR in this manner?
    - What are the disadvantages of having such a tightly-focused AAR?
- How effectively did the hierarchy of training objectives capture all the tasks you were required to perform as a team during the EBS08 missions?
    - What aspects of your task were not captured by the goal hierarchy?
- The Team Goal Hierarchy contains very specific, low-level criteria against which performance is rated:
    - What are the benefits associated with having such specific criteria?
    - In what ways could this specificity of performance criteria detract from the feedback you receive from the assessor?
- Does using a team goal hierarchy result in a more objective assessment of your performance than the methods that are currently used?
- Do you think that using a team goal hierarchy is more likely is more likely to result in a true representation of your performance than the methods that are currently used?
- What are the benefits of being assessed as a team rather than as an individual?
- What are the drawbacks to being assessed as a team?
- Would there be any benefit to being able to compare the performance of individual team members across specific objectives?

- In what ways would this focus on the individual enhance the training benefit of the team overall?
- Within the Team Goal Hierarchy, there are very specific performance criteria.
    - Would you have preferred to have been given a more comprehensive description of these criteria prior to the first mission?
    - How would a better knowledge of these specific criteria have influenced the way you approached the mission?
- Using the team goal hierarchy allows the assessor to link specific mission tasks with high-level Air Force goals:
    - How useful is it for you to be able to see the relation between your specific duties and these high-level Air Force tactical goals?
    - Do you think that this approach to identifying training goals – associating specific tasks with high level tactical goals – could be implemented more widely in the Air Force?
- Do you have any further comments on the ABM DCA team goal hierarchy?

### C.2.3 The *AWAR* tool

- Team performance on each specific objective in the Team Goal Hierarchy was rated on a 5 point scale – from TERRIBLE to EXCELLENT. How effective were these ratings in providing a true representation of the performance of your team?
- What are the advantages of using *AWAR* populated with the Team Goal Hierarchy over the approach to performance assessment currently employed by the RAAF?
- What are the disadvantages of using *AWAR* populated with the Team Goal Hierarchy over the approach to performance assessment currently employed by the RAAF?
- Do you think that the *AWAR* tool provides a good structure for the AAR?
    - What did you like about *AWAR*?
    - What did you *not like* about *AWAR*?
- How easy was it to understand the organisation of the material displayed during the AAR?
    - What aspects of the material displayed were not easy to understand?
- *AWAR* permits performance data to be provided in an electronic format immediately following a training exercise or mission:
    - Is this desirable?
    - Does immediate feedback increase the training benefit of an exercise?
- In the *AWAR* system, the scores from lower-level measures were averaged to give a score for higher-level objectives. How valid are these aggregate scores as performance measures?
- The ratings of you performance were colour coded. Did this colour coding facilitate your interpretation of the performance feedback? How?
- In the AAR, a graph was presented which summarised the assessor's ratings in the six broad categories of training objectives:
    - Did this graph increase your understanding of the good and bad aspects of your performance? How?
    - Did the graph help you to understand the areas in which you need to improve?
- How could *AWAR* be modified to be more useful or effective as an AAR tool?

- Would you be happy for 3CRU to implement the *AWAR* system for training purposes?
- Do you have any further comments on *AWAR*?

### C.2.4 The *TDT* intervention

- How easy to understand were the four teamwork concepts (information exchange, communication delivery, supporting behaviour, and initiative/leadership)?
    - Which dimension of teamwork is most relevant to an ABM team?
- Were you already familiar with these teamwork concepts?
    - Does the Air Force promote these teamwork behaviours in a less formal way, or in different language?
- What are the advantages of having a performance review focused solely on teamwork processes?
    - Are there any drawbacks to having a team-process specific AAR?
- Do you think that there is a place for a more overt focus on teamwork processes in the training delivered by SACTU?
    - How do you think trainees at SACTU would respond to *TDT*-style performance reviews?
    - What are some of the possible learning benefits of introducing team-process training at SACTU?
    - What disadvantages might there be to introducing this kind of team-process training at SACTU?
- At the beginning of the week we gave you a vignette about the friendly-fire incident involving the US Blackhawks, and asked you to pick out instances of good and bad teamwork behaviours. How useful was this exercise in illustrating the *TDT* concept?
- Did the *TDT* briefing and Blackhawk exercise improve your ability to recognise good and bad teamwork behaviours during the mission you controlled today?
- What kind of impact, if any, did a greater awareness of team behaviours have on your mission performance?
- Following every EBS08 mission, you were debriefed on both your task work performance and teamwork performance using the *TDT* approach.
    - Did you find the summary of how well you worked together as a team useful?
    - What impact did this *TDT* summary have on your performance as a team in subsequent missions?
    - How well were the two separate debriefs integrated?
    - In what way could we improve the integration of these two types of debrief?
- What aspects of the *TDT* will you carry through to future exercises?
    - How do you think these aspects of *TDT* will improve your work performance?
    - What kind of impact do you think *TDT*-style assessment and feedback would have on your performance as members of an ABM team in a live control environment?
- Do you have any further comments on *TDT*?

### C.2.5 The *ADGESIM* suite of tools

- Did red and blue entities fly as you would expect in terms of acceleration, climb rate, turn rates, etc.?
- Did red and blue groups appear to hold formation?
- Were the weapons ranges of red and blue aircraft accurately represented?
    - Were they consistent, realistic, etc.?
    - Were there any problems with the accuracy of missile hits?
- What was the effect of any lack of realism that was experienced?
- Did you find comms between yourself and red aircraft pilots realistic in terms of radio quality (e.g., degradation with increased distance from radio towers)?
    - What were the effects of this?
- Did you experience any problems relating to the number of aircraft the sim operators were required to control, e.g., delays, unresponsiveness, different pilots with same voice...
- How realistic were the sensor inputs to Solypsis?
- Was the detection and tracking performance of the system realistic?
    - If not, what was odd?
    - What was the effect of this lack of realism?
- Do you have any further comments on *ADGESIM*?

## C.3. SIMOPs and WFMD

### C.3.1 The *ADGESIM* suite of tools

- How many formations have you been controlling over the course of this exercise (range, average)?
- What level of workload have you experienced at EBS08?
    - Has your workload changed over the course of the week? How? Why?
- What problems resulted from your high level of workload?
    - What effect did these problems have?
- Did red and blue entities fly as you would expect in terms of acceleration, climb rate, turn rates, etc.?
    - Did red and blue groups appear to hold formation?
    - Were the weapons ranges of red and blue aircraft accurately represented? Were they consistent, realistic, etc.?
    - Were there any problems with the accuracy of missile hits?
    - Do you have any further comments on realism?
- How did you find the organisation of the information you had available to you?
    - Did you have access to all of the information you needed?
    - Did you find it easy to access the right information at the right time?
- How did you find the presentation of information?
    - Was the information organised such that you were able to transmit information from the sim interface to the ABM team over the radio with minimal translation?
- The *ADGESIM* developers attempted to organise information in an intuitive way so that you had spare cognitive capacity for higher level tasks, e.g., thinking ahead?
    - Did this work as they intended?

  - Do you feel that you were able to control more aircraft as a result of the organisation of information?
  - Do you have any further comments on information management?
- How did you find the general layout of the display?
  - Did you find the colour coding of fuel states useful?
  - Did you find the colour by IFF feature useful?
  - Did you find the callsign colour coding useful?
  - Do you feel that all relevant information was accessible?
  - How did you find the filtering of information?
- Do you have any further comments on ADEGESIM?

# Appendix D: *TDT* introductory vignette

**Instructions:** Read through this true story of the shooting down of two US Army helicopters by friendly fire in Iraq and identify examples, either positive or negative, of behaviours related to the four teamwork dimensions of INFORMATION EXCHANGE, COMMUNICATION DELIVERY, SUPPORTING BEHAVIOUR, and LEADERSHIP / INITIATIVE.

On April 14, 1994, at 07:36, a USAF E-3 AWACS aircraft from the 963rd Airborne Air Control Squadron (based at Tinker Air Force Base, Oklahoma) departed Incirlik Air Base (AB), Turkey as the lead aircraft of 52 coalition air missions scheduled for that day in support of Operation Provide Comfort (OPC). The AWACS, under the command of Major Lawrence Tracy, was to provide airborne threat warning and air control for all OPC aircraft during its time aloft. The AWACS crew reported on station at its assigned surveillance orbit located inside Turkey just north of the northern border of Iraq at 08:45. The weather that day was fair and clear over northern Iraq.

At 08:22, two U.S. Army UH-60 Blackhawk helicopters from the 6th Battalion, 159th Aviation Regiment, called Eagle Flight, departed Pirinçlik, near Diyarbakır, Turkey headed for the OPC military coordination center (MCC) located in Zakhu, Iraq. Both helicopters were fitted with external, 230-gallon fuel sponsons mounted below each side door with each sponson emblazoned with large American flags. In addition to the flags on the sponsons, each helicopter was marked with American flags on each side door, on the nose, and on the belly.

At 09:21, the Blackhawks reported their entry into the no fly zone in Northern Iraq by radio on the en-route frequency to the AWACS en-route controller, Lieutenant Joseph Halcli, and then landed six minutes later at the MCC. Halcli and his superior officer, Captain Jim Wang, the AWACS crew's Senior Director, added "friendly helicopter" symbology to their radar scopes, noted that both helicopters were displaying Identification friend or foe (IFF) signals, and then suspended the symbology after the Blackhawks disappeared from their scopes upon landing at the MCC at 09:24. Although the helicopters were squawking the wrong IFF code for the no fly zone (called the Tactical Area of Responsibility or TAOR), neither Wang nor Halcli informed them of that. Wang and Halcli also neglected to direct the Blackhawks to begin using the TAOR radio frequency instead of the en-route frequency.

At the MCC, the Blackhawks picked up 16 members of the UN Provide Comfort coalition leadership team including five Kurdish civilians, three Turkish, two British, and one French military officers, plus five U.S. civilians and military officials. At 09:54, the helicopters departed from the MCC for Arbil, Iraq. The Blackhawks reported their departure, flight route, and destinations by radio which was acknowledged by Halcli. Halcli reinitiated the friendly helicopter track on his scope. Halcli placed symbology on his radar screen to show the two Blackhawk tracks and notified Wang of the helicopters' movement. In addition to Halcli's screen, the friendly helicopter symbology was visible on the radar screens of Wang, Tracy, and USAF Major Doug Martin. Martin was the "Duke" or airborne command element on the

AWACS, meaning that he was a rated pilot assigned to the crew to ensure that all engagement (combat) mandates were adhered to and executed as written.

Enroute to Arbil, at 10:12, the Blackhawks entered mountainous terrain and their radar returns disappeared from the AWACS' scopes. Captain Dierdre Bell, an air surveillance officer on the AWACS, noticed that the Blackhawks' radar returns had disappeared and sent an electronic "attention arrow" to Wang's scope. Wang took no action and the large blinking green arrow automatically disappeared from his screen after one minute.

Meanwhile, at 09:35, two USAF F-15C fighter aircraft from the 53rd Fighter Squadron, piloted by Captain Eric Wickson and Lieutenant Colonel Randy W. May, departed Incirlik AB. Their mission was to perform an initial fighter sweep of the TAOR to clear the area of any hostile aircraft prior to the entry of coalition forces. The air tasking order (ATO) that was supposed to list all scheduled coalition aircraft missions for that day and which the two pilots reviewed before takeoff, mentioned that there would be U.S. Army Blackhawk helicopters operating in the TAOR that day but didn't list takeoff times for them. At 10:15, Wickson radioed Martin on the AWACS and asked if he had any information to pass to them. Martin replied in the negative.

At 10:20 Wickson, the F-15C flight lead, reported entering northern Iraq to the AWACS controller responsible for air traffic inside the TAOR, USAF Lieutenant Ricky Wilson. The TAOR frequency that the F-15s were using was different than the en-route frequency being used by the two Blackhawks. Wilson, however, was monitoring both frequencies as well as being able to see both Blackhawks on his radar scope before they disappeared at 10:12. Wilson and the other AWACS crewmembers, many of whom were monitoring the F-15s radio frequency, did not inform the F-15s that Blackhawks were currently operating in the TAOR. At 10:21, Wilson, believing that the Blackhawks had landed again, asked Wang if he could drop the friendly helicopter symbology from the AWACS' scopes and Wang approved the request. An AWACS crew instructor, Captain Mark Cathy, who was on the mission to assist the AWACS crew and supervise Wilson on this, his first mission into the TAOR, had retired to the back of the airplane at 10:00 to take a nap.

At 10:22, Wickson reported a radar contact on a low-flying, slow-moving aircraft 40 miles (64 km) southeast of his current position. Wilson acknowledged Wickson's report with a "clear there" response, meaning that he had no radar contacts in that area. Unbeknownst to the two F-15 pilots, the unidentified aircraft were the two U.S. Army Blackhawks. Neither Tracy or Wang spoke up at this point to request that AWACS crewmembers attempt to identify the F-15s radar contacts.

Both F-15 pilots then electronically interrogated the radar target with their on-board IFF systems across two different frequencies. Their IFF systems responded negatively to the attempt to identify the contact and the F-15s moved to intercept the unidentified aircraft. Intermittent IFF returns from the Blackhawks now began to show on Wilson's and other AWACS crewmember's scopes and friendly helicopter symbology reappeared on Wang's scope. After closing to 20 miles (32 km) of the radar contacts, at 10:25 the F-15s again reported the contact to the AWACS and Wilson this time responded that he now had a radar contact at that reported location. Although the Blackhawk intermittent radar and now steady IFF returns on the AWACS scopes were in the same location as the unidentified contacts being tracked by the F-15s, none of the AWACS controllers advised Wickson or May that the contacts that they were tracking might be friendly helicopters.

The two F-15s now initiated a visual identification (VID) pass of the contact. The VID pass entailed violating one of OPC's rules of engagement, which prohibited fighter aircraft from operating below 10,000 feet. At this time the two Army Blackhawks had entered a deep valley and were cruising at a speed of 130 knots about 200 feet (60 m) above the ground. Wickson's VID pass was conducted at a speed of about 450 knots (833 km / h), 500 feet (150 m) above and 1,000 feet (300 m) to the left of the helicopters. At 10:28 Wickson reported *"Tally 2 Hinds"* and then passed by and to the front of the two blackhawks. "Hind" is the North Atlantic Treaty Organisation (NATO) designation for the Mil Mi-24 helicopter, a helicopter that the Iraqi, Turkish, and Syrian militaries operated and was usually configured with armament on small, side-mounted wings. Wilson responded with *"Copy, Hinds"* and asked Wang, "*Sir, are you listening to this?"* Wang responded, *"Affirmative"* but offered no further guidance or comments.

May then conducted his own VID pass about 1,500 feet (500 m) above the helicopters and reported, *"Tally 2"*. May later stated to a USAF accident investigation board that his "Tally 2" call meant that he saw two helicopters but didn't mean that he was confirming May's indentification of them as Iraqi Hinds. Neither F-15 pilot had been informed that U.S. Army Blackhawks participating in OPC often carried auxiliary fuel tanks mounted on wings nor had either been instructed in the paint scheme that Iraqi Hind helicopters used, light brown and desert tan, which was different than the dark brown / green color used by the Blackhawks. Wickson later stated that, *"I had no doubt when I looked at him that he was a Hind...The Blackhawk did not even cross my mind."*

Following their VID passes, Wickson and May circled back behind the helicopters approximately 10 miles (16 km). At 10:28, Wickson notified the AWACS that he and May were *"engaged"* and instructed May to *"arm hot"*. At 10:30, Wickson fired an AIM-120 AMRAAM missile at the trail helicopter from a range of about four nautical miles. The missile hit and destroyed the trailing helicopter seconds later. In response, the lead Blackhawk immediately turned and dived for lower altitude in an apparent attempt to evade the unexpected attack. Shortly thereafter, May fired an AIM-9 Sidewinder missile at the lead helicopter from a range of about one and a half nautical miles, hitting and shooting it down also. All 26 people on board the two Blackhawks were killed. After flying over the wreckage of the two Blackhawks lying burning on the ground, May radioed Wickson, *"Stick a fork in them, they're done"*.

“1994 Black Hawk shootdown incident” (n.d.).

**DEFENCE SCIENCE AND TECHNOLOGY ORGANISATION**
**DOCUMENT CONTROL DATA**

1. PRIVACY MARKING/CAVEAT (OF DOCUMENT)

| 2. TITLE | 3. SECURITY CLASSIFICATION (FOR UNCLASSIFIED REPORTS THAT ARE LIMITED RELEASE USE (L) NEXT TO DOCUMENT CLASSIFICATION) |
|---|---|
| Exercise Black Skies 2008: Enhancing Live Training Through Virtual Preparation<br><br>Part Two: An Evaluation of Tools and Techniques | Document (U)<br>Title (U)<br>Abstract (U) |

| 4. AUTHOR(S) | 5. CORPORATE AUTHOR |
|---|---|
| Eleanore Tracey, Sam Hasenbosch, Julian Vince, Daniel Pope, Aaron Stott, Christopher Best, Christopher Shanahan and Melanie Finch | DSTO Defence Science and Technology Organisation<br>506 Lorimer St<br>Fishermans Bend Victoria 3207 Australia |

| 6a. DSTO NUMBER | 6b. AR NUMBER | 6c. TYPE OF REPORT | 7. DOCUMENT DATE |
|---|---|---|---|
| DSTO-TR-2305 | AR-014-564 | Technical Report | June 2009 |

| 8. FILE NUMBER | 9. TASK NUMBER | 10. TASK SPONSOR | 11. NO. OF PAGES | 12. NO. OF REFERENCES |
|---|---|---|---|---|
| 2009/1028074 | AIR 07/232 | DCAF | 55 | 9 |

| 13. URL on the World Wide Web | 14. RELEASE AUTHORITY |
|---|---|
| http://www.dsto.defence.gov.au/corporate/reports/DSTO-TR-2305.pdf | Chief, Air Operations Division |

15. SECONDARY RELEASE STATEMENT OF THIS DOCUMENT

*Approved for public release*

OVERSEAS ENQUIRIES OUTSIDE STATED LIMITATIONS SHOULD BE REFERRED THROUGH DOCUMENT EXCHANGE, PO BOX 1500, EDINBURGH, SA 5111

16. DELIBERATE ANNOUNCEMENT

No Limitations

17. CITATION IN OTHER DOCUMENTS Yes

18. DSTO RESEARCH LIBRARY THESAURUS http://web-vic.dsto.defence.gov.au/workareas/library/resources/dsto_thesaurus.shtml

Team training, distributed mission training (DMT), virtual preparation, simulation systems, team performance assessment and review

19. ABSTRACT

Exercise Black Skies 2008 (EBS08) was a simulation exercise conducted in the weeks prior to the live air combat training exercise, Pitch Black 2008 (PB08). During EBS08, a Royal Australian Air Force (RAAF) air battle management (ABM) team conducted a series of vignettes designed to prepare them for their tactical command and control role in PB08. A broad goal of EBS08 was to provide an environment within which a number of new simulation tools and training techniques could be evaluated and developed for future implementation within the Royal Australian Air Force (RAAF). The tools and techniques evaluated were: the Air Defence Ground Environment Simulator (ADGESIM), Toteboard, Air Warfare Assessment and Review tool, After Action Review tools and the Team-Dimensional Training framework. These evaluations, which form the basis or this report, were informed by the observations of human factors scientists and interviews with the RAAF participants. Participants provided extensive feedback on each tool, highlighting shortcomings and providing valuable suggestions for improvement. This feedback will guide the further development of these tools for implementation in future research exercises and in RAAF training programs. Overall, the participants found EBS08 extremely valuable and were satisfied with the quality of training they received. The simulation tools and training techniques assisted in promoting the quality of this experience and prompted participants to consider methods by which the current RAAF approach to training could be improved.