



Model Driven Aviation Training Family of Systems Architecture Volume I

By

Trevor Holden, B.Eng., B.Eng (Hons), MScRes.

Date: October 2016

A Doctoral Thesis submitted in partial fulfilment of the requirements for the award of Doctor
of Philosophy of Loughborough University

© by Trevor Holden 2016

PREFACE

“If you make people think they’re thinking, they’ll love you; but if you really make them think, they’ll hate you”
- Harlan Jay Ellison

This thesis is a record of my research efforts over the past four years; it is in essence the end result of a number of assault courses to which I have personally had to run through to reach a conclusion that is this thesis. My time at the university has included teaching both academic and industrial students which has led to interesting outcomes; and the realisation that it is not easy completing a project with substantial distractions. I have been popular and unpopular depending on the circumstances and sometimes I have had to draw a line in the sand. All the days at the university have been an experience (some good, some bad) and have added to the knowledge and memories I gathered from my time in industry.

The realisation of the disconnect / gap between academia and industry served as the basis for the production of this thesis. The industrial sponsor became the largest of the obstacles to overcome with the constant change of management and PhD research direction; nonetheless a non-fictional and hopefully rigorous story has been written. Through hard work (no exaggeration) and patience the research has hopefully opened a new direction and thought processes to the management of a flight training system and with it the governance and configuration control of the advanced technology being used within complex systems.

The end point of the research was decided by time and the sand; the output is by no means the end to the research, but a new beginning and new way of thinking. Systems Engineering is difficult to complete as an individual project, but like the quote by Robert Orben, should lead to success. This notwithstanding the journey has been an experience that I will take with me wherever the future leads and the knowledge gained I will pass on and hopefully start someone else’s journey through systems engineering issues and capabilities.

“A Graduation Ceremony is an event where the commencement speaker tells thousands of students’ dressed in identical caps and gown that ‘individuality’ is the key to success.”
- Robert Orben

Trevor Holden

October 2016

EXECUTIVE SUMMARY

The PhD project has evolved from focusing on the technical problem of the integration and interoperability of an assemblage of complex systems and SoS within a flight training system to development of a workflow process using frameworks to aid the decision making process for the selection of optimal flight training blending mixes. The focus of the research involved developing a methodology to satisfy research project proposal requirements agreed upon with the industrial sponsor. This thesis investigates the complexity of a modern flight training systems and the need for understanding that it is supported by a complex Family of Systems (FoS) including Virtual Reality Training Environments such as flight simulators, to live training aircraft with various configurations of avionic controls. One of the key technical problems today is how best to develop and assemble a family of flight training system into an integrated Live / Synthetic mix for aircrew training to optimise organisation and training objectives.

With the increased use of emulation/synthetic data on aircraft for live training, the synthetic boundary is becoming increasingly blurred. Systematic consideration of the most appropriate blend is needed. The methodology used in the research is model driven and the architecture produced is described at a level of abstraction to enable communication to all stakeholders for the means of understanding the structure involved in the system design process. Relational Oriented Systems Engineering and Technology Trade-Off Analysis (ROSETTA) frameworks are described using Model Based Systems Engineering (MBSE) techniques for supporting capability based trade-off decisions for selection of optimal flight training FoS mixes dependent on capability. The research proposes a methodology and associated methods including a high-level systematic closed loop information management structure for blended device/tool aircrew training and a modelling and analysis approach for the FoS aviation training problem to enhance the existing training programmes to provide a more efficient and agile training environment. The mathematical formalisms used provide a method of quantifying subjective opinions and judgements for trade studies to be accomplished on the suitability of technology for each student pilot in relation to training and organisational objectives. The methodology presented is by no means a final solution, but a path for further research to enable a greater understanding of the suitability of training tools / technology used to train individual pilots at various stages throughout the training pipeline lifecycle(s).

TABLE OF CONTENTS

Preface.....	II
Executive Summary	III
Table of Contents	IV
List of Figures	IX
List of Tables	XIII
Glossary	XIV
Acknowledgements.....	XV
Authors Declaration	XVI
Chapter 1 Introduction & Background	1
1.1 Why is Research Needed Within the Flight Training Domain	3
1.2 Statement of Problem	5
1.2.1 Research Enterprise Viewpoint Needs.....	8
1.3 Research Methodology.....	9
1.3.1 Solving Multi-Attribute Problems	9
1.3.2 Methodology Direction.....	11
1.4 Research Aims / Objectives	11
1.4.1 Research Deliverables.....	14
1.5 Scope of Research	15
1.6 Research Novelty	17
1.7 Structure of Thesis	18
Chapter 2 Wider Systems Engineering Context	21
2.1 Theory of Systems Design	23
2.2 What is a Model?.....	25
2.3 Model Ontologies.....	27
2.4 Complex Integrated Systems.....	29
2.4.1 Model Driven Architecture.	31
2.4.2 Brief Summary of MBSE Constraints In Proposed Methodology.....	33
2.5 Representing FTS Systems with Models	34
2.5.1 Managing Complexity of the FTS SoS / FoS Problem.....	35

2.6	Chapter Summary and Relationship to Research Problem	39
	Chapter 3 Decision Making Techniques	41
3.1	The Evolution of Decision Support	42
3.1.1	The Decision Making Process	43
3.2	Decision Modelling of FTS	45
3.2.1	DSS Evaluation Methodology	48
	Chapter 4 Complexity of Modern & Future Flight Training Systems	51
4.1	Virtual Environments Critical Factors	53
4.2	Evaluating FTS Family of Systems	60
4.2.1	Pilot Education & Performance Considerations for Blending Mix Choice	63
4.3	Pilot Attributes in Relation to the Systems of the FoS	69
4.4	FTS Decision Making Strategies	75
4.4.1	Human Performance Factors Affecting Choice Of Blending Mix	77
4.4.2	FTS Effectiveness Measures	80
4.4.3	Transfer of Learning	82
4.5	FTS Current Trade-off Options	84
4.5.1	Understanding Trade-off Relationships	86
4.6	Chapter Summary	86
	Chapter 5 A ROSETTA Framework for Modelling and Simulation	88
5.1	Introduction to Relational Orientation	89
5.1.1	Why use a Specialisation of ROSETTA for the FTS	90
5.1.2	Modelling Differential Equations	92
5.2	FoS Model For Assessment Of Blending Mix	93
5.3	Analysis of Organisational Objectives	97
5.5	Chapter Summary	99
	Chapter 6 Methodology For FTS Design Process	101
6.1	Logical Modelling of FTS Requirements	103
6.2	Requirements Modelling	104
6.2.1	FTS PhD Requirements Modelling	108
6.3	Flight Training Enterprise	120
6.4	Use Case Development	123
6.4.1	Use Cases for the Decision Support System	125
6.5	The Readiness Metamodel	126

6.6	FTS Process Workflow Design for Methodology	130
6.7	Chapter Summary	133
Chapter 7 Architecture Approach to Process Workflow		135
7.1	Mission Planning and Baseline Simulation Model(s) (General Overview)	139
7.1.1	Architecting Flight Training Mission Scenarios	139
7.1.2	Training Attributes for Mission Planning Tasks	140
7.2	Pre-Pilot Assessment Implementation Design	142
7.2.1	Index of Learning Style (ILS) for FTS	144
7.2.2	Knowledge of Strengths and Weaknesses Of Students	145
7.2.3	ILS Evaluation Process	146
7.2.4	Personal Allocation Attention Assessment	149
7.2.5	Pre-Study Summary	150
7.3	Handling Qualities Workload Scale (HQWS) Model	152
7.3.1	Handling Qualities Workload Scale (HQWS) Implementation Design	153
7.3.2	HQWS Evaluation Process	155
7.3.3	HQWS Summary	157
7.4	ROSETTA Level 0 Model	158
7.4.1	ROSETTA 0 Implementation Design	159
7.4.2	ROSETTA 0 Evaluation Process	161
7.4.3	ROSETTA 0 Summary	163
7.5	Pilot Workload Evaluation	164
7.5.1	Instructor Evaluation Of Pilot Workload with Scenario Tasks	165
7.5.2	Workload Implementation With ROSETTA Design	166
7.5.3	Workload Summary	168
7.6	ROSETTA Level 1 Model	169
7.6.1	K&S Relationship To MEC Identification For Technology Suitability	169
7.6.2	ROSETTA 1 Implementation Design	170
7.6.3	ROSETTA 1 Evaluation Process	179
7.6.4	ROSETTA 1 Summary	186
7.7	Pilot Awareness Rating Scale Model (PARS)	187
7.7.1	Student Understanding of Scenario, State of Mind and Success	189
7.7.2	PARS Implementation Design	189
7.7.3	PARS Evaluation Process	190

7.7.4	PARS Summary	191
7.8	ROSETTA Level 2 Model	192
7.8.1	Organisational Objectives to Fidelity Characteristics of Training Tool	194
7.8.2	ROSETTA 2 Implementation Design	194
7.8.3	ROSETTA 2 Evaluation Process	197
7.8.4	ROSETTA 2 Summary	200
7.9	Task Load Model	201
7.9.1	Task Load Implementation Design	201
7.9.2	Task Load Evaluation	202
7.9.3	Task Load Summary	204
7.10	Prediction of Task Performance with Chosen Training Tool	204
7.10.1	Performance Prediction Implementation Design	205
7.10.2	Performance Evaluation Process	207
7.10.3	Performance Summary	208
7.11	Pre-Flight SA Assessments	209
7.11.1	Pre-Flight Implementation Design and Evaluation	209
7.11.2	Pre-Flight SA Summary	210
7.12	Goal Modelling Techniques for FTS	211
7.12.1	Why Model to Estimate Success of Mission Objectives	213
7.12.2	Goal Model Implementation Design	214
7.12.3	Goal Modelling Evaluation Process	217
7.12.4	Goal Modelling Summary	220
7.13	Performance based Comparison	222
7.13.1	Performance Based Comparison Implementation design	223
7.13.2	Performance Base Comparison Evaluation Process	225
7.13.4	Performance Base Comparison Summary	229
7.14	After Execution Assessments	231
7.14.1	Observer-Rating Assessments	232
7.14.2	Post-Flight Self-Assessments	234
7.14.3	Pilot Flight Success Evaluation	236
7.14.4	Mission Operability Assessment	237
7.14.5	Paper Based Assessments	238
7.14.6	Simulator Configuration Assessments	243

7.14.7 After Execution Assessment Summary	245
7.15 Workflow Process Tool Summary	246
Chapter 8 Summary and Conclusion	249
8.1 Summary of Research Project	249
8.1.1 Summary of Satisfying Research Objectives	251
8.2 Review of Research Problem and Novelty	252
8.3 Significance of Research	254
8.4 Are the Developed Frameworks usable in Practice?	256
8.5 Conclusion of Research Domain & Methodology	257
8.5.1 Discussions	259
8.5.2 Method Limitations	259
8.6 Further Work	260
8.7 Final Remarks on Research Domain	261
References	265
Appendix	See Volume II
	305

LIST OF FIGURES

Figure 1 Subset of Systems of Interest Within a FTFoS	8
Figure 2 Relationship between Competency Framework and PhD Research Boundary.....	16
Figure 3 Structure of Thesis.....	20
Figure 4 Concept of Modelling Real World Entity(s)	25
Figure 5 FoS Core Ontology Meta-model	28
Figure 6 Four Domain MDA Partitioning (amended from Elammari & Issa, 2013)	32
Figure 7 Abstract Representation of the Model Construction Process	36
Figure 8 Recognition-Primed Decision Model (adapted from Klein, 1997)	41
Figure 9 Provisions to DSS development (adapted from Finlay (1989))	43
Figure 10 The Decision Making Process (adapted from Crooke & Slack, 1991)	44
Figure 11 Holistic View of Learning With VRTEs	61
Figure 12 Integrated Performance Model for FTS	64
Figure 13 Functional Relationship between SA and Decision Making Factors (Meystel, 2001)	72
Figure 14 Risk Elements Affecting Situation Awareness.....	76
Figure 15 Goal Oriented Flow of Control Resulting from Task Analysis (adapted from Keller, 2003)	77
Figure 16 Human-Centred Design and Evaluation Process (amended from Hooey, 2002)	79
Figure 17 Process and Structure for Linguistic requirements Modelling for Systems Design	107
Figure 18 Simple Communication Ontology Model	108
Figure 19 FoS Architecture Stakeholders Requirement Diagram	109
Figure 20 Logical Model of Requirement_01	111
Figure 21 Logical Model of Requirement_02	111
Figure 22 Logical Model of Requirement_03	112
Figure 23 Logical Model of Requirement_04	113
Figure 24 Logical Model of Requirement_05	114
Figure 25 Logical Model of Requirement_06	115
Figure 26 Logical Model of Requirement_07	116
Figure 27 Logical Model of the DSS	117
Figure 28 Amended Logical Model of Requirement_02	118
Figure 29 Amended Logical Model of Requirement_04.....	119

Figure 30 Amended Logical Model of Requirement_06.....	119
Figure 31 Flight Training Enterprise System Metamodel	120
Figure 32 Domain Model using MDA Approach	122
Figure 33 The DSS High Level Use Case Diagram	125
Figure 34 The Readiness Metamodel	129
Figure 35 FTS Workflow Process.....	131
Figure 36 Workflow Process Stage Option	131
Figure 37 Continuing the Workflow Process Prompt.....	132
Figure 38 Simplified Architectural Approach	136
Figure 39 Database Architecture Format	137
Figure 40 Pipeline Model for Mission Scenario	139
Figure 41 MEC Relationship to the Mission Scenario and K&S	141
Figure 42 Pre-Study Questionnaire Result Window.....	143
Figure 43 Indication of Simulator Sickness	144
Figure 44 ILS Linear Scale Report Form Example	147
Figure 45 ILS Assessment Result Example.....	149
Figure 46 Pre-Study Architecture	151
Figure 47 Pilot Qualities Evaluation Scale Architecture	154
Figure 48 Commencement of HQWS.....	154
Figure 49 Ending the Mission Handling Qualities Stage.....	155
Figure 50 Inputting assessment scores into the HQWS Sub-system	155
Figure 51 ROSETTA 0 Architecture	159
Figure 52 Flow for Selection of Mission Related Competencies	160
Figure 53 Prompt to Decision maker for ROSETTA 0	162
Figure 54 ROSETTA 0 Elimination Stage	163
Figure 55 Workload Planning Stage Architecture	165
Figure 56 Entering Workload Values	166
Figure 57 Task Importance Related to Mission Goals.....	167
Figure 58 ROSETTA 1 Design Architecture.....	171
Figure 59 ROSETTA 0 Elimination Database File for OSCAR123	172
Figure 60 Identification of Level of Preparedness.....	172
Figure 61 RSE Shape between K&S and MEC	174
Figure 62 Creation of RSE Shape.....	175
Figure 63 Identification of Relationship between Experience to MEC	177

Figure 64 Subset of ROSETTA 1 Design Database	178
Figure 65 Subset of ROSETTA 1 Analysis Database	178
Figure 66 Summary of ROSETTA 1 Framework Analysis Method	180
Figure 67 ROSETTA 1 Analysis Architecture	181
Figure 68 Example of Analysis Frame for One Technology	183
Figure 69 ROSETTA 1 Elimination Function	185
Figure 70 PARS Design Architecture	190
Figure 71 Calculation of PAR specific Attributes	191
Figure 72 Demonstration of High-Level Abstraction Scales from Fidelity Database File....	195
Figure 73 ROSETTA 2 Design Architecture	195
Figure 74 ROSETTA 2 Analysis Architecture	197
Figure 75 Summary of ROSETTA 1 Framework Analysis Method	198
Figure 76 Elimination Warning Messages for Decision Maker	200
Figure 77 Task Load architecture	202
Figure 78 Graph Illustrating difference Between Accuracy and Precision	206
Figure 79 Prediction of Accuracy and Precision Architecture	206
Figure 80 Accuracy Predictions for Subset of Oscar123	207
Figure 81 Pre-Flight Assessment Architecture	210
Figure 82 Goal Modelling State Behaviour	214
Figure 83 Goal Modelling Assessment Technique	215
Figure 84 Goal Propagation Algorithm	216
Figure 85 Goal Modelling Architecture.....	217
Figure 86 Inputting Data into Goal Modelling Assessment.	217
Figure 87 Example of Assessment on Student Pilot Success for Task 1	218
Figure 88 Example of Assessment on Student Pilot Success for Task 2	219
Figure 89 Potential Goal Modelling Grading System.....	220
Figure 90 Simple Block Diagram Describing Method of Assessment for Performance	223
Figure 91 Performance Based Comparison Design Architecture	224
Figure 92 3-D Comparison of Flight with Baseline.....	225
Figure 93 Analysis Using 2-D plots of Flight Attributes.....	226
Figure 94 Subset of Performance Per Task Analysis.....	228
Figure 95 Example of GUI Arrangement for Pilot Control Adjustment Assessment (subset of task).....	229
Figure 96 Illustration of Semantic Issues Relating to Database Search	231

Figure 97 Instructor Observer Rating Assessment Architecture	233
Figure 98 Post Flight After Execution Assessment Architecture	235
Figure 99 Subjective Confidence Rating Of Pilots Abilities	236
Figure 100 Total Average Rating of Post Flight Success Evaluation.....	237
Figure 101 Design Characteristics and Operational Factors (amended from Crocker, 2010)	241
Figure 102 Output Example from the Paper Based Assessments.....	243
Figure 103 Triangular Distribution Results from Simulator Configuration Assessment	244
Figure 104 Final UI Example for Simulator Configuration Assessment.....	244

LIST OF TABLES

Table 1 Subset of VRTE (Flight Simulation) Tools	54
Table 2 Fidelity Dimensions (advanced from Estock et. al., 2006).....	96
Table 3 Felder Learning Styles Dimensions (adapted from Felder & Silverman, 1988)	146
Table 4 ILS Grading Sheet (amended from work from Felder, 2008)	147
Table 5 MissionRelatedHQWS Assessment Output File (Subset)	156
Table 6 Subset of TaskNoandDetails Mission File.....	157
Table 7 Output File From ROSETTA 0 Design	160
Table 8 Scale descriptors for ROSETTA 0.....	162
Table 9 Example of a ROSETTA 0 Frame	163
Table 10 Workload Design Evaluation Results	167
Table 11 Rating of Each MEC.....	173
Table 12 K&S and MEC Ranges for Mission per FoS System	173
Table 13 ROSETTA 1 Analysis Results for Elimination of Technology.....	185
Table 14 Checklist of Factors to be Considered for Fidelity Trade-off Studies.....	193
Table 15 Stages of Expertise for the Knowledge Dimension	196
Table 16 Stages of Expertise for the Skills Dimension	196
Table 17 Example ofOutput from Task Load Assessment	204
Table 18 Example of Results from the Sub-set ofAccuracy Predications.....	208
Table 19 Goal Assessement Results	219
Table 20 Database for Paper Based Assessment	242
Table 21 Header Content of the Main Student Database File	246

GLOSSARY

ATC	Air Traffic Control	MBSE	Model Based Systems Engineering
BAE	British Aerospace	MDA	Model Driven Architecture
CIM	Computational Independent Model	MEC	Mission Essential Competency
ConOps	Concept of Operations	MFD	Multi-Functional Display
COTS	Commercial off the shelf	MFTS	Military Flight Training System
DEVS	Discrete Event Simulation	MoE	Measures of Effectiveness
DSM	Dependence Structure Matrix	MoP	Measures of Performance
DSS	Decision Support System	NLS	Natural language Sentences
EA	Enterprise Architecture	OMG	Object Management Group
EIC	External Input Coupling	OOTW	Out Of The Window
EOC	External Output Coupling	PIM	Platform Independent Model
FMS	Full Mission Simulators	PSM	Platform Specific Model
FoV	Field of View	QFD	Quality Function Deployment
FoS	Family of Systems	ROSE	Relational Oriented Systems Engineering
FSTD	Flight Simulator Training Device	ROSETTA	Relational Oriented Systems Engineering and Technology trade-off Analysis
FTFoS	Flight Training Family of Systems	RSE	Response Surface Equation
FTS	Flight Training System	RSM	Response Surface Methodology
GBS	Ground Based Systems	SA	Situational Awareness
GBT	Ground Based Technologies	SME	Subject Matter Experts
HITL	Human in the Loop	SoI	Systems Of Interest
HMD	Head Mounted Display	SOP	Standard Operating Procedures
HOTAS	Hands On Throttle-And-Stick	SoS	Systems of Systems
HPM	Human Performance Modelling	SOSE	Systems Of Systems Engineering
HUD	Heads-up Display	TNA	Training Needs Analysis
ISO	International Organisation for Standardization	ToT	Transfer of Training
IT	Information Technology	UI	User Interface
ITAR	International Traffic in Arms Regulations	V&V	Verification and Validation
LVC	Live, Virtual, Constructive	VE	Virtual Environments
M&S	Model and Simulation	VRTE	Virtual Reality Training Environment

ACKNOWLEDGEMENTS

I would like to express the most sincere gratitude to Professor Charles Dickerson and Dr. David Mulvaney, for all the invaluable discussions and consistent support throughout. A special note to my industrial supervisor, Dr. Andrew Bradley who started the original research project with me but for whom, unexpectedly, had to bow out due to ill health.

I would also like to thank Professor Roger Dixon, who began the research project as my internal examiner, for his advice and support through the viva-voca's throughout the first two years, but who then for various reasons unattached himself from the research project due to the complexity of the new research direction. A special acknowledgement goes out to Professor Ron Summers for agreeing, upon short notice, to become the new internal examiner, without him undertaking the Viva; I would not have progressed into the final year.

A special note for John Scott, who is a previous senior manager from BAE (now not with the company), who became involved in the project and completely changed the direction half way through the project time inviting a new literature review etc. and the realisation of the wide scope area which needed to be covered. Hence, this journey has been far from uneventful and I have no idea how I managed to be at this stage or how I've finished on time – but for the experience, thank you.

I could not have started this research project without the support of friends and family, who have placed a great deal of faith in me of being the first of the family to obtain a doctorate. Oh and to my girlfriend, thank you for your patience and understanding. The last year of the project was incredibly difficult to progress through having lost my father to phase 4 Cancer; and my only regret is that I could not be finished in time so my late father could see his only son graduate – destiny (if there is such a thing) was always against me finishing early. So dad, I eventually did finish it despite all the obstacles!!!!

I have experienced a lot over the years at Loughborough, from writing lectures to teaching, waiting for an imminent lab move that never happened until at most inconvenient time, through to architecting a course for international students – it has been a very busy four years especially since the complete change of direction half way through. What has been produced is directed towards an academic piece of work, but industry and not just the aerospace industry will hopefully be able to adapt some of the methodologies into their own processes to streamline and make more formal some complex decision making tasks.

AUTHORS DECLARATION

This is to certify that I am responsible for the work submitted in this thesis; that the original work is my own except as specified in acknowledgements or in footnotes, and that neither the thesis nor the original work contained herein has been submitted to this or any other institution for a higher degree.

Signed:

Print:

INTRODUCTION & BACKGROUND

This thesis will briefly describe the work completed to date, reiterate the novelty of the research and communicate succinctly the justification for inclusion of various methods and assessments within the proposed methodology. Central to the research in this thesis is the investigation, development, and assessment of the formal decision support framework to develop a methodology to satisfy the PhD project requirements. These begin in Chapter 1.3.1 with an initial critique of the mathematical formalization of system modelling using design matrix methods. Combined with requirements from the industrial Sponsor in Chapter 1.4, this leads to the consideration of ROSETTA (Dickerson & Mavris, 2011) as an alternative to other approaches. After summarising the requirements, Chapter 1.4 introduces a route map to the investigation and assessment of ROSETTA established on model based approaches, which includes relational orientation techniques to system engineering that will be continued in the wider systems context in the remainder of the thesis.

The diversity of the operational environment in the new globalised world has increased the complexity involved in global missions for military scenarios. This places a huge burden on advancing training systems to accommodate scenarios involving environments in each corner of the earth; furthermore, this places a strain on present and future training budgets (McHigh & Casey, Jr, 2010). The technical problem of developing and assembling flight training systems embodies the core challenge of the System of Systems (SoS) and / or Family of Systems (FoS) engineering that encompasses both commercial and military domains. Flight training systems are information intensive and currently developed and assembled in an ad hoc and bespoke manner involving labour intensive and document based management systems, which are problematic to control.

“The general unreliability of all information presents a special problem: all action takes place, so to speak, in a kind of twilight.....like fog”

- Carl von Clausewitz, On War

The research is motivated by the current lack of rigorous framework structures that are able to impart knowledge to Subject Matter Experts (SME) of the numerous systems within the FoS utilising information management techniques to permit traceability of each student pilots progress to organisational objectives in relation to the training technology/system(s) capability. Systems of Systems Engineering (SoSE) approaches are favoured to solve such problems as: integrate, test, and assess the interoperability of required components and thus the development of new architectures are continuously evolving (Lane, 2007). One such

system engineering approach, Enterprise Architecture (EA) is “*the continuous practice of describing the essential elements of socio-technical organisation, their relationships to each other and to the environment, in order to understand complexity and manage change*” (Blevins, 2006). Thus, EA is a method of systematically representing an enterprise including the business processes, systems and software applications, information flows, and information technology (IT) involvement.

Systems which are classed as SoS or FoS are composed of heterogeneous systems that are developed and procured separately and incur a level of interoperability to complete a mission or task. The management of developing such SoS are laden by the difficulty of gaining confidence that a particular architecture of subsystems will achieve a global SoS-level requirement. A model based approach to flight training as an enterprise has never been successfully attempted before; however the e-enabling of the commercial airlines in (Wilber, 2007) identifies that the aircraft was one component of a much larger system. The FoS problem of interoperation and communication involves numerous diverse elements that constitute the operational system including: flight operations, Virtual Reality Training Environment (VRTE) integration, and various management systems. This is a hugely important area of research for the military domain and is at the forefront of Human in the Loop (HITL) interaction with technology in mission critical and life critical situations. VRTEs can be used to prepare skills and mental capabilities for a subsequent live demonstration; the skills dimensions of a VRTE consist of: adaptability, situational awareness, performance monitoring and feedback, leadership and team management, communication, and decision making. The design task for technology trainers is to implement fast-paced decisions by supplying the pilot with as much information necessary while reducing the amount of distracting information, which is difficult to achieve in a compressed cockpit (IFALPA, 2012). However, are there questions regarding the use of such technology along with VRTEs and game based technology used for training. Most of the effects of game based technology are subjective, lacking an empirical foundation (Hays, 2005).

Graessar and King (2008), explain the need for intelligent training on complex domains that involve conceptualization and verbal reasoning that are not mathematically well formed.

‘The Challenge of combining entertainment and pedagogical content is the foundational question of serious games’
 – (Graessar & King, 2008)

The new generation of students have a different set of skills than the previous and other emerging technology requires that existing methods of instruction be reviewed and adapted to leverage these new technologies (Bullen et. al., 2011). However, VRTEs along with simulations are limited in training due to the developers of the system having insufficient training in andragogy, behavioural sciences, and learning technologies (Hamill, 2005). It is clear the requirements and characteristics of the learner have to be considered when developing high-tech learning environments and training programmes involving consultation with SME within relevant domains; there are far too many VRTEs introduced without the required verification on usability, engagement, and learning gains (Gee, 2003);(Holden & Westfall, 2010).

Currently, the simulations for military purposes are designed to be as realistic as possible. It is considered that the simulations can then be used to provide measureable and repeatable results that can be validated by real world observations. To be more formal, the simulations are used to bring precision to events that can teach someone how to do something, and the process of making such models. The art of understanding complex systems problems is to design architecture, which describes the model, at multiple levels of abstraction to obtain greater knowledge and understanding of the relevant entities that need to integrate and interoperate together.

1.1 Why is Research Needed Within the Flight Training Domain

‘Great Pilots are made not born.....A man may possess good eyesight, sensitive hands, and perfect coordination, but the end product is only fashioned by steady coaching, much practice, and experience’

- Air Vice marshal J.E. RAF.

Present training, no matter which domain, is all about optimizing performance. Pilots need to be able to perform efficiently, individually or part of a team, in terms of exploiting the tools available, namely the VRTEs and the aircraft. Inadequate investment in the right tools to perform the associated operational task, no matter how good the training programme is, would lead to significant problems in terms of economic waste of money, time and effort, but also could force poor decision making strategies in training individuals and/or teams using the incorrect technology for training (MOD, 2015). As most training is directed at the individual level, it is perceived that VRTEs will provide low cost effective training. Decisions of which VRTE to acquire is generally open to interpretation between SMEs, all of whom tend to have different opinions and priorities. There is increasing concern that reducing the number of

flight hours and increasing the number of simulated missions only shifts the expense to the simulator environment because the fidelity value of training must be maintained (Longridge et. al., 2001). Development of LVC capability (see Volume II, Table I) to utilize bending technologies to permit inflight simulated and constructive threats has been seen as one solution to maintain the fidelity value of training. For the training organisation, training within a VRTE is used to record compliance with SOPs, and real live flight is used to validate the pilot training without the emphasis being on SOPs (Giles, 2013); but rely on SME opinion on pilot progress and decision making ability (Hosemann, 2013).

Observation is one of the most important techniques to assess students and is the prime consideration for training design. This has steered to instructor led training, where the description of the task is given, the student is allowed to practice, and thus it is perceived training has occurred (Woods, 2004). Fortunately, this process has worked well for years, however, with the added complexity of aircraft all with disparate cockpit layouts and the mass widespread use of VRTEs, this process is being strained as it cannot be clear whether the student is following the correct mental model for the task within the flight simulator; as repercussions of incorrect processes are relatively unclear (RAS, 2009). Thus, task analysis is needed to identify what students do, not what they or instructors think they do (Chandra & Gray, 2012).

Military Flight Training System (MFTS) development concerns the integration of combat multi-role aircraft that are reliant on precise system management capabilities requiring significant knowledge of the pilots on weapon systems and computer systems than previous fighter pilots before them (Mason, 2013). The RAF training system capabilities comprises of full-motion simulators together with an assortment of non-motion flight training devices supported by sophisticated computer based learning and information systems. The full mission simulators (FMS) provide advanced air combat training, operational scenarios, synthetic radar, surface-to-air missiles, air-to-air missiles and even decoy systems. The new Hawk MK2 AJT aircraft incorporates additional synthetic functionality for the next generation of pilot training offering some of the capabilities of the FMS, whilst live flying. However, the previous training Hawk aircraft is arguably easier to fly in terms of workload as the new Hawk with inbuilt multi-function displays (MFDs) relies on large scale system management that requires intellectually focussed pilots (Mitchell, 2013). Furthermore, concentration has to be given to the acquisition of knowledge and how individuals manage

the multitude of sensors to assimilate and process large amounts of data, which modern aircraft and VRTEs provide on displays and/or available at the touch of a button (Sumwalt et. al., 2002).

It has become clear that a holistic and adaptive training system is needed that enables a pilot undertaking training to address the entirety of physical and mental challenges, nevertheless, in modern economics cost effective training using blending mixes is one of the main requirements. It is perceived that this can be accomplished through a blend of Ground Based Technologies (GBT) and in a synthetically enhanced aircraft, such as the AJT Hawk, when correctly operated to maximise the training value of each training sortie. With greater emphasis on cost, the need to better understand how to use the technology to eliminate waste with the assurance that pilots will still acquire necessary knowledge and skills (K&S) using synthetic based technology, has become a higher priority. Having the ability to make best use of existing, and possible future technology (i.e. trainers involved in development of new GBT) using more synthetic based training, it is perceived that real flight hours can be traded for virtual hours.

'Flying is so many parts skill, so many parts planning, so many parts maintenance, and so many parts luck. The trick is to reduce the luck by increasing the others'

- David L. Baker

The goal for a successful training programme, therefore, is to expedite the acquisition of expertise in student pilots to provide effective training for the front line using the correct training technology to accelerate progress through the pipeline and allow students to achieve the minimal acceptable performance to receive their wings, with minimal cost (Pease, 2009). The aim/objective is to develop a training programme not to deliver pilots who do well in training, but, to deliver pilots who can perform well in the real world using a combination of real and virtual training (blending mix) (NAP, 2015).

1.2 Statement of Problem

The globalised world has many challenges in an increased diversity of operational environments, which incur competing requirements for training resources that placed increased burden of present and future training budgets (Mchugh & Casey, Jr., 2010). Along with these stressors, management of training resources is directly under pressure with the lack of understanding of how to use technology within the training context for preparation of pilot readiness. Research has shown that if training goals are not correctly embedded into the

‘game’, the only transfer of knowledge gained is how to play the game, rather than the intended knowledge and skills the ‘game’ was supposed to achieve (Belanich et. al., 2004).

Military training is expensive and the current philosophy to the training problem is with the belief that training is effective. However more prevailing question is: Is an expensive piece of simulation technology effective in training combat pilots?. Although the technology used for training goes through a rigorous verification and validation (V&V) process and acceptance testing, the flight training system does not and is predominantly for SMEs (instructors) to evaluate the learning effectiveness of the training system (FAA, 2014). The purpose of training is to achieve a level of readiness for pilots, however, currently training is a ‘check-box’ system and student pilots are evaluated only to the degree to which they complete training and not necessarily to the ability to perform trained behaviours in an operational setting or the effectiveness of training using blending mixes (ICF, 2013). It is heavily reliant on instructor opinion meaning that it is difficult to identify the strengths and weaknesses of training technology/systems; and as a result, improving the level of mission essential competencies (MEC) and quantifying the relationships between it and knowledge and skills (K&S) with disparate training technology/systems has proven difficult if not impossible in present training systems (Alliger et. al., 2013).

Most flight training instructors are confident of training student pilots in real aircraft, but they are not confident in how to teach pilots using the VRTEs in replacement of live flying because they are not taught how to use the VRTE as an effective instructional tool (Champney et. al., 2006). Low fidelity VRTEs are often freely available for student pilots at relatively low cost, whereas high fidelity trainers are much more expensive to acquire and maintain, require specialised personnel to run, and are less accessible. Time spent in the VRTEs is usually driven by availability or the planned class curriculum with little guidance on how best to distribute the available VRTEs to pre-specified training regimes (Taylor et. al., 2005).

The procedures used by UKMOD to procure technology for training were difficult to apply and the results revealed that many training technologies were poorly specified to support operational needs (MOD, 2004). In an attempt to resolve this problem ad-hoc complementary set of procedures, known as training needs analysis (TNA) was developed, which concentrated on a number of design principles (Van Der Pal, 2003) to formally delineate considerations of training effectiveness from those of cost effectiveness. However, this type

of analysis are often minimal and hence a reliable and comprehensive TNA is rarely accomplished, and often lacks the inclusion of a solid fidelity analysis including human-in-the-loop (HITL) considerations (Van der Pal et. al., 2010);(EASA, 2014) (see Chapters 3, 5 & 7 for further discussions). These issues have forced training developers to look at unique ways to enhance the current training system using new technology to save resource costs; thus, there is a need to seek appropriate blending mix training methods that will ensure the highest levels of readiness whilst being cost effective in training (Buxbaum, 2010).

With the increased use of emulation/synthetic data on aircraft for live training, the synthetic boundary is becoming increasingly blurred; therefore, with the abundance of training technology available systematic consideration of the most appropriate blend is needed (ICAO, 2014). Evaluation techniques are necessary to analyse blending of live and synthetic technologies and methodologies need to be developed to assist decision makers in procuring effective and blended training solutions. The methodology required for management of resources in a flight training system has to support transfer of training (ToT) with time savings in a live environment. This should work towards reducing training risk and maintain or improve training quality, this is hoped will drive changes in technology development training management to best suit training needs and allow a firm understanding of how technology characteristics and configurations affect training outcomes.

Methods of assessing blending mix consider objective assessments based on human considerations, functional fidelity, learning complexity factors and conditional factors (environmental) that allows a decision maker a choice using a rating system based on semantic descriptors, however, the task analysis regarding blending mix is actioned with a binary live/virtual choice with little or no direct cohesion to the objective assessment and as such choice between training blends is still heavily reliant of subjective opinions (Freitas & Jarvis, 2006);(ICF, 2013). The competency-based training (Van de Pal et. al., 2010) provide methods in which a training program can be evolved within the training programme, furthermore, it gives a set of parameters with which the selection of an effective learning environment or blending mix for the desired learning outcomes should be related to (TNO.NLR, 2007). Thus, concentration is needed in the ability to design training programmes that will lessen the student pilot learning curve and optimize his/her workload by concentrating of the capability of the blending mix to effectively evaluate and practice the planned mission tasks to effectively move the student pilot swiftly through the training pipeline.

1.2.1 Research Enterprise Viewpoint Needs

Continuing discussions with BAE and a progress review meeting with the industrial sponsor Mid-2013, their business needs involves issues surrounding the enterprise viewpoint, which is briefly discussed below.

Currently, the Flight Training Family of Systems (FTFoS) is supported by a complex family of systems (FoS) from flight simulators to aircraft avionics. The FoS are a mix of live, virtual and constructive (LVC) systems, which include documents, desktop PCs, ground based simulators (GBS), training aircraft and a number of other subsystems that training aircraft carry for flight realism as shown in Figure 1.

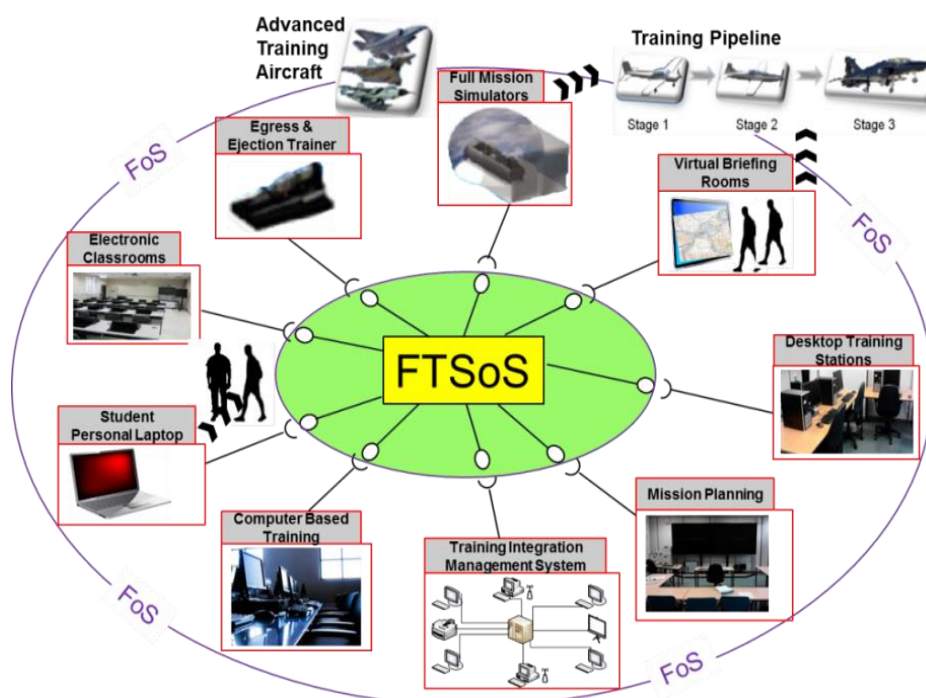


Figure 1 Subset of Systems of Interest Within a FTFoS

One of the key technical problems is how to evaluate, develop and assemble a family of flight training system into mixed media training environments where there is an integrated Live / Synthetic blending mix for aircrew training. This mix is generally constructed in an ad hoc manner that is not repeatable and often paper documents are used across domains with no direct convergence between them (Doiron, 2014). The key driver for BAE is to sell the Hawk T. Mk2. advanced jet, which is marginally more expensive than the competitor's equivalent. Furthermore, BAE currently has little expertise within the training environment to advise how best to use the advanced technology of the aircraft i.e. live with synthetic mix, within a blended FTFoS.

1.3 Research Methodology

In addressing the research problem, the methods employed involves the underlying ‘system research’ paradigm comprising knowledge and methods from studies of the social science of organisational behaviour. The social and human attributes within complex socio-technical systems are inherently difficult to quantify and thus has an increased reliance upon human judgement, intuition and understanding of the context of the system that forms the overall study (Hodgson et. al., 2013). The reductionist approach reduces the complex phenomenon of the system to a series of simple, controllable and measureable attributes within sub-systems with the aim of achieving scientifically viable results (Zanetti, 2013). Systems design attempts to consider the interactions between the sub-systems and aims to understand the ‘emergent’ behaviour resulting in unwanted system responses. The complexity of human activity combined with the socio-technical systems interactions is the perfect choice for using the systems engineering approach to solve complex organisational problems. The method of abstraction and the ability to draw boundaries between the various systems of interest (SOI) incurs the ability to leave out unnecessary detail and still consider an all-encompassing view of the whole system, which promotes ease of understanding of a complex multifaceted system or FoS/SoS problems (Dumitresco, et. al., 2013).

The research methodology adapted for the PhD project is directly related to the research techniques and engineering techniques at the Scientific / Interpretivist boundary identified by Clarke (2000). The methodology describes a number of research stages involving, conception, design and prototyping of information technology artefacts, followed by the generation of new information regarding the existing class(s) of technology. The techniques to complete the methodology relating to this research topic include:

- Interview based techniques – subjective views.
- Subjective assessment techniques – feeling based on experience.
- Mixed methods research – qualitative and quantitative data.
- Decision Support tool development – decision between solutions.
- Decision Support System verification methods – proof of concept.

1.3.1 Solving Multi-Attribute Problems

Generally a system is regarded as an assembly of objects or entities each with attributes and operations with relations between attributes together with relations between entities. The establishment of model based approaches for system descriptions and analysis (INCOSE,

2015) has led to mathematical formalization of system modelling using the principles of homomorphism of relational structures (Tokunaga & Fujimura, 2013). The organisation of system entities forms the principle of structured engineering design; additionally the principle of analysis asserts that the specification of design should be separate from the solution (Guenov & Barker, 2005). A design matrix can be used to mathematically transform between functional requirements and design parameters has the ability to identify in what ways elements of a system or model can relate to or be dependent upon each other (Browning, 2001). The general use of mappings is to seek a solution that best meets a set of requirements.

The goal of engineering analysis in design and technology trade-off is to compare and evaluate a number of solutions against a set of requirements, which makes the process of decision making a multi-attribute problem. The management and planning tool referred to as the House of Quality (HoQ) uses a form of mapping technique that uses qualitative mappings to propagate customer requirements and preferences (elicit tacit knowledge from SME) to engineering parameters through a series of more detailed parameterizes frameworks until the series of mapping are traceable from customer to the system component level to identify key design drivers and trades (Jaiswal, 2012). The Quality Function Deployment (QFD) approach (Iqbal et. al., 2014) originated in manufacturing where relationships have been observed to be linear and non-negative, thus the relations in the QFD are also linear (Fehlmann, 2015); in addition, it is difficult to measure the correspondence between requirements and engineering parameters on a uniform scale to obtain a Pareto improvement measure for the system design process (Bhattachary et. al., 2010).

An alternative and a more generalised approach is to use relational frames that are suited to capture relationships and dependencies between system elements using sensitivities derived from analytics (Dickerson & Mavris, 2013). The ROSETTA framework offers a more structured and mathematical based approach to the QFD methodology. ROSETTA, proposed in Dickerson and Mavris (2011), provides a means to translate between theoretical mathematics, subject-matter-expert (SME) driven analysis, and M&S, by representing a single problem using all three types of analysis and highlight the commonalities and differences between the different representations of the problem. The framework permits a more accurate Pareto measure for decision support between design solutions that is based on sensitivity analysis, using M&S techniques, to describe the relationships between requirements and system elements. This entails considering the outcomes of the analysis, the

socio-technical constraints and processes, and the development and implementation of a model based design in a multi-attribute system for decision support in an area that is currently dominated by technology driven principles.

The model based system engineering (MBSE) method used allows multiple perspectives or viewpoints (Murray, 2012) that facilitate knowledge of the entire socio-technical FoS to be considered, which forms the domain of interest (described by the requirements narrative in Chapter 1.4) and can help identify the attributes needed for an in-depth study that enables close ties to the target application environment of a flight training system (FTS).

1.3.2 Methodology Direction

The research will describe a methodology at a high level of abstraction that will investigate the development of rigorous frameworks and a prototype workflow process tool to support a number of domain training needs simultaneously in addition to providing decision support to allow integration of performance metrics for efficient progression through the training pipeline. The workflow process incurs the ability to capture knowledge of SME subjective / objective assessments and evaluations to gain understanding of the suitability of GBT and blending mixes used for disparate training scenarios. Hence, the PhD research is directed to the development of a ‘proof of functional concept’ Decision Support System (DSS) with the investigation of the feasibility and use of a specialization of the ROSETTA frameworks (see Chapter 5) and prototype workflow tool (see Chapter 7) to assist decision makers in procuring effective training mixes. (See Holden & Dickerson, 2013 for further details). These intended techniques and feedback methods prescribed by the workflow of the methodology are used to provide features that a decision maker needs to assess suitability of a particular training technology to a student pilot for a more efficient training system at the FoS level.

Planning the research involved the use of mind mapping techniques that defined the research problem, technical approach, potential problems and solutions, and practical consideration of verification of the developed ROSETTA frameworks and process workflow tool (See Volume II, Figure A).

1.4 Research Aims / Objectives

The aim of the research is to develop a conceptual model of the FTFoS enterprise that can support the simulation of the enterprise for the purpose of analysis and decision support considering the available training technology/blending mix and suitability to the student

pilot(s) using the research problem of the requirements narrative below for the basis of systems design.

This research will prescribe a methodology inclusive of ROSETTA framework(s) to provide structure for Live / Synthetic (airborne and ground based) aircrew training and a modelling, simulation and analysis approach for the FoS aviation training problem that can be used to support capability based trade-off decisions to select optimal flight training FoS mixes. The ROSETTA framework will permit allocation of systems in the mix at the task and activity level for Mission Essential Competency (MEC) but will provide automated assessments of the cohesion of the FoS architecture. Simulation of the FoS architecture will be used to assess the capability to support a coherent scheme of training for the aircrew under realistic operational conditions. Pre-flight mixes can be assessed in trade-off analysis; and the simulated performance of the selected mix can also be compared to actual performance in the post-mission analysis. The ROSETTA framework structure, once all the factors and variables are established, can be advanced to assist in the delivery of a more efficient and cost effective training environment at the FoS level.

Research Proposal Requirements Narrative Agreed With Industrial Sponsor

Academic objectives are:

Req1: Development of a foundational meta-model for the flight training system. This model will provide concordance between models of the VRTE and the training with disparate aircraft and systems for choosing correct blending mix.

Req 1.1: Development of an open architecture specification for Training Family of system.

Req 1.2: Facilitate the transfer of methods and knowledge from academia to industry, strengthening the relationship links.

Req2: Analyse the employment of model based techniques within a Decision Support System (DSS) to resolve the issue of representation, utilisation and decision support for complex systems involving socio-technical processes within the FTFoS.

Req 2.1: Improve the understanding of how DSS tools and techniques can be advanced using mathematical based relationships within a ROSETTA framework to produce a more robust and repeatable method for trade-off analysis for selection and suitability of blending mix, as per requirements narrative.

Req3: Produce ROSETTA style DSS frameworks which can assist in the integration of quantitative and qualitative type data to aid ‘unbiased’ decision support and trade-off analysis for selection of blending mix, as per requirements narrative.

Req3.1: Provide mathematical and model based formalisms for a relational oriented viewpoint for the training FoS architecture.

Req3.2: Investigation of the feasibility of using the ROSETTA methodology to specify framework structures for systems engineering.

Req4: Development of a prototype ‘proof of concept’ workflow process for administration of the ROSETTA framework(s) and associated databases.

Req5: Set foundations for further research collaborations with industry.

Industrial Objectives include:

Req6: Investigate and assess the potential for a mathematical based ROSETTA framework, employing various modelling techniques and tools, to resolve multifaceted applications for decision support with respect to a complex flight training system to give a competitive advantage.

Req6.1: Assessment of techniques used to develop the ROSETTA framework(s) and the workflow process tool through a ‘proof of functional concept’ DSS application to evaluate the limitations of the proposed DSS design.

The industrial sponsor ‘BAE Systems’ main focus in the research project is leverage of capability-based assets for competitive advantage. In this case the asset is deemed to be a DSS for the organisation and management of technical FoS within a flight training system for additional support to current and future customers.

Regarding **Req6**, in this thesis the critique of a framework that is capable of supporting both modelling (and simulation) as well as a decision support system has begun initially with Chapter 1.3.1, which considered ROSETTA as an alternative to other approaches. In keeping with **Req6**, the investigation and assessment of ROSETTA will be continued in the wider systems context in the remainder of the thesis.

Specifically, Chapter 2 establishes the wider systems context for the research. Later, Chapter 5 describes the advantages of using the strengths of both qualitative and quantitative data within one framework structure proposed by an implementation of the theory of ROSETTA that can be used for modelling and simulation of a complex multi-attribute problems to aid in

the decision making task by providing measures relating to design solutions. Chapter 7 then describes the design of the methods along with arguments for the feasibility of using the ROSETTA framework(s) for decision support. Chapter 8 (Summary and Conclusion) includes an assessment and final critique (Chapter 8.4) of the methodology for the decision support system integrating the ROSETTA framework(s) to identify causal relationships between system elements of the flight training system (FTS).

The research output will include specialized implementation of the ROSETTA framework(s), as per agreed requirements narrative summarised at the start of this chapter, that can be used for assigning MEC models to specific training and tactical models within a training scenario and the allocation of Live / Synthetic systems for the purpose of simulation, interoperability analysis, and trade-offs for optimisation decisions. The repeatable methodology offered by ROSETTA can be used to develop and assemble a family of flight training systems for efficient and cost effective blended training to help guide and improve training levels. The systems in the FoS are seen as an integrated collection of domain models (relational frames) which are referred to as the conceptual model of the system. Quantitative knowledge for assemblage and analysis of the flight training FoS is typically provided through system attributes and mathematical equations that relate the attributes; every row in a matrix representation of a ROSETTA framework then corresponds to an equation or a simulation. The cells / slots in the matrix can be used to represent sensitivities between attributes of the systems within the FoS. The system responses can then modelled by a transformational frame of response surface equations or simulations or both (Jou, et. al., 2014). The integration of the FoS and scenario models, using the transformation frame of the framework, provides a formal conceptual framework that can support powerful innovations in design and analysis for the training FoS problem.

1.4.1 Research Deliverables

Research deliverables are directly related to the requirements narrative given for the research described at the start of Chapter 1.4. In order to achieve the objectives of the research, key deliverables were identified as being:

- Use model based systems engineering (MBSE) techniques for the context of developing the DSS methodology for the flight training enterprise to produce mathematically robust models that reduces complexity within large multifaceted problems.

- Demonstrate an understanding of DSS tools and techniques that can be used to solve complex organisational processes.
- Assessment of proposed techniques for the DSS application.
- Investigate and implement a specialisation of the ROSETTA framework theory into a flight training system described by the requirements narrative to facilitate decision support for training tool blending mix for pilots to expedite the decision making process to enable a coherent scheme of training between decision makers.

1.5 Scope of Research

The research described in this thesis concentrates on methodological considerations and therefore does not seek a full solution to the requirements narrative, rather inquires whether a decision support system, inclusive of multiple ROSETTA frameworks, can be developed inclusive of multiple viewpoints with the inherent ability to either identify bias or reduce its affect and assist in acquisition of knowledge where uncertainty exists with concentration given to technical aspects of the FTFoS.

To enable focus on relevant attributes of concern, the scope of research has been confined to enable a systems engineering viewpoint on the FTS described by the requirements narrative. FTS are, generally, governed and operated by training contractors who have a critical interest to protect their intellectual property (IP) regarding useful processes; in addition, government sponsored defence contractors have strict controls regarding defence related technology and systems (re: International Traffic in Arms Regulations (ITAR) restrictions (IEEE, 2014)). Thus, information and documents used in the development of the conceptual model, which encompasses the proposed workflow process, is based on an extensive literature review, interviews with training contractors, and various assumptions made, which have been acknowledged by the industrial sponsor. Hence, the produced methodology has been designed on the information gathered by the literature and interviews with assumptions based on current information as to how a real FTFoS for training military pilots is governed and operated. Furthermore, research into availability and management of facilities with the FTS is being actioned both as part of BAE industrial research and maintenance factors are being researched by another university, therefore, these aspects are accommodated in the methodology at a high level of abstraction.

The proposed specialisation of the ROSETTA framework(s) requires access to training attributes, which in a real system has been derived by SMEs for the human system of the FTS

as part of a competency framework (ICAO, 2014), to perform sensitivity analysis between system attributes. Nevertheless, for the methodology the specifics of the competency framework are deemed already to be in existence and the methods described has an indirect relationship (via a shared database) to the existing training attributes. The proposed methodology does not define or evolve the competencies or identify what competencies are required by the student pilot for human performance, as described in Figure 2.

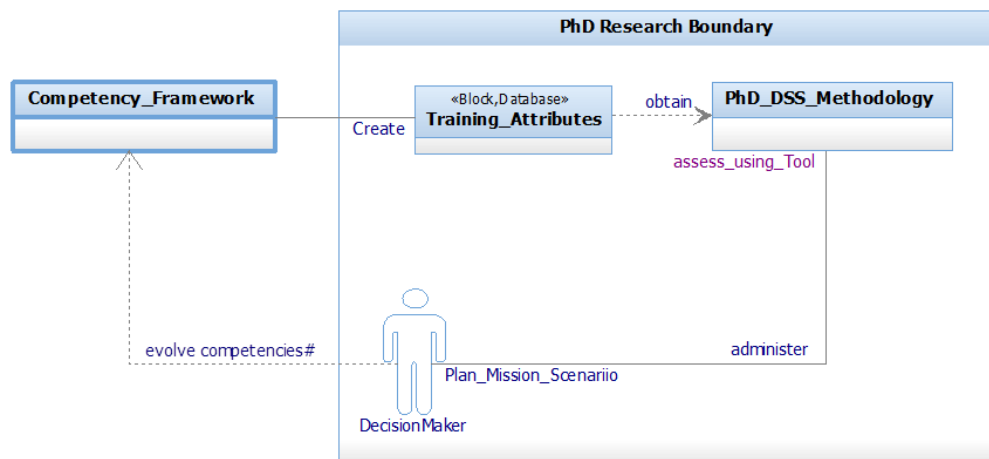


Figure 2 Relationship between Competency Framework and PhD Research Boundary

The training attribute descriptions are saved in a database for a given system reflecting in a broad chronological order the information on specific tasks performed and competencies gained dependent on the mission task to be completed (some mission attributes emerge with the result of identification of others so that one may have its start at the end of another whilst progressing through a training mission). The methods described in the thesis access a generalised training attribute database, for ‘proof of functional concept’ basis, to obtain a list of training attributes to be selected on a per task basis to create a full mission list for mission planning and subsequent sensitivity analysis to technical system characteristics, which encompasses the PhD research boundary.

Due to highly independent performance variables and limited flight experience from the project principle investigator, a holistic systems perspective in developing the DSS was considered important. This approach does not seek to reduce the complex functioning of the FTS, but for verification the number of variables and analysis/test parameters has been more focussed to match the technology available for the research project and is described within the system model artefacts. The focus of the socio-technical factors considers a multitude of attributes from instructor’s subjective opinion (current FTS is reliant on) that influences the qualitative assessments of student pilot’s performance, feedback from the student pilots,

feedback analysis from the flight training technology, and subjective and dynamic mathematical relationships based on statistical data from all actors. The research highlights the focus on performance management systems upon ‘hard’ technical issues in operations, with metrication of subjective feedback closely associated with mission scenario outcomes and the suitability/capability of training technology/systems for training correct levels of training attributes.

The assessment of the DSS occurred in a laboratory environment, as a means of proof of feasibility and utility of the methodology for its intended application in the flight training organisational context in question – validation of the methods and DSS, therefore, is beyond the scope of this research project and the continued development and validation is held in the context of the capability development team of the industrial sponsor. This industry-driven research, especially the methodology, is applicable to multiple domains and disciplines.

1.6 Research Novelty

This research will provide valuable insights into management practices for FoS/SoS including providing framework structures to support the management of elements of inter-related systems that are functioning as a FoS. The complex socio-technical systems involved within the flight training domain has always been managed in an ad-hoc manner using instructor observed feedback mechanisms as the basis for pilot readiness evaluations; However, the problem of blending mixes becomes more than an optimization issue; the problem includes how decision makers can assess procuring the blending technologies. The novelty of quantifying complex evaluation processes within a simple workflow process facilitating requirements traceability for lifecycle management has never been successfully achieved before. The advanced relational modelling and transformation approaches can provide a structured engineering environment to enable end-to-end design and analysis of the FoS architecture.

Requirements traceability through the architecture will be an important aspect of the research to facilitate compliance with training objectives. The MBSE approach used will provide mathematically based formalisms that are precise and repeatable while offering agility in design and analysis; thereby offering a significant advancement to SoSE. The unique approach in developing relationships between flight training parameters and training needs / devices provides additional knowledge available to SMEs that will lead to future developments of advanced flight training processes and technology (including aircraft and

VRTEs). As the models of flight training, their relationships (human and technical) and transformations are developed, the readiness metamodel and FTS workflow process (Chapter 6) can evolve to abstract training parameters (MEC, K&S, performance accuracy, etc.) that supports architecting and engineering other flight systems for knowledge and information management for existing and future organisational needs. The combination of subjective and objective data retrieved from both the decision maker (instructor) and the student pilot, is used to evolve quantitative relationships stored in the frameworks allows for knowledge transfer to SMEs regarding technology suitability for training attribute types and levels.

1.7 Structure of Thesis

The wider Systems Engineering (SE) context is introduced in Chapter 2, where the importance of the system structure when it relates to the representation of the system in a model is discussed along with the method of abstraction to reduce the complexity of the model throughout the lifecycle. The chapter then introduces the concepts of model ontologies relevant to the research problem. A core ontology meta-model is presented that will be used to control the scale of the model architecture including considerations on Model Based Systems Engineering (MBSE) methodology in representing complex systems and closes with a brief discussion on the management of complexity of the methods of the proposed methodology and how this relates to systems engineering techniques.

The thesis concentrates in formalising the decision making process, consequently, Chapter 3 discusses decision making techniques in the environment to which a decision support system has to function and the benefits of using frameworks to assist in decision making tasks. A brief discussion about the selection of attributes that form the parameters of the framework and cost associated with detailed decision making techniques conclude the chapter.

Chapter 4 discusses the complexity involved in training with VRTEs with an introduction to the types of VRTEs used with the flight training domain along with the critical factors of interest with humans training on such technology. A portrayal of the specific training attributes of consideration for training is discussed incorporating assessment and evaluation criteria's with an introduction to pilot decision making activities when coping with workload and situation awareness stressors. The flight training attributes and performance considerations is referred to giving details of the problem regarding transfer of training from training technology exercises to real operational environments with further details regarding the transfer of learning (Classroom to practical tasks) to achieve competence. The chapter

concludes with a brief account of current training technology trade-off options between live and virtual mixes to account for efficient management of current flight training systems.

The proposed ROSETTA framework for modelling and simulation of the flight training system is introduced in Chapter 5, which debates the use of surrogate models that describe the quantitative relationships between framework parameters with which to base trade study decision on and assist in associating training technology and human attributes to performance outcomes within a workflow process for training management. Further details on the fidelity characteristics of training technology are discussed including the analysis of organisational objectives within the software tool that relates to the proposed methodology of the research project problem.

Chapter 6 begins discussions of the methods used in the proposed methodology and commences with logical modelling of the research requirements and discusses the importance in gathering system knowledge through both the natural language statements and human perspectives from experienced individuals in an existing human system for the purpose of requirements modelling. Meta-modelling of the system is introduced for the purpose of gaining knowledge of working practices and roles systems elements performs within the system followed by a description of the flight training system domain model using a model driven architecture approach to modelling. To emphasise the relationship between training outcomes and training attributes the readiness metamodel is produced to give an indication of specific attributes with which trade-studies can be based upon. The chapter concludes with a description of the proposed flight training system workflow process that defines the stages of assessment for training technology elimination and relationship evolution.

The architecture used to develop the workflow process for the methodology is discussed in Chapter 7 which includes a high level description of the methods employed and discusses the feasibility of each in relation to the research problem. The software-based information management tool offered by the methodology presents a ‘proof of functional concept’ understanding of using the output data from the evaluations of each assessment stage as the basis for argument. Included in this chapter is a simplified (abstract) architectural approach to the workflow process for ease of understanding and a method of data management involving database architecture.

The conclusion assesses how the research objectives have been met, reviews the research novelty and the evaluation of the decision support system developed within the research. The significance of the research in relation to the problem of quantifying training aims is

discussed followed by thoughts on the findings with regards to the methodology, the tools and techniques for ‘proof of concept’ design to the future evolution of the development of training technology and training programmes. The structure of the thesis can be seen more clearly in Figure 3 where each ‘box’ describes briefly what specific areas are discussed in the thesis.

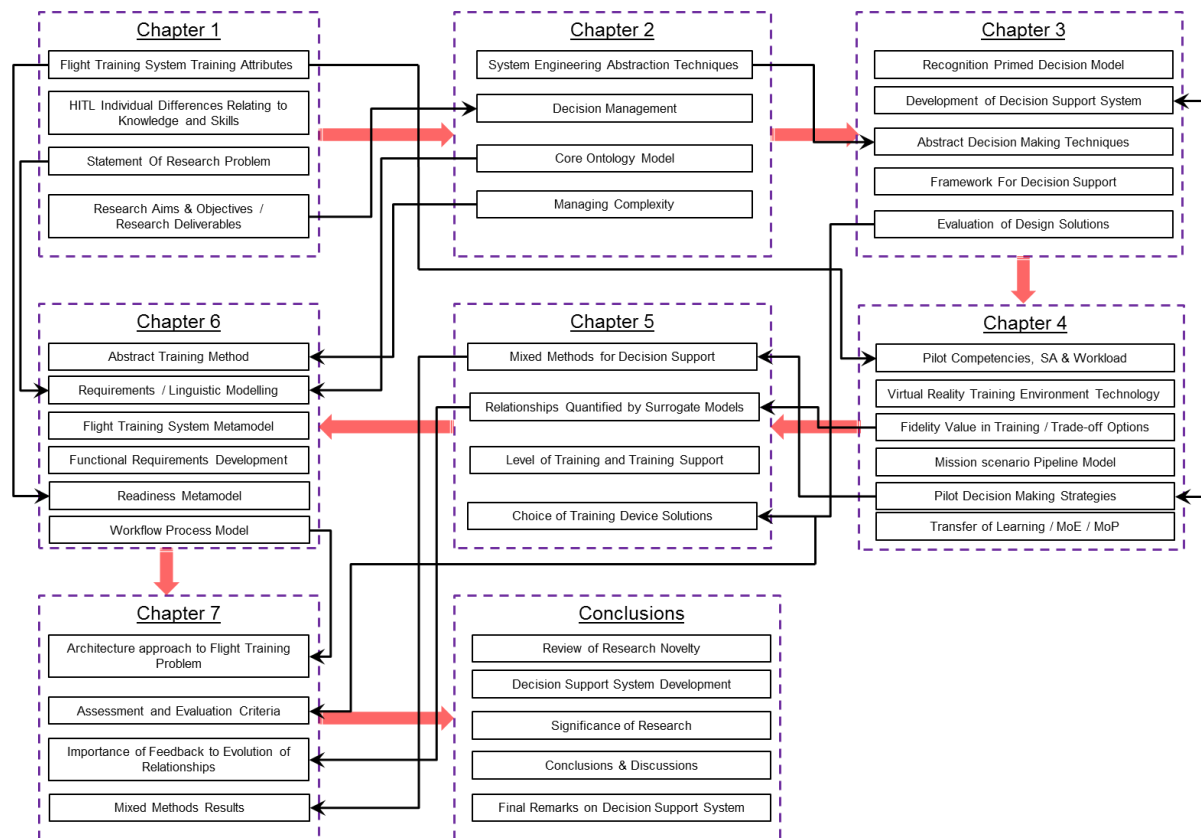


Figure 3 Structure of Thesis

In this chapter, it has been established that the research methodology adapted for the PhD project is directly related to the research techniques and engineering techniques at the Scientific / Interpretivist boundary. Ultimately the research outcome will describe methods and a structure that can blend the precision of engineering with the intuition and knowledge of SMEs to support decisions about Live / Synthetic mixes for FTFoS design. Chapter 6 begins discussions of the methods used in the proposed methodology; and the architecture used to develop the workflow process for the methodology is elaborated in Chapter 7.

WIDER SYSTEMS ENGINEERING CONTEXT

Outline of Chapter

This chapter introduces the concepts and methods of using a systems design approach to model a complex system to gain knowledge and understanding of the system domain problem for the purpose of producing the methodology described in this thesis. The complexity surrounding the FTFoS in the research problem will drive the systems context of the solution to be much broader and deeper in scope than the technical processes of the ISO15288 standard for Software and Systems Engineering (ISO, 2010). In this chapter, a brief overview of using abstraction methods is presented for removal of unnecessary details from the architecture and to decompose the system representation for ease of managing complexity in system design; along with discussions about the advantages of using a common ontology model with both MBSE and MDA techniques for the development of the systems architecture concentrating on the identifications of relationships between system entities. Thus, this chapter introduces the wider system context within which the systems design approach will be developed.

Elaborations of the concepts and methods of the systems design approach within this broader context using MBSE and MDA to prescribe an executable methodology aligned with ISO15288 can be found in subsequent chapters of this thesis. For example, Chapter 5 introduces a structured analytical framework for characterizing and evaluating design alternatives for the purpose of optimising architecture designs; and Chapter 6 prescribes methods for logical semantic modelling (Chapter 6.1), requirements modelling (Chapter 6.2), and use case modelling (Chapter 6.4).

Systems Approach to Solving Problems

A systems approach to the problem of identification of relationships between entities within the system of FoS/SoS is to identify the interactions of and between the entities (Regli et. al., 2014). The approach bounds the problem and analyses it in specific component pieces concentrating in the modelling of the system or SoS at various levels of abstraction. This technique encompasses balanced judgement about the benefit, cost, performance and safety of any proposed system or system change. This evaluation takes a pragmatic approach to dealing with intended functionality and constraints of the system. The elements of the system

are modelled in terms of their independent attributes and the formal mathematical relationships between them (Eltag et. al., 2015). The relationships within a system can be expressed in a hierarchy of system elements describing the architecture of the system from the holistic viewpoint through to component level at various levels of abstraction dependent on design interests (Levenchuk, 2015). The system elements can be configured in several ways including: hardware, software, data, actors, processes and procedures. The interrelationships between the elements are of vital importance in analysing the system as a whole. The manipulation of the attributes values, and performing trade-offs allows for optimization of the system and identification of solution alternatives (BKCASE, 2014). The use of graphical modelling languages and parametric analysing tools can permit the structure of the system to be captured and the subjective region of interest and associated boundaries to be identified and communicated (OMG, 2012). Many of the attributes and relationships between other attributes may not be fully known with certainty; in this case response surface equations (RSE) can be used to describe these relationships, which can be optimized as the system evolves over time (Jou et. al., 2014). By manipulation of the mathematical relationships of the variables, between variables and performing trade-offs studies permits optimization of the system. . In ISO 15288 a system is described as *‘man-made, created and utilized to provide products or services in prescribed environments for the benefit of users and other stakeholders’*. The ISO 15288 standard for Software and Systems Engineering does not prescribe an executable methodology, model or technique for executing a system engineering project; the standard provides a reference model to support process assessment as specified in ISO/IEC 15504-2:2003. The reference model is in relation to requirements for a number of processes suitable for usage during life-cycle of a system or system-of-system

Human factors research is a multidisciplinary field devoted to optimizing human performance and reducing human error; it incorporates the methods and principles of the behavioural and social sciences, engineering, and physiology (ICAO, 2014). Human factors are the applied science that studies people working together and aspects of HITL with technology. It is recognized that inadequate system design or inadequate operator training can contribute to individual human error that leads to system performance degradation (FAA, AC 120 51E, 2004). Current engineering management culture, which is largely data-driven, there is an affinity to look directly at a group of engineering parameter metrics to provide the focus for human effort concerning performance control and improvement (Numrich & Tolk, 2010);(Barite & Saviano, 2014). Hence, the well-known quoted proverb ‘you can’t manage

what you can't measure', which leads to a quantitative process control as best practice. Gray (2001), provides insight in developing methods to systematically identify, measure and manage subjective data which influences organisational performance; the direct conclusion of the discussion surrounds the integration of human intellectual ability in the organisation with technological and physical components that ultimately affect the effectiveness of an organisation. Edvinsson & Malone (1997), consider the subjective factors to be variables driven by 'hearsay', intuition, gut feeling and inside information. However, most complex organisations are becoming increasingly knowledge intensive requiring more feedback from employees/colleagues, suppliers and end users to integrate into organisational processes to improve general functioning and strategic decision making (Johnson, 2012).

The development of the system considers decision management to provide a structured analytical framework for characterizing and evaluating design alternatives for the purpose of optimising architecture designs for lifecycle processes (Miragila, 2014). The decisions are based on trade alternatives centred on requirement criterion and objectives, which can be conflicting objectives for optimised system design. The alternatives have to be ranked, via a suitable selection model developed exclusively for the trade-off study. The decision rationale is generally based on performance criterion that should be robust and repeatable for different system design projects. Within this thesis a FoS and SoS are comparable problems to the development of complex systems, where technology system solutions require trade studies to be performed to optimise organisational objectives.

2.1 Theory of Systems Design

Systems engineering is embracing the challenge of complex systems and FoS/SoS concentrating mainly on human centred and organisational domains. Ferris (2008), suggests systems engineering design can be considered research when done in a setting of considerable complexity and lack of clarity of objectives. In essence, systems engineering research is best described by combining techniques from disparate components of management science and systems practice into a methodology through the process of trial and error (Muller, 2013). Systems engineering requires a number of considerations (Madni & Sievers, 2014) including:

- An understanding of the context of the project including requirements and aims.
- An understanding of the concepts and theoretical principles on which the design will be based.

- Consideration on the design of alternatives.
- Consideration on how the design will be validated, preferable at an early stage within the design process.

By understanding alternative philosophical viewpoints the research being accomplished can critically evaluate the design and thus making it more rigorous for all the end users of the system. Hence, system engineering has a flexible infrastructure which is moreover reliant on the ontology view that is taken through the design process (See Chapter 2.3). Models within the systems paradigm are similar to the concepts of object-oriented programming, as both object and system models share the concept of internal state. The aim of modelling should always be focussed around the objectives, management, or control (Buhari & Mohammed, 2014).

‘General Systems theory is a name which has come into use to describe a level of theoretical model building which lies somewhere between the highly generalized constructions of pure mathematics and the specific theories of the specialised disciplines’
– (Boulding, p.197, 1956)

In system design, the system does not exist yet and the objective is to implement one which encompasses the desired functionality. Furthermore, the task is to create a system structure containing a number of system parts or components; these parts/components are interconnected by the relations specified by the system structure, with the end result of the system behaviour being as desired (INCOSE, 2015). The objectives of system design are the measurement of effectiveness (MoE) of the system in accomplishing its goal(s) that are required to evaluate the design alternatives. Thus, the model must include attributes whose variable values are calculated during model execution. The construction of the system specification is the foundational activity of modelling and simulation (M&S) where concentration is given to establishing and capturing the relationships between system descriptions (Grogan et. al., 2015). The hierarchical representation of system structure provides a way to represent and work with these relationships and system elements can be related by morphisms at each level of the hierarchy. Homomorphism is at the state transition level where mapping between states is preserved between the state transitions and the outputs (Dickerson & Mavris, 2013).

2.2 What is a Model?

“Will one size fit all to provide adequate performance, predictability and safety whilst managing complexity, affordability and scalability.”

- (Systems Engineering Research Centre, 2010)

A model is perceived as any physical, mathematical, or logical representation of a system, entity, phenomenon, or process. However, the most simplistic explanation of a model is a system specification that has a sound mathematical foundation with defined semantics (Holt et. al., 2014). In the context of M&S the system specification is done at the state transition and the coupling of components levels (level 3 and 4 respectively), with respect to Klir’s generative and structure levels (Klir, 1991). A model can be considered as an approximation, representation, or idealization of selected aspects of the structure, behaviour, operation, or other characteristic of a real-world process, concept, or system (IEEE 610.12-1990), as seen in Figure 4. The system description is developed from what the model designer knows about the real world problem including assumptions on aspects of the problem that knowledge is currently absent; this forms a partial description of the real world. A conceptual model is then produced using abstraction techniques to model relevant details of the real world that is desirable. The model describes the functionality gathered from the system description; once the behavioural model is verified with the requirements, the methods that describe the operations of the real world are then added to the model. The designed model is then transformed into code for simulation and execution which describes the computer representation of the real world object.

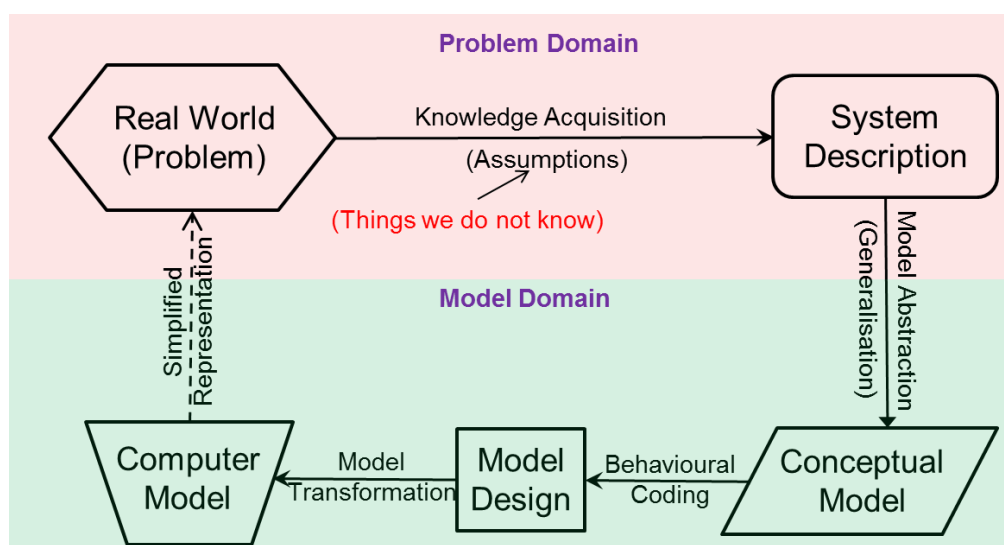


Figure 4 Concept of Modelling Real World Entity(s)

Formal system models introduce an aspect of rigour and flexibility that is both machine and human readable to establish a common understanding. The common concept of a simulation model incurs the following characteristics: a set of – instructions, rules, equations, and/or constraints – for generating system behaviour. Therefore, a model is created with a state transition and output generation processes to accept input paths and generate output paths depending on its initial state setting(s).

To allow for processor and memory limitations of computers, reduction of the complexity in models is generally needed to allow the model to be executed in resource-limited simulators. But the simplified model must also be valid, at some level, and within a current frame of interest. The issues of time become apparent when human-in-the-loop (HITL) simulations are to be valid (Kleer et. al., 2014).

Abstraction is an important method (or algorithm) in system design to reduce complexity while preserving validity in a frame. It permits viewpoints of models to be produced with varying degree of abstraction to support system design. Evidently, abstraction is used to preserve resources; more detailed the model, the greater the resources required (Evora, et. al., 2013). In a virtual simulation, scope is how much of the real world is represented; resolution is the number of variables in the model, their precision or granularity. The trade-off between model abstractions against cost is a constant concern when developing a simulation model. For example, benefits in reduced run-time and memory requirements may be accompanied by an inevitable loss in predictive accuracy. The benefits of abstraction include a more rapid analysis of the model at lower cost (Quadri et. al., 2012);(Horsinka, et. al., 2014).

Two of the main tasks of systems theory and design is to provide: recommendation on the level of abstraction needed to adequately represent the real world problem and a common language by merging modelling and communication; where the model provides precision while communication provides comprehension, both of which are needed for the design of systems. One of the main challenges of a model is requirements are expressed in natural language and are thus open to interpretation. To enable the interpretation to be explicit, conceptual models are created that formalizes the requirements to a given standard (Rolland, 2013). The UML and SysML modelling languages were designed to assist in the conceptualisation processes (Zoughbi, 2010); however, once deemed to be a formal language, the model artefacts produced, unless strict ontology rules are followed and clarity of description within the models are standardised, can still be interpreted not as intended.

2.3 Model Ontologies

Technology and management techniques are becoming increasingly intricate with data and information management becoming more of a concern for an efficient training pipeline. Unified information technology framework(s) proposed in this thesis are required to support a number of domain training needs simultaneously using a common semantic design ontology that can be advanced to model the entire FoS within the FTS. These models can be used to deal with complexity in a visual approach providing a primary data source using a common language that is easy to understand between all relevant stakeholders.

Object-oriented modelling techniques enable architects to create abstractions that are domain- and application -specific and are seen as a means to manage complexity in systems and software engineering projects. The models used are abstractions of reality that are prescriptive (include rules and directions) to form templates which the system is later implemented, and descriptive (for shared understanding of relationships and roles) to form a semantic elucidation of the architecture of the model (Henning, 2013). Ontologies are special types of models that utilise a form of partial description or under specification as an important means of abstraction. The ontology is a form of metamodel that describes how to build the system models to ensure that the domain is described as completely as possible and redundancy in the models is reduced, i.e. all system elements/objects must have a role to perform in the system, have attributes and behaviours governed by rules relevant to system goals, and identified relations to one or more system elements. An ontology model is similar in design to a domain model developed in the Computational Independent Model (CIM) within the Model Driven Architecture (MDA) methodology (Magableh et. al., 2012). The metamodel defines a specification of a human-readable semantic notation for elements and relationships that describe the design conventions required for the system models; in addition, a metamodel can be used for logical modelling of system requirements to identify and clarify understanding of relations between qualified constituents of the stakeholder's requirements narrative (N.B. there are other definitions available that are domain specific definitions). The basic core ontology meta-model used in this thesis, illustrated in Figure 5, an advancement for the suggestion made by Simperl (2009), is used to control the complexity of the model artefacts used in the architecture of the conceptual model of the FTFoS Enterprise described by the requirements narrative and designed to be reusable for other domains where complexity of design is required to be understood.

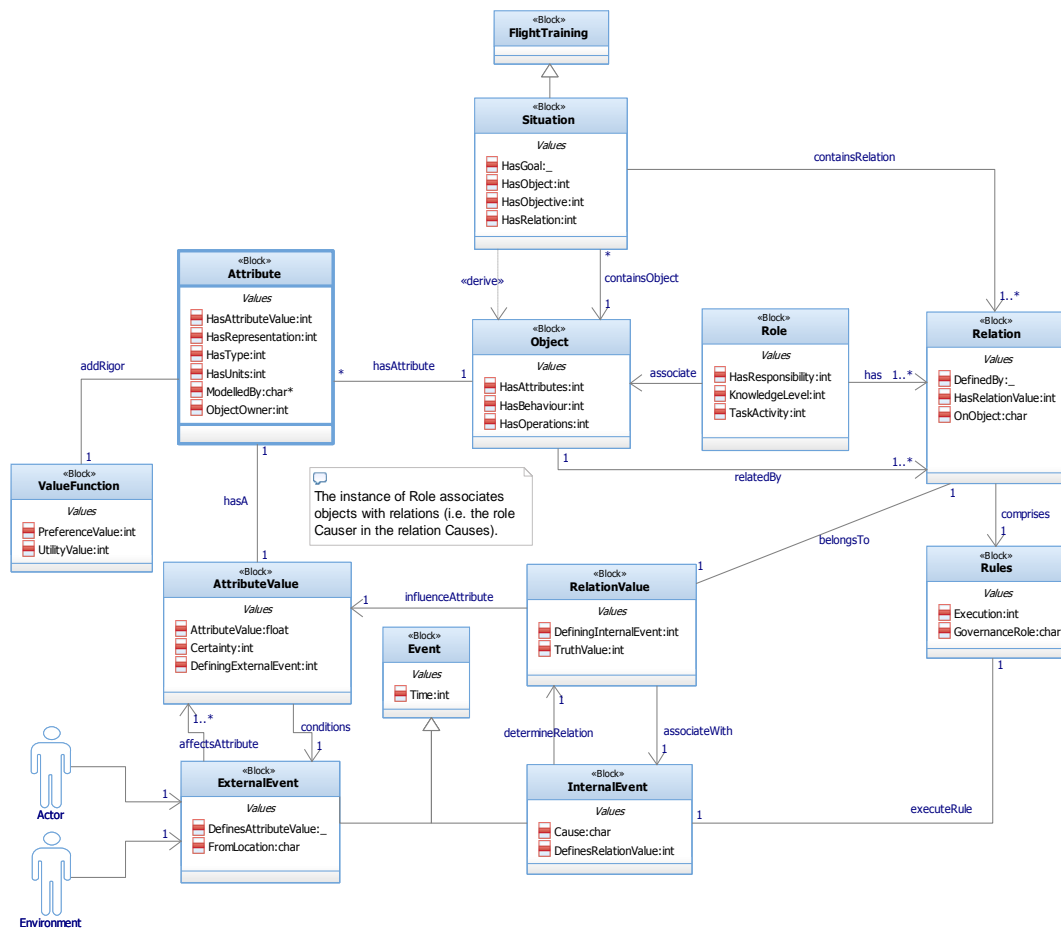


Figure 5 FoS Core Ontology Meta-model

The ontology initially concentrates on domain specific knowledge, i.e. the flight training system and examines key parameters for the design process using abstraction techniques to keep the model uncomplicated. The ‘Situation’ block describes aspects that are required to be present for the development of mission scenarios but also to ensure properties to enable traceability through the architecture to requirement goals. The situations, which define circumstances that an entity of a system has to cope with, are integrated as objects and participate in relations. The object, an instance of a block, describes a real-world entity by the means of attributes, operations and behaviours; the attributes of an object must add value to the system and the object must perform a role within it. Each attribute has a property (e.g. type, unit, representation, etc.) that needs to be identified before any system model is produced to ensure full interoperability when models produced by different departments or businesses are integrated together. Any identified object in the architecture must interact with other objects in the system (or with itself) for the benefit of the system in achieving its goals; if no interaction is identified then a decision of whether the object is needed within the system is required. The interactions are described by relationships, which must be known by the architect, to provide meaningful knowledge of the specifics of the interaction (i.e. the rules),

why it's needed (i.e. the value of the relation) and to understand the effect of an attribute change in one object has on a receiving object through an interaction. The object attribute values are affected by internal (time or behavioural) and external (human, environment, or other systems) events, established by the relationships between objects (generally modelled by sequence diagrams), which can alter system states. From the events, the system is required to meet the goals and objective constraints gathered from the requirements. The ontology model should permit a greater degree of governance in ensuring that the architect models what is required using abstraction as the means to minimise detail in the model but still maintain important aspects of interest of the real world system and permit animation or simulation of the model(s) for correct behaviour for verification of requirements.

The ontology model provides a mechanism to control the design process and give an overarching viewpoint through the design and communication with various domain experts. It is perceived that the ontology model and associated rules are hard to keep consistent within the design process (Pirro, 2010) when behaviour of the system model is being consistently amended to ensure requirements are satisfied, this notwithstanding the ontology model can be used to maintain as much as practically possible the design context of the UML or SysML architecture. Issues of geographical location of engineers involved in the development of the system adds a further challenge, as keeping with ontology rules reduces flexibility in the design process and attribute properties are difficult to manage to ensure zero interoperability issues between models.

2.4 Complex Integrated Systems

SoSE is considered a multidisciplinary area that on the technical facet includes systems engineering specialities and software and information management specialities (Horvath & Rudus, 2014). As these systems become more complex and extensive, unique aspects of systems working on the edge of chaos requires more innovative approaches for SoSE (Northrop, 2006). Lane et. al. (2007), describe the term 'system of systems' as representing many things to many different people and organisations. In the business domain, a SoS and FoS can be seen as the enterprise-wide or enterprise integration and sharing of core business information across functional and geographical areas. In the military domain, a FoS/SoS is a communication infrastructure and a configurable set of constituent systems to support operations in a constantly changing environment (Ricci, et. al., 2014). Additional viewpoints include a FoS/SoS being an architecture that evolves overtime, often driven by organisational

requirements, new technologies, and available budget and schedules. The evolutionary FoS/SoS architecture is more of a network architecture that is reconfigured, evolves and develops with needs and available resources. SoSs and FoSs are unique as they are comprised of constituent systems that possess the following characteristics (Sage, 2001), (DoD, 2008):

- Operationally independent – all of the constituent systems can perform useful functions both within the FoS/SoS and outside the FoS/SoS
- Managerial independent – all of the constituent systems are managed and maintained for their own purposes.

As complex systems become more of ‘the norm’, the associated need for flexible, adaptable, agile development processes to deal with rapid change and to facilitate innovative solutions in ultra-large solution spaces, are needed (Boehm, 2006).

Modern defence operations are characterised by the requirement to integrate individual systems, platforms and infrastructure to validate the operational requirements. There is also a strong need to introduce new technologies rapidly where they can add benefit (HoC, 2013). The designed systems architecture represents the description and the design of the complex system; in essence, a system’s architecture is a ‘blueprint’ of an actual system. It deals with the overall functionality of a system at a level of abstraction that is useful to the architect and adds value to the system verification and validation method.

“Architecting deals largely with unmeasured-ables using non-quantitative tools and guidelines based on practical lessons learned; that is, architecting is an inductive process. At a more detailed level, engineering is concerned with quantifiable costs, architecting with qualitative worth”
 -(Maier, 2000)

Moreover, systems architecting is a process driven by a stakeholder’s requirements; if a system is to be successful it must satisfy a useful purpose at an affordable cost for an acceptable period of time. An architecture consists of a number of related views of the system under consideration. These views and their associated relationships are defined by an architectural framework (Benkamoun et. al., 2014).

“The fundamental organisation of a system embodied in its components, their relationships to each other, and to the environment, and the principles guiding its design and evolution” - ANSI/IEEE Std. 1471-2000, Recommended Practice for Architectural Description of Software-intensive systems

The systems architecture sets out the context of how the system operates within its immediate environment. Well understood architectures are used to provide a clear context for the insertion of new technologies and the architect must also understand how the system or FoS/SoS will operate and the environment in which it will function in otherwise the system design will be incomplete and emergent behaviour will be of concern once the system is operational

2.4.1 Model Driven Architecture.

'Model-driven Engineering, is an emerging new paradigm in software engineering which bases system development on (meta-)modelling and model transformations, and provides methods to build bridges between similar or different technical spaces and domains'
 – (Gargantini, p.1, 2010)

Model Driven Architecture (MDA) developed by the Object Management Group (OMG[®]) is the international standard for system specification and interoperability based on formal models (D'Souza, 2001). It represents a coherent synthesis of a number of well-established and proven software engineering techniques. The process provides a level of sophistication that places software engineering on par with other, more mature, engineering disciplines, such as hardware, aeronautical and civil engineering. The MDA process concentrates on the separation of the specification of the operation of a system from the specifics of the system's platform (Muller & Mukerji, 2003). This separation technique ensures, as much as reasonable practicable, the issues of entangling functionality with implementation, generally encountered in standard parametric approaches to systems design, is reduced. Concentration is then given to core solution of the problem and incurs the ability to separate this from the technology based implementation of the solution. The viewpoint on the system is an abstraction of supressing selected details to establish a simplified model and thus lead to concentrate on important system characteristics (Dumitresco et. al., 2013).

The process of acquiring expertise, capturing it in a form, which is uncontaminated by other subject matters and making it accessible to others is the essence of the MDA process (Raistrick, 2004). Subject matter partitioning is one of the best accepted and mature partitioning strategies. A vital system design decision is the choice of the software implementation technologies on which system will be built: - programme languages, operating systems, data storage technologies (databases), etc. The MDA domains and associated model descriptions can be seen in Figure 6.

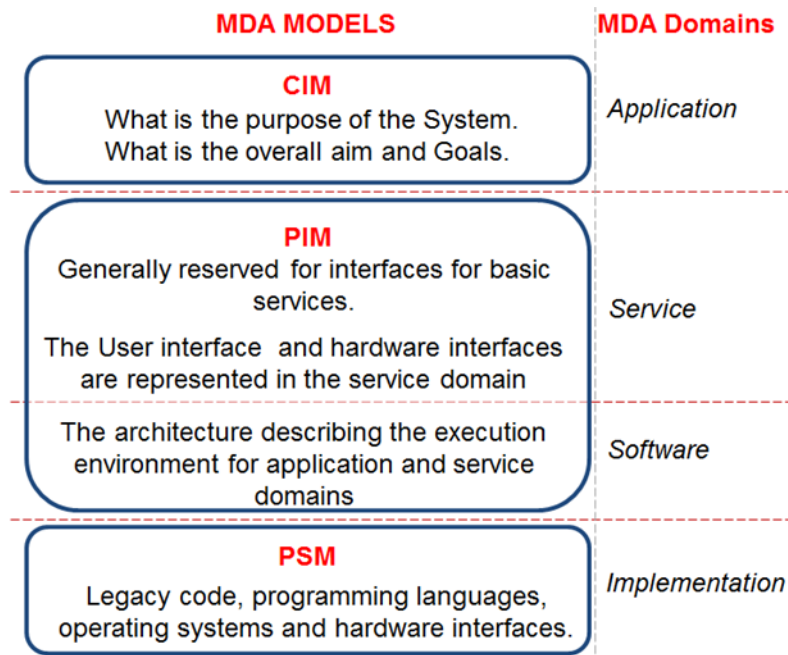


Figure 6 Four Domain MDA Partitioning (amended from Elammari & Issa, 2013)

The Computational Independent Model (CIM) is used to capture and model the concept of operation of the system as well as detailing how the system will interact with other systems and the external actors/users; additionally the CIM will consider system goals, requirements, stakeholder needs, and business rules: generally modelled using Use Cases. The diagrams will then be used for the next layer of the design process to specify in greater details the system in the PIM view. Here the system 'HOWs' with reference to capabilities and/or functions are defined, which captures the functionality and structure required to permit the implementation of that functionality across different technology platforms. The PIM concentrates on the operations of the system, whilst hiding the platform dependent implementation (i.e. behaviour of the system is captured within the designed architecture using classes/blocks, state machines, sequence diagrams, etc., whilst implementation specifics are abstracted out). This method should ensure the designed behavioural specification, within the architecture, does not change between platforms and is uncontaminated by knowledge of other system aspects (although there are opportunities for research in model interchange between platforms) (Alanen & Porres, 2010). A multilevel PIM view is necessary to capture the level of detail in any single model. In MBSE, the PIM view decomposes the business rules and system requirements to specify a detailed model of the system; as a result the capabilities of the system are more thoroughly defined and derived. The next stage of the MDA process is to develop Platform Specific Models (PSM), which uses the behavioural models or artefacts in the architecture developed within the PIM stage and adds platform specific details within the implementation of operations within transitions and/or states within

the architecture (e.g. Java, C#, LabView, Matlab etc.). The final stage of the MDA Process is the generation of a Platform Model using model transformation of the PSM into platform specific code. This model transformation concept is the key to the MDA process (Kriovile, et. al., 2015). The role of model transformation in MDA is to flow between the stages of the process, i.e. from CIM \rightarrow PIM \rightarrow PM whilst preserving entity relationships between the different models; the conceptual integrity of the system is also preserved.

'Model Transformation is the process of converting one model to another model of the same system'
- (Miller and Mukerji, p.2-1, 2003)

The transformation process relies on the ability to preserve the mappings between the models. However, the mapping between the CIM and PIM is a manual process and is prone to error and continuity issues are common place (Kriovile et. al., 2015). One of the key advantages of MDA is the rigorous and repeatability of the approach and the removal (as much as practically possible) of most of the ambiguity involved within the systems design process (Hazard, 2014). It permits semantic capturing at a high level of abstraction through to component level to software and is defined with a robust process and flows with concentration given to relationships between them. Hence, traceability through the initial requirements elicitation process through to the PM provides architectural integrity of the satisfaction of requirements and aids in verification of the design architecture.

2.4.2 Brief Summary of MBSE Constraints In Proposed Methodology

"The UML is simply a standard diagramming notation – boxes, lines, etc. Visual modelling with common notation can be a great aid, but it's hardly as important as knowing how to design and think as objects. Such design is a very different and more important skill, and is not mastered by learning UML notation or using a CASE or MDA tool. A person not having good OO design and programming skills who draws UML is just drawing bad designs."
- (Bell, 2004)

The advantages of using model based approaches for design and verification has not been fully explored: model based approaches such as UML and SysML are used to capture semantics and notation for object oriented problem solving but these models are left behind once the initial project is completed (Heena & Rangna, 2011). The understanding of MDA in developing software and systems means that the system can be developed using PIM with concentration given to the functions of the system (i.e. black box design), thus, modifying how a system works involves revision of functions, furthermore, it allows the models to be used throughout the lifecycle of the system to enable reuse. However, the execution aspect

of the parametric models and algorithm development within the modelling environments themselves is limited: parametric diagrams have a key role in systems engineering modelling (Sakarri et. al., 2013). The realisation that the modelling languages, moreover SysML, is complementary to the popular simulation tools in the market, i.e. Matlab and CAD/CAM, is something that has not been fully explored (Vogel-Heuser et. al., 2014); being involved in the new ISO-15288 (software and systems life cycle process) standard and INCOSE, the modelling constructs and tools are continually improving as the community is moving forward with MBSE initiatives (ISO, 2015).

One of the issues with the use of modelling language is that they are a ‘language’ as a result is open to different semantic interpretations of requirements and modelling structures. It has become apparent with the experience gained from working with relational orientation that without a robust design methodology offered by MDA and the complementary ROSE and ROSETTA frameworks (Holden, 2012) to bring precision in modelling, faith in these modelling tools will decline. The tool set available for MBSE design are varied, most tools vendors offer the basic modelling constructs and fairly recently began to offer add-on software for basic requirements verification, but at an additional cost. The major issue encountered is with regard to the support for the tool itself. Rational Rhapsody has been obtained through the academic initiative, but the tool itself is not user friendly and the support from IBM is expensive. The SysML profile within the tool is used for requirement elicitation and knowledge acquisition during the early stages of the systems design process. Due to limited simulation capability of the MBSE tools, the PIM model, which describes the behaviour of the system, is transformed into the PSM using a parametric analysis tool (Labview) for verification of the Workflow model and associated ROSETTA frameworks.

2.5 Representing FTS Systems with Models

The FTS integrates disparate technology and human social interaction with the technology and training mission goals. Buchanan and Huczynski (1997), describe various opportunities applying an empirical, natural sciences approach for this type of socio-technical organisational system problem that are inherently difficult to quantify using robust mathematical methods. Senior (2002), advocates a ‘soft’ systems approach to gain a clear understanding of the complexity in socio-technical and organisational systems. Empirical reductionists’ approaches, as described in Checkland (1999), seek to abstract complex systems into a series of controllable and measureable variables to yield scientifically

reproducible findings. The emergent behaviour of the interactions external to the system boundary is also complex to model and understand; and the decision of what levels of abstraction are sufficient to gain an all-encompassing view of critical attributes and operations, are important factors to study for this type of complex socio-technical system. Clear and concise analysis is needed with all affected parties (i.e. stakeholders) involved in decisions to ensure accurate trade-offs that can impact both the FoS level and the organisation process is known to all (Friedman, 2005).

2.5.1 Managing Complexity of the FTS SoS / FoS Problem

For management of the systems in the FTFoS, strategic planning and decision making techniques for recognising and controlling the integration of disparate complex systems into a FoS, is needed. Mastashari et. al. (2012) recognised and analysed the roots of errors when multiple / large numbers of stakeholders are involved, which is common place in SoSE, some of the key recommendations made are as follows:-

- Understand goals in concrete terms, balance contradictory or incompatible goals, and establish priorities before focusing on planning and gathering information for a solution.
- Avoid ‘economising’ up-front since this encourages one to omit crucial steps through the process – Clarifying complex relationships among variables (or system components) before down-selecting to the variable(s) or systems of interest may avoid problems of ‘unintended’ side-effects or long term repercussions (undesired emergent behaviours).

‘...as the complexity of a system, increases, our ability to make precise and yet significant statements about its behaviour diminishes until a threshold is reached beyond which precision and significance (or relevance) become almost mutually exclusive characteristic.’
- (Zadeh, 1973)

Systems can be integrated together through interaction forming a SoS or FoS to achieve more functionality than the individual systems operating alone. FoS/SoS are higher order complex systems that are composed of independent component systems. Due to the managerial independence, the dynamic, time-dependent composition of the FoS/SoS, management of this complexity requires intensive decision making during all phases of the systems engineering lifecycle (Verbeek, 2013).

This research should provide some insights into how declarative, procedural, and structured knowledge impact performance and how problem solving can be measured more effectively. The methods within the thesis should identify in a holistic way, how to use information more effectively to drive future technology requirement needs. Partial knowledge about the current state of a system is held by SME's and the difficulty is unifying such partial views into a more unified whole (collaboration in confidential areas, such as military pilots, is further impeded by International Traffic in Arms Regulations (ITAR) and Intellectual Property (IP) restrictions). M&S comprises complex and multifaceted set of activities that require abstractions of information in order to build a valid model. The phases of the process include: clarifying objectives, collecting data (knowledge of how things work, or observable numerical data), model building (architecture design and behaviour), and model validation; the process can be simply illustrated in Figure 7.

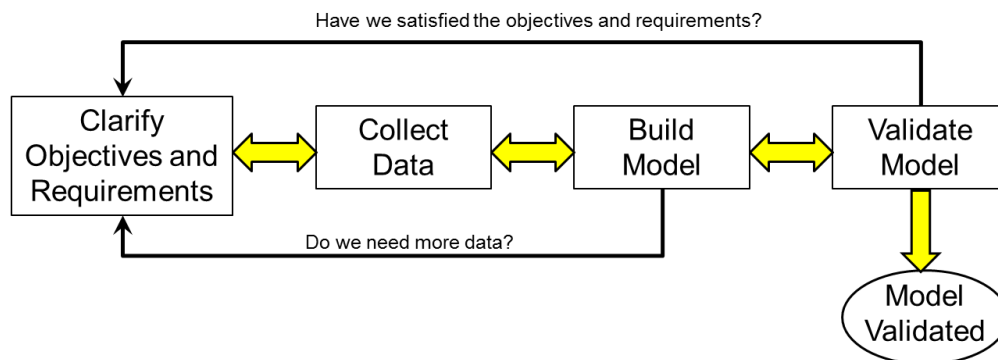


Figure 7 Abstract Representation of the Model Construction Process

The problem space and the stakeholders' needs are examined before the formal definition of the system of interest is developed. This defines the operation aspects of a potential solution from the stakeholder's viewpoint, independent of any specific solution and should describe what the solution should accomplish and how the solution is defined and developed (Kossiakoff & Sweet, 2005). From the stakeholder analysis, a mission analysis is performed to understand the socio-technical and –economic context in which potential problems or opportunities reside (see Chapter 6). This task is generally performed iteratively with stakeholders' involvement to gain a more robust understanding of the problem space. Analysis of the problem space is performed in relation to organisation needs, capability gaps, or opportunities and solutions that can be used to evolve organisational strategies for its business objectives; in the military context mission analysis is referred to as concept of operations (ConOps) (ESA, 2008). The mission analysis is based on fundamental concepts, such as states of the system, scenarios (of actions), functions (identified in use case diagrams),

etc., for more information of ConOps and Mission analysis, refer to ISO/IEC (2011). The steps and position of the stakeholder requirements and system requirements in the engineering cycle can be summarised by Faisandier (2012), which illustrates the cycle of needs required to permit a system to satisfy them and to ensure the architect obtains enough knowledge on the System-of-Interest (SoI) before developing a solution to the problem.

With the system needs identified, the system is required to be defined in detail. This process is executed iteratively and recursively and includes the definition of system requirements, (functional, behavioural, and temporal models), the physical (hardware, software, human roles, etc.) architecture, and system analysis. In each iteration of the process, a gap analysis is performed to ensure all system requirements have been mapped to elements in the architecture and the inputs of the system are identified and realized. The interactions between the system elements are defined by interfaces which are dependent of the structure of the architecture. The architecture design has been heavily influence by ISO/IEC 15288 (2008), which gives an implicit view of architecture generally defined in a hierarchal structure. ISO/IEC/IEEE 42010 (2011) gives a useful viewpoint of the structure of the architecture considering stakeholder concerns, viewpoints, and lifecycle modelling (Maier & Rechtin, 2009).

The design process provides an association between the system architecture and the implementation of technological aspects of the system elements that compose the physical architecture of the system. The analysis of the system architecture permits quantitative assessments of the system to design choices and provides a rigorous approach to technical decision making (Schmid, 2013). The analysis uses M&S, cost analysis, technical risk analysis and effectiveness analysis. Analysis of the system is a critical task and provides a basis for assessing solution alternatives based on assessment criteria gathered from the first two stages of the system engineering process. In complex systems or FoS/SoS, ensuring the integration of individual system elements to function properly as a whole to satisfy the design requirements and prepare the system for final validation against requirements, is a must (Valbuena, 2013). Integration consists of a progressive process of linking elements that compose the SoI and checking the correctness of static and dynamic aspects of interfaces between implemented elements (DAU, 2010).

Verification can be considered as checking the ‘correctness’ of any system element or interconnection (integration) of system elements (i.e. building the system right). The

verification procedure is performed in parallel with the system definition and system realization processes and is defined by the objective evidence that specified requirements have been fulfilled (ISO/IEC/IEEE 15288, 2015). The purpose of verification is to identify defects within the system at any time of any transformation of inputs into outputs. Verification is used to provide information such as, inspection, testing, analysis, etc., that the system under consideration can demonstrate satisfaction of requirements with no erroneous error occurring during the integration stages. Validation is strongly linked to verification tasks, and is concerned to prove the system has the right features to produce expected effects (i.e. building the right system). When building the model with the collected data, we need to partake in a recursion back to the objectives to ensure the designed model captures all the parameters in the requirements. Once confidence has been gained in the model, the process of validating the model can begin. Once this has been satisfied, the designer can, with confidence, commission a valid model which the simulator can execute (Hussain & Senere, 2013).

‘A model may have demonstrated an adequate level of practical utility by repeatedly producing satisfactory answers to real-world engineering questions. Ultimately it is the user, not the model developer, who decides if the model has sufficient utility. To determine whether or not this is true, the user needs access to comparisons of model predictions and experimental results relevant to the applications of interest’
- (Sargent, 2004)

Model credibility is primarily concerned with the confidence that the model and the information generated from the model is credible. Moreover, verification and validation (V&V) considers factors such as model recipients’ (researchers, managers, decision makers, and policy makers) understanding of the models assumptions, outputs; and the complete understanding of the constraints that the model is designed for. An important aspect of modelling is to emphasis which dimensions in the model are fundamental and which are incidental. It is also important to distinguish between qualitative findings and quantitative results that inherently carry with them large margins of error.

The validation phase concentrates on the ability of the model to provide predictions within certain constraints; in the FTS these constraints include: human, phase of flight, cockpit configurations, or types of procedures, weather conditions, etc. The validation phase approach is to model a baseline scenario in which some type of human performance data from executing a mission scenario using a chosen blending mix is available. As such, the model can only be validated against a given data set. The comparison with the ideal model requires

metrics (goodness of performance) and a tolerance (is the performance good enough). When the performance is outside the tolerance, the model trajectory is judged not to qualify as valid within the particular frame. If the model fits within tolerance, the model validity is confirmed and performance is of acceptable standard. Statistical techniques are employed if comparison involves further consideration and calculation of their variance in their output values.

2.6 Chapter Summary and Relationship to Research Problem

The complexity surrounding the FTFoS is the main reason for further understanding of how to manage complexity. MBSE approaches tackle the issue of complexity by separating aspects of the FoS into easy manageable ‘chunks’ by identifying system elements from the system description and using abstraction techniques to understand the relationships of and between each system element at a level of detail to which the ‘actors’ / stakeholders of the system can understand. The designed architecture and behavioural models bring a level of precision to the complexity with events declaring the level of interaction between each object (software description of a real world entity). MDA techniques assist in understanding by separation of concerns between each identifiable domain of the FTS and therefore allow each domain to be concentrated on in isolation of others. In this way the human and technical aspects can be investigated separately and then the interfaces between each can be identified and actioned by using M&S techniques to verify operation.

The model based approach being used in this thesis is used to present a proposed methodology to understand relationships between mission planning scenarios and system analysis across disparate domains, including technology and human factors that exists in the training lifecycle. The system analysis concept is to characterize the impact of the system from a set of parameters, set by mission planning, to achieve specific system performance goals. The goal of mission planning is to parametrize an operational mission scenario to improve effectiveness of the blending mix (and pilot readiness) with an indirect intent of assessing some technological improvement for increased efficiency. The result of mission modelling, therefore, is to identify a desired technology to optimize a scenario for specific training attributes for a specific mission goal; and is generally in the domain and responsibility of the decision makers (instructors). Generally, SMEs focus on mission scenarios using their own experience as they do not have the expertise to describe the problem of training effectiveness in a manner amenable to technology system development.

Similarly, a systems engineer (technology focussed) does not have the SME skills to translate mission scenarios into product design specifications.

To achieve some integration between the two domains (mission and system), implementation using a form of MBSE is investigated. The models produced are an abstraction of the flight training system (FTS) concentrating on details described in the requirements narrative used to conceptualise the organisation and describe the system using ontology as the basis for producing the systems architecture used for analysis. A static description of the mission scenario will be created and communication between this model and a dynamic representation of the mission model is described to permit system optimization at the mission level. When the goals of the mission scenario are known and the student current readiness is estimated the emphasis switches the analysis using a high level system representation of the FoS in the system analysis context.

DECISION MAKING TECHNIQUES

'Decision making is knowing if to decide, then when and what to decide. It includes understanding the consequences of decision. Decisions are the means by which the commander translates his vision of the end state into actions'.

– The military Decision-Making Process (Department of the Army, p.5-1, 1997)

Outline of Chapter

The chapter introduces the complex decision making tasks and associated issues with developing a system and/or methodology based on assumption made and commences with discussions on formal decision making tasks, which centres on a model that includes domain expert experience to develop comprehension on a situation. Further discussions on what constitutes a decision support system and how it integrates with the decision making process is presented. The research specific domain is then introduced with the concepts of a decision support system in relation to the advantages of using frameworks to organise information in a structures manner. The chapter concludes with a brief outline on how the decision support system should be evaluated for use in a practical way.

Naturalist Decision Making

Naturalistic decision making describes proficient decision making strategies based on recognition processes (Klein and Calderwood, 1991). Klein's recognition primed decision model, which has been evolved and illustrated in Figure 8 is based on research and feedback conducted in an operational context.

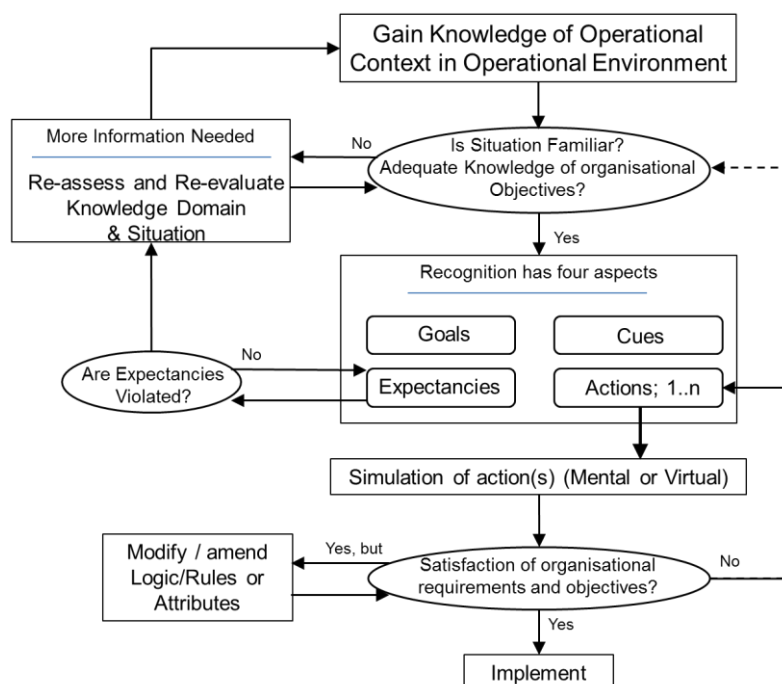


Figure 8 Recognition-Primed Decision Model (adapted from Klein, 1997)

Klein's findings indicates that decision makers base their choices on mainly prior experiences, planning and quick ad-hoc modification of plans to suit current constraints based on mission status; there is no consideration on what the optimal decision strategy should be. Without gaining knowledge about people, organisational politics, and operational procedures it is improbable that a link between the current situation and experience can be made. The level of knowledge should be 'consistently' re-evaluated as the domain model is being developed to answer further questions and to fill gaps in understanding. Once constraints are met, the rules used to produce the model can be assessed to generate improvement to the model outcomes. As with most requirement analysis tasks, decision making tasks may present the decision maker with conflicting or incompatible goals (Pourshadid et. al., 2014). Evolving assessments of the situation necessitates rapid judgment of the situation, prioritization of goals and potential actions, and implementation of appropriate task strategies. Formal decision making strategies in the planning context directly maps to SE methods of iteration and recursion for analysis of system design behaviour for satisfying customer requirements. However, formal decision making strategies in an operational context is deemed too slow and consequently ignored in practice (Thunholm, 2006);(Bushey and Forsyth, 2006).

3.1 The Evolution of Decision Support

The term Decision Support System (DSS) was devised by Gory and Scott Morton (1971) to describe the role of a framework for supporting management decisions using the aid of technology and models within the decision-making process. Finlay (1989) advocates that optimal decision making involves both information and intelligence. The differentiation between technology supporting decisions and actual decision making is distinguished by the way technology provides the ability to influence decisions based on the state of the system(s) of interest (SoI) (Iyengar et. al., 2014). Displaying current attribute values of the system(s) and the ability of technology based on simulations or user inputs to restrict further options for the decision maker can force a reconsideration of actions or current beliefs (Liu & Zarate, 2014). This provides the basis for a more accurate description of a DSS: '*A technology system that postulates information and/or data to facilitate a course of action in harmony with pre-defined objectives to permit trade-offs between solutions*'. There exist conceptual frameworks for a DSS that has their origin in management science, computer science, decision analysis, organisational behaviour and decision research, as shown in Figure 9. The multi-disciplinary nature of decision making makes it a prime motivator for the development

of DSS tools to assist in reducing the number of alternative solutions and make the process more efficient. The production of models allows complex system, organisation designs and processes to be understood at an abstract level and invites the ability to integrate system data within the analysis and has the ability to permit mixed method quantities to be incorporated into the DSS (Karlsen, 2014).

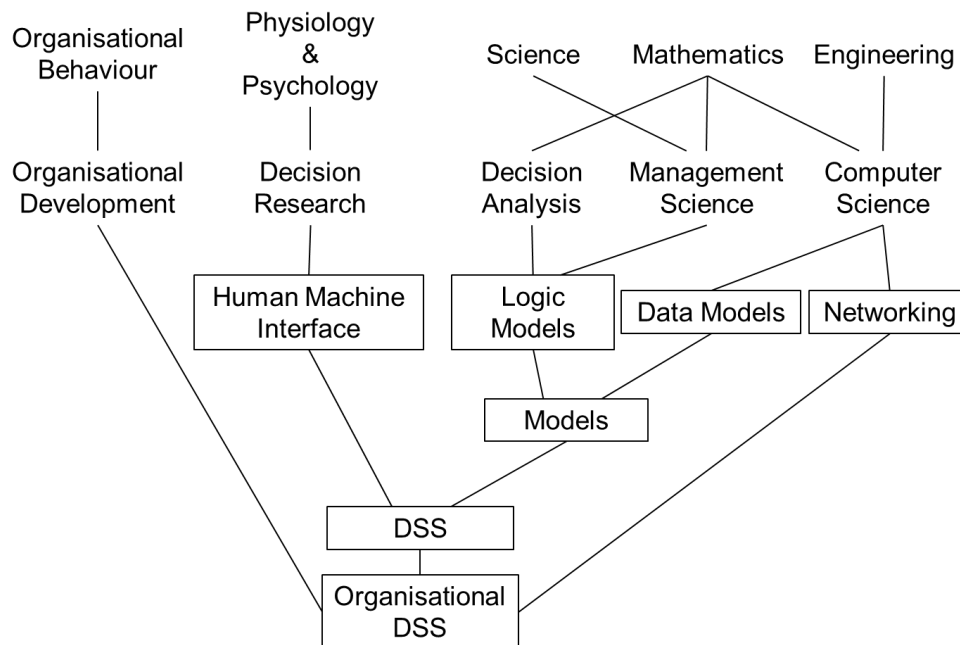


Figure 9 Provisions to DSS development (adapted from Finlay (1989))

The role and importance of a DSS within a large complex organisation has been identified by the research conducted by Nemati et. al. (2002), Bolloju et. al. (2002), Ribino et. al. (2011) and Belling et. al. (2013); where the DSS assists decision makers in solving various wicked and semi-structured problems that involve several attributes, conflicting objectives and multiple organisational goals.

3.1.1 The Decision Making Process

Crooke & Slack (1991), describe decisions as '*a good decision is one where the decision maker fully understands the background, objectives, alternative courses of action and range of possible consequences of a decision*'; the decision making process described by Crook & Slack is illustrated in Figure 10. Observation is the first stage, which influences the problem statement; once the problem is fully understood, the objectives of the DSS can be elicited and derived, which will assist gaining detailed understanding of the problem. Once these stages have been complete it should be possible to identify what courses of action are available and identify decision boundaries. These boundaries can be used to eliminate courses of action,

supported by technology, from the decision making process. It is then required to map the solutions (available courses of action) with the pre-defined objectives. The decision maker then has to choose an available option, implement the decision and gain feedback from the execution of the decision so future decisions can be influenced by previous choices and consequences of the decision can be realised.

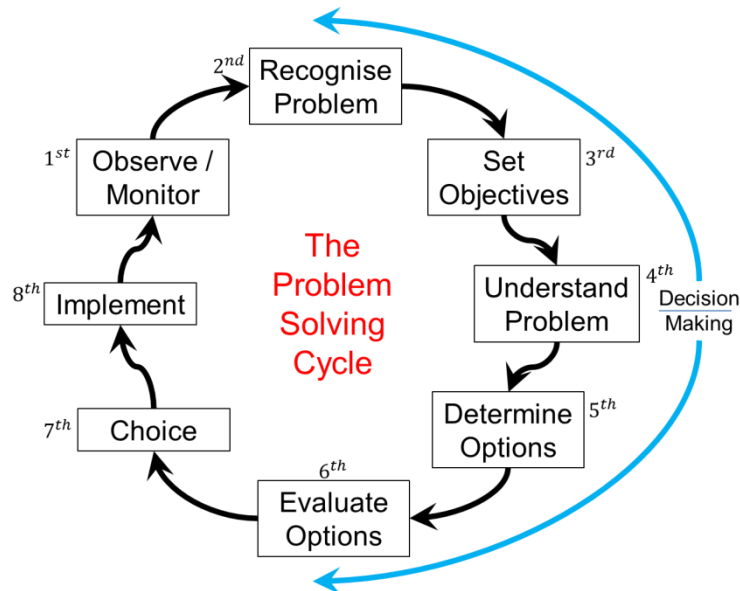


Figure 10 The Decision Making Process (adapted from Crooke & Slack, 1991)

To ensure simplification of the decision making process, the development of the DSS has to consider at least three interrelated characteristics (De-Brain, 2007), :-

- The level of abstraction of the decision (strategic, tactical and operational)
- The dependency of the decision
- The degree of structure and justification of the decision

Boundaries that separate the elements of a DSS can be considered as identical to methods used to develop models within a model based management system (MBMS) (Tariq & Rofi, 2012); where each MBMS is capable of storing, adapting and editing models within their environment.

The DSS should also allow multiple perspectives to be considered to resolve ‘wicked’ problems, which result from ‘big data’ of knowledge and information within complex organisational processes. These perspectives can have drastic implications for the design of the DSS (Courtney, 2001). These perspectives can be summarised as:

- Optimised viewpoint – focus is on the logic of optimal choice.
- Process Oriented viewpoint – focus is on constraints during real-world operations.

- Organisational viewpoint – focus is on the structure of the organisation, Standard Operational Procedures (SoP), channels of communication and interactions between actors of different departments or domains (Keen, 1993).
- Individual viewpoint – focus is on individual's characteristics, abilities, strategies, and beliefs to gain an understanding to assist in prediction of behaviour.
- Political viewpoint – focus is on constraints and interactions, moreover 'bickering' within organisations (Peagnat et. al., 2013). (This viewpoint is often not taken as relevant for the design of a DSS, but has also been accountable for new technology implementation failure). This viewpoint has a strong association with strategic decision making, as a result strategic decisions are often motivated by political acumen and indirectly manoeuvre decisions made at the tactical and operational level – moreover it is frequently an unspoken truth. Feedback from actors/users in the system, sometimes given in confidence, is vital to gain knowledge on suitability of technology to achieve objectives and for ease the use on behalf of actors to achieve goals.

Obviously, without detailed knowledge of organisational practices and politics along with personnel behaviour and characteristics, integration of a number of viewpoints into the design of the DSS in the proposed methodology will occur at a highest level of abstraction, based on a number of assumptions. The DSS is designed to be applicable for a number of organisations and domains; however, with information deficit detailed integration of viewpoints is beyond the scope of the research thesis.

3.2 Decision Modelling of FTS

"The intellectual equipment needed for the job of the future is an ability to define problems, quickly assimilate relevant data, conceptualize and reorganize the information, make deductive and inductive leaps with it, ask hard questions about it, discuss findings with colleagues, work collaboratively to find solutions and then convince others."

– (Robert B. Reich (1946 -), American Politician and Writer

Decision making is required to be available at the time when the decision is made so information has to be processed and any simulation needs to be executed quickly and efficiently. The DSS is designed to integrate complex decisions into a simple form with an accurate and precise synopsis to inform the decision maker of design alternatives for the strategic planning of training. The foundation for the design and development of the DSS begins with the analysis of the requirements (Husack & Papadice, 2014). A model driven approach is identified as a solution to the parametric driven approach generally used when engineering analysis is employed (Liu et. al., 2015). The approach forces precision in the

requirements analysis phase, hence, requirements that are expressed in natural language are generally modelled to assist in preventing misinterpretation issues through the design process. The DSS should also permit a structured approach to separate the problem space into domains of interest that can show different levels of understanding of the problem and associated solution space including knowledge management and information handling (Shaw et al. (2002), & Petrie et. al. (2001)). Thus, the DSS is developed and evaluated from personal, technical and abstract organisational perspective or viewpoints.

Attributes to consider within the DSS framework are essential prerequisites to the development of the system model representing the DSS. Acquiring the key attributes for decision support is directly dependent on the level of abstraction used for the decision process. Understanding of the 'cause & effect' and dependency between attributes can assist in the development of mathematical functions within the model to describe the relationships between key parameters. Hernandez et. al. (2011), advocate qualities of an effective model are simplicity, robustness, ease of control, completeness, ease of communication and adaptability. The level of complexity and abstraction used for these models can be seen to influence the characteristics of the domain being modelled. By allowing multiple perspectives and viewpoints to be available within the model; numerous model artefacts can give the perception of reducing the apparent complexity of the model (Taentzer & Bordeleau, 2015). Accordingly, models form the core element in the development of a DSS.

One of the most widely understood modelling techniques is data modelling (Benyon, 1997);(Hernandez et. al., 2012). These types of models consist of structured models (assembled by objects, relationships and rules) and have strong correlations to logical models (petri net, predicate calculus, etc.) that are conceptual representations of the data structures. The models illustrate how data flows through a system and how data can be manipulated by the components of the system. The type of modelling methods include: data flow diagrams, state machines, and ER diagrams. The characteristics of the DSS for the proposed methodology are an evolution of the work presented by Sojda (2006), Turban et. al. (2005) and Ridha (2013) and include:

- Information selective to avoid information overload.
- Ability to combine the use of models and mathematical analysis techniques.
- Include features that are easily understood by non-engineers in an interactive approach.
- Present a common ontology that is familiar with all decision makers using the DSS.

- Aimed at unstructured and ‘wicked’ problems to gain understanding of organisational and operational structure.
- Emphasis flexibility and adaptability to accommodate changes in the environment and decision making approaches of the decision maker.

Usability is an important consideration for both the DSS framework and the workflow process for decision making. Usability definition can be found in ISO 9241 (ISO, 2010), which states that usability is the extent to which a product or system can be used by identified users to achieve specific goals with effectiveness, efficiency and satisfaction in a distinctive context of use. Maguire (2003), identified some benefits of usability including: increased productivity, providing a competitive advantage, a reduction of cost and time, and reducing probability of error. Usability can be considered as an important assessment of a DSS and is strongly related to the ease of following a workflow process for using the framework as a decision support tool. Usability in this context will be based around a metamodel identifying relationships within the training system and providing an easy guide for the identification of mathematical relationships within the framework.

A decision support framework for a FTS has to consider multi-attribute trade-offs involving the capability of both FoS technology and the ability of student pilots who are using them, the availability and deployment of the blending mix within the training system, and how they have been/are allocated. To gain additional knowledge, which the DSS has to consider within the workflow process, lessons can be elicited from research performed within the capability and deployment domain. According to Siemieniuch & Sinclair (2000b), a capability development and Deployment system, directly related to a decision support system, would facilitate the following characteristics:

- Be supported by a knowledge management infrastructure.
- Be supported by a number of COTS tools and simulation packages.
- Enable improvement, deployment, and evaluation of capability with a global aim of cohesion and re-usability.

For the proposed methodology, additional properties are required to be included within the framework system, these include:

- Identification of process steps and workflow patterns within an easy to follow metamodel to reduce human error during the process.
- Identification of methods, tools and techniques to carry out these steps.
- Identification of the appropriate level of technological detail and capability of the systems (for training goals) within the FoS for non-engineering decision makers to base their analysis and decisions on.

- Robust and repeatable techniques for trade study analysis concerning the correct deployment of training technology to the appropriate student at the appropriate stage within the training pipeline: from the analysis performed with the data and visualised within the ROSETTA framework(s) and supporting databases.

The decision support system considers the FoS with common characteristics. These characteristics include functional and non-functional requirements, which are used to inform the decision maker of certain constraints. The functional requirements (What the system needs to do) form the basis of scenarios which describe the interaction of an actor(s) to the system (DSS Framework); these form the basis for the use cases within the model. The non-functional requirements (constraints which the system has to operate within) are described by the attributes of the system and are used for analysis and to further inform the decision maker of suitable alternatives to the optimal. Included within the system model is the physical and organisational environment and the associated relationships between them. These non-functional requirements will be regarded as FoS attributes rather than being assigned to individual systems within the framework.

Included within the FoS (technical or human), the resources are evaluated for the inherent skills and knowledge (K&S) needed to achieve the respected level of training attributes for the mission scenario at the stage within the pipeline. The relationships between the technical and training attributes are given by mathematical functions to assess the suitability and availability of all alternative training systems/technology; these relationships can be classed as assigning priority weightings to recognise the strength of the relationships between the attributes of the systems.

3.2.1 DSS Evaluation Methodology

The increased complexity of technology used in today's training systems adds to the integration issues within an organisation that relies on structures and work processes to successfully integrate required knowledge for the correct functioning of collaborating entities. The accuracy and capability of the integration of the socio-technical systems becomes critical to achieve contractual commitments within industry. The 'soft metric' method used to 'fine-tune' the framework(s) within this research project may be defined by the following statement:

The application of measurement-based techniques to socio-technical variables within the DSS influence the performance of training management processes to provide useful management information regarding causal process factors that

influence the quality and accuracy of the training programme. Such information may then be used for flight training management process improvement to enhance the quality of the end product.

The dynamic nature of socio-technical systems inherently necessitates more sophisticated performance monitoring methods and systems to understand the factors which influence decisions in organisational processes and outputs (Courtney, 2001). Various features and characteristics of the flight training system are responsible for complex emergent behaviour that invariably affects process outcomes including, accuracy, autonomy, performance and cost. Costs associated with the deployment of the DSS ROSETTA framework would relate to the evolution of this proof of concept through pilot studies to full deployment, all of which it is perceived to be time consuming at the initial stages. Categories to consider for full conception into an operational environment include: hardware and software costs, staffing costs, training and additional support. Once the process workflow, mission scenarios, and sensitivity relationships are fully developed within the operational domain, maintenance of the information and data used for the DSS and framework, time and staffing costs are not perceived to be excessive.

Identifying the dependency between attributes of consideration within the ROSETTA framework is of prime concern and incurs a method of sensitivity analysis represented by a mathematical relationship. The sensitivity relationships are used in the main decision body of the framework with visual information to assist the decision maker to identify important beliefs. Visual displays within the framework present information in a certain way that entails a decision based on what is being presented to the decision maker. The information shapes findings in certain ways and directs, to some degree, of how to read them. The visual graphs or charts give material form and scientific visibility by simplification to data sets that were previously immaterial and invisible due to the complexity of representation (i.e. large excel spreadsheet or textual data). (See Chapters 5 & 7 for detailed discussions of ROSETTA frameworks for decision support).

The key to a good decision is to first understand the problem then the other stages that follow are, solution generation, solution analysis and choice between solution alternatives that meet the requirements and finally solution implementation, similar to the concept described by Unsworth and West (2000). Where operational performance is the main concern, three performance attributes are considered: quality (achieved functionality), scheduled

performance, and economic performance. A trade-off analysis can be performed against each of these attributes, thus, the DSS should ideally permit a compromise between attributes, which is generally dependent on agreement with the 'customers' requirement constraints, (for more information about trade-offs in operational management, see Burke (2003)). Slack et. al. (2001) states quality means 'Doing this right' and providing error-free services that are 'fit for purpose'. The decision of what is 'fit for purpose' involves the trade-off activity within the DSS and it is assumed that all alternatives meet the 'fit for purpose' criteria although with varying levels of operational performance. Therefore, the goal of the DSS is not to automate management decisions, but rather support the intuition of the decision maker in choosing the correct blending mix solution to match current training mission requirements.

COMPLEXITY OF MODERN & FUTURE FLIGHT TRAINING SYSTEMS

“...the advice, guidance and assistance given to the Army to train effectively, efficiently and economically in order to fulfil its operational role. It ensures the systematic development, application and evaluation of training techniques and methods, providing Army training with both quality control and quality assurance”
- (AGAI, 1994)

Outline of Chapter

The chapter begins discussions on the research problem of the requirements narrative and includes the concept of training pilots using disparate training technology, especially VRTEs. How humans integrate with the interfaces of the training technology along with the connection to training attributes required to gain readiness is introduced (i.e. a brief study of internal and external factors that contribute to the selection of blending mix solutions). A brief description of transfer of learning from the classroom to task activities and what can be classed as legitimate measurements of the effectiveness of training is discussed. The chapter closes with an outline of the complex flight training system integration problem with introducing the concepts of human performance considerations with trade-off options available for blending mix (training technology) choice.

Training with VRTEs

Training is the process of developing an individual's knowledge, skills, and behaviour through practice and instruction. Education equips individual's intellectual property, knowledge and understanding, which leads to reasoned decisions, judgments and conclusions (Aquinis & Kraiger, 2009). There are a number of functions that need to be considered for the efficient management of a training system (Wagenerf, 2014), these include:

- Training Assurance
- Training / Educational Support
- Resourcing of training and education
- Determine priorities in order to direct resources appropriately to individuals
- Stakeholder inter-relationships and responsibilities
- Defining qualitative and quantitative requirements (time to train, performance attributes).

An efficient training system is one which manages the pipeline to minimise the time spent training between stages (McAllister, 2013). It is vital to provide a training regime to suit individual's ability to retain knowledge and gain skills as efficiently as possible (ICAEA, 2009).

Technology advances have provided the military with substantial capability which places more emphasis on training than in past times (RAF, 2014). However, budget restraints and time to train is becoming more constrained meaning that training resources and facilities are being overtaxed (Hartley, 2015). Flight hours are amongst the long list of expendable items being considered to be traded off in training. VRTE training is used when resources are not available or when safety precludes the use of live training technology; with the added benefit of capturing performance data that can aid in creating accurate and timely post mission reports. However, there are stringent requirements for VRTE that ensure training is enhanced through identifiable use of training goals, otherwise VRTEs can be seen as nothing more than glorified, high-tech video game(s).

VRTE's are generally used to familiarise the student pilot with standard operation procedure (SOP), navigation skills, terrain/environmental familiarization, and most importantly decision making skills (based on dynamically changing scenarios which should affect changes of behaviour of the student pilot) (Nisansala et. al., 2015). Learning to use the VRTE and supporting technology can be time consuming, cumbersome and frustrating. In addition, developing the training programme and evolving the training system to include VRTE's can be equally troublesome. The student pilot should become familiar with the input setup for devices such as keyboard, joystick, or any other non-intuitive input device and be comfortable with information being provided, such as Head Mounted Displays (HMD) or screen displays, are aspects that are generally overlooked in acquiring VRTE's. However, VRTEs cannot replace full scale live training, but there is a general consensus that they can be used to supplement live training and allow users to gain cognitive and procedural skills using real life scenarios to enhance live training (McGrath, 2005). Nevertheless, VRTE's have been known to be detrimental to training; the main factor considered to affect training in a negative manner is the lack of realism (e.g. cues, Field of View (FoV), motion, etc.) provided by the technology (Johnson et. al., 1999);(Smith, 2006).

The key to identifying the VRTE requirements is to analyse the most appropriate delivery method (lecture, discussion, self-study, etc.) and then evaluate the most suitable training

tool/technology. The general guidance on the choices of technology includes a list of possible benefits and metrics including: reduced cost, reduced time, increased availability for learning, and improved performance (EASA, 2013). To this extent, the defence systems approach to training quality standard (DSAT QS) based on the provisions of BS EN ISO 9001:2000, should be applied across all defence training. The standard prescribes a management framework for agile training that can be responsive to feedback from evaluation of a training exercise. Once a need has been identified, the standard necessitates that all training be derived from an analysis of the operational requirements.

Live training can be classified as the exercise of the operational platform while simulated training suggest the operational platform is not exercised but is replaced by another technology (i.e. the simulation) (Abma, & Maig, 2011). JSP 822 (2012) defines blended learning as: “*A blended learning solution combines educational and training methods, media and environments to increase learning effectiveness and efficiency to meet specific training and education needs. These solutions can then be considered and prioritised within practical constraints such as cost, time, political and legal*”. Thus, it is perceived that focus will shift from reasons for using VRTE technology to methods of developing VRTE and blended training capabilities at a lower cost (Hurley, 2011). Currently, VRTEs are used for the following goals, but it is essential for budget restraints to expand these goals to enable more training time on simulators than in live aircraft sorties (Kozuba & Bondaruk, 2014):

- Ease the workload on aging and expensive aircraft
- Maintain or increase combat readiness
- Build pilot experience base
- Serve as an aid to aircraft sorties
- Offset range invasion and weapons training

A better understanding of how virtual environments can be used to focus training efforts to improve transfer of training (ToT) is needed, which will invariably invite the necessity of knowledge acquisition of key stakeholders (instructors and other SMEs) on how best to use the technology. This will lead to a more agile training system that can justify the use of the blending mix for each planned training mission scenario and cost efficient management of resources (Gustavsson et. al., 2013).




4.1 Virtual Environments Critical Factors

VRTE systems convey a level of personal presence within the synthetic environment to an extent that participants have felt ‘immersed’ within the virtual environment. Currently there

are insufficient metrics to define the level of immersion conveyed by virtual environments (Kronqvist et. al., 2016). Of particular interest, especially in military FoS decisions is how this sense of immersion relates to sensorimotor and cognitive performance of the student pilot. A fully interactive VE embodies visual, auditory, and haptic/kinaesthetic interactive component environments. However, due to cost and convenience many VE are created using subsets of the three components.

Computer simulations that generate the virtual environment (VE) can present the real world in an abstract manner and the human operator is allowed to interact with components of the VE through their responses being sensed appropriately and then coupled into the VE simulation. To this extent, field of view, display resolution, level of interaction, feedback mechanisms, etc. all contribute to the level of immersion experienced in addition to the level of task involvement which is pivotal to the sense of immersion (Stevens & Kinkaid, 2015). Haptic feedback provides tactile cues about the object that has been touched, collided or for pilot training - the sensation of the gravitational forces acted upon the pilot. Haptic / kinaesthetic feedback is difficult to achieve in real time scenarios and causes issues with human proprioception that relies on integrating forces that our whole body experiences to maintain a mental model of the surrounding VE and timing issues between all three environments. Visual Cues have been found to improve immersion and help the user gain a sense of presence within the simulation (Meyer et. al., 2012). The main types of VRTEs used for flight simulation training are described in Table 1 along with some advantages and disadvantage of acquisition.

Table 1 Subset of VRTE (Flight Simulation) Tools

VRTE	Advantage	Disadvantage	Visual Representation
Reconfigurable Skills Trainer	<ul style="list-style-type: none"> Negates the requirement for expensive multi console training equipment. Allows team, sub-team and individual training. Hardware can support other training media. 	<ul style="list-style-type: none"> Cost of hardware, software and development. Some loss of fidelity may be required to allow reconfiguration. Long development time. 	
Part Task Trainers	<ul style="list-style-type: none"> High level of fidelity. Permits multiple fault injections whilst not contravening Health & Safety regulations. Provides realism for specific tasks 	<ul style="list-style-type: none"> Cost. Size. Mobility, Accessibility. May support training in small groups only. Expensive and difficult to modify. At risk of subsequent updates to Government Furnished Equipment (GFE). 	
Full Mission Simulator	<ul style="list-style-type: none"> Allows students to experience and gain ideas of actual situation. Environment and teaching situation controlled. May be the only possible training medium due to danger of real environment. May be networked for federated and confederated training systems. Ability to replicate most fidelity requirements. Can provide the opportunity to improve unit collective performance wherever people need to practice expensive or dangerous activities under realistic conditions. Actively involve the learners in making decisions, playing roles and adopting attitudes. Simulators allow instructors to progress from simple to complex scenarios. 	<ul style="list-style-type: none"> Can be expensive. Requires instructor to brief, monitor, and debrief for learning transfer to occur. Assessments tend to be subjective, relying on the experience of the instructor. Must be well designed to ensure maximum benefits both intellectually and emotionally. Requires careful project management to ensure that only required learning outcomes are provided. Requirements likely to be over specified. New software must be developed each time a change is made to operational equipment. 	

Fidelity is the level of realism presented to the learner, and is focussed on the technology used to simulate a particular learning environment. Hays and Singer (1989) defines fidelity as *‘the degree of similarity between the training situation and the operational situation which is simulated. It is a two dimensional measurement of this similarity in terms of : (1) the physical characteristics, for example visual, spatial, kinaesthetic, etc.; and (2) the functional characteristics, for example the informational, stimulus, and response options of the training situation (p. 50)’*. The unique measure for a VRTE can simply be described as how well the simulation formulates its representation, or its fidelity, which in turn can be described as the accuracy of the representation when compared to the real world (Stotz, 2000). To be more specific the formal definition of fidelity is *‘the degree to which a model or simulation reproduces the state and behaviour of a real world object or the perception of the real world object, feature, condition, or chosen standard in a measureable or perceived manner. The methods, metrics, semantics of models or simulations used to compare these models or simulations to real world referents in such terms as accuracy, scope, resolution, level of detail, level of abstraction and repeatability’* (Mowbray et. al., 2003). Resolution, error/accuracy, precision, sensitivity and capacity define the requirements of the simulation or model capabilities.

The fidelity of the model or simulation is expressed in terms of the relevant referent and the model or simulation capabilities. The training system/technology often attempts to emulate as many physical and functional stimuli as experienced in the real world. The current direction is to increase the physical fidelity (look and feel – displays, switches, HOTAS) and functional fidelity (dynamics or actions – flight model, weapons models), (Kozuba & Bondaruk, 2014). Fidelity for Human-in-the-loop (HITL) applications, such as VRTEs, can be categorised by specific sensory inputs such as: visual cues, motion cues, agent intelligence, noise effects, cockpit displays, tactile feedback, temperature effects, aircraft models, weapons models, sensor models, and environmental models. Physical fidelity refers to the degree of similarity of layout and feel of the VRTE compared to the real entity; visual fidelity refers to the degree data is presented to the user as a surrogate for the real world view to in effect minimise judgement errors during execution of training mission; environment fidelity refers to the correlation with the natural and tactile environment for the VRTE such as visual range being affected by rain. Physical, visual and verbal fidelity are fairly realistic, however, weather simulation often has a distinct lack of fidelity (Perey, 2008). Psychological fidelity is

the degree which the simulation manufactures the sensory and cognitive processes within the student pilot as experienced in the real world (Ghani, 2016).

For some time it is thought that high fidelity simulations (both visual and auditory) the feeling of presence can be very powerful (Vafadar, 2013), however, the lack of a complete objective measure this feeling cannot be verified. Nevertheless, high fidelity simulations does not necessarily transfer to more effective training and some lower fidelity simulations can assist in acquiring the details of training and education without the added complexity and physical configurations of high fidelity VRTEs (Wickens & Hollands, 2012). It may be that the lack of fidelity enhances the focus of the student pilot on general principles of communication, coordination and problem solving. Park et. al. (2005) has found that high fidelity training simulation can actually hinder effective training and learning especially because of over stimulus of novice students. This argument is strengthened by the work conducted by Hays and Singer (1989) whose research concluded that the VRTE does not need to exactly replicate the real world to provide effective training. They suggest that departing from realism may improve the training efficiency.

‘...there is some evidence from flight simulation that higher levels of fidelity have little or no effect on skill transfer and reductions in fidelity actually improve training. Reductions of complexity may aid working memory and attention as skills and knowledge are initially acquired’
– (Caird, p.128, 1996)

Hence, is the perceived objectives on improving the ‘look’ of the VRTE using visual, sound, and physical fidelity for improved training effectiveness pertinent, or should research be concentrating on using training goals and objectives to drive improvements in VRTE technology and therefore concentrate on the task which the VRTEs are expected to perform along with the level(s) of fidelity required to complete the training tasks? Irrespective, fidelity has proven difficult to clarify and apply in practice (especially for training), the qualitative terms such as high, medium, and low is generally used to describe a model or simulation (DoD, 2010). Since fidelity is regarded at the ‘goodness’ for simulations, an objective measurement offers benefit for describing and choosing VRTE for specific objectives and goals (Korteling et. al., 2013); the relationship between fidelity and training benefit has not been identified. If this definition could be determined, a fitness assessment of the simulation for a specific purpose can be implemented (sometimes referred to as appropriateness or suitability for a desired application). However, the benefits and cost effectiveness of training systems have not been fully investigated (Cohn et. al., 2009).

Measurement is focussed on the experience and opinions of users with regard to training effectiveness of the specific training technology used, thus, lacks the measurement of the real transfer of learning and retention of training for real operational conditions. The level to which training leads to readiness and enhancement of actual behaviour on the job should be the standard of measuring effectiveness of training (Alvarez et. al., 2004). (Fidelity is further discussed in relation to the methodology in Chapter 5.2 and throughout Chapter 7.)

For a decision on which VRTE system and blending solution to use during flight training, a firm understanding of the performance, complexity and capability of each system is a good starting point. Many researchers have discussed aspects of performance: Liang et. al. (2008), expressed the necessity of rapid response for real time audio images; Friston & Steed (2014), suggests latency is more important than visual resolution – a number of other researchers also carried the same view (Meehan et. al., 2003);(Phillips et. al. 2005);(Lippi et. al., 2010). Typically, the performance of a simulation will increase as the complexity of a model decreases; and it could be said that reducing scope/resolution often entails a loss of validity of the model. Thus, a trade-off exists between performance and validity. The purpose of the model, however, should give an indication of the acceptable threshold limits of these two parameters (Deniaud, 2015). Occasionally, the inadequacies cause perceptual conflicts that contribute to simulator sickness (Young, 2006). Thus, the understanding of the human perceptual system can permit trade-offs in various performance factors of the VRTE system.

The granting of VRTE approval has been primarily driven by the level of fidelity giving correlation between the features of the VRTE and associated equipment simulated to the production of ToT that is both positive and high (Hamblin, 2005). In human processing terms, controlling the aircraft is a divided-attention task, requiring the coordination and integration of big data, and procedures. Thus, pilots need to establish under what conditions to make control alterations with reference to the outside world or the virtual environment in order to assess their control accuracy. The objective with using VRTEs, especially low fidelity simulators, in the initial stages of instruction should be to maximise the quality of instruction time in live real-world flying (IFALPA, 2011). The major deficiency with PC-based simulators is the restriction of field of view (FoV) generally confined to a forward view; the virtual/physical aircraft instruments are quite small and sometimes of non-standard format, the aerodynamic aircraft model can be elementary, and the flight control can lack ‘real feel’ (Strachan, 2014). This notwithstanding, high physical fidelity may necessarily not be

required for training utility. In general, student pilots should perform better if feedback and a review of a previous executed flight and a preview of the manoeuvres in the forthcoming flight using desktop simulators are given (Rantz et. al., 2009). It is also reasonable to hypothesize that student pilots would not develop sufficient psychomotor skills for combat flight from training on desktop flight simulators, but they would be adequate for development of cognitive skills required to control an aircraft's flight path (McLean et. al., 2012).

State of the art displays and control technologies within the cockpits had the premise of introducing computer based technology and automation and with it more capable systems. However, Multifunctional displays (MFD) led to an increase in pilot workload where the pilot is in danger of information saturation (Barbe et. al., 2012). The touch screen technology offered by the MFDs bring some advantages including, pilot interaction – intuitive operations, cockpit definition – space optimization and software flexibility. However, if they are not properly located, this could increase workload and degrade human-machine interaction (Young et. al., 2012), (see Volume II, Table III for criteria). The operational intent requires to be decomposed into specific functional requirements before dealing with allocation of pilot cognitive resource (Bestoso, 2005). There are strict specifications about the visual aspects of cockpit design within the military standards. Internal and external visual fields are of prime concern, the internal visual field signifies the cockpit scope that the pilot can see with natural vision line and the external visual field signifies the cockpit scope that the pilot can see without interference of airplane and cockpit structure (Zhang et. al., 2007);(Yeh et. al., 2013).

In most applications there exists a trade-off in fidelity for affordable or performance goals; a typical example is the resolution of representation for aspects of real world which are less relevant to achieve the simulation goals (McGrah, 2005);(Stewart et. al., 2008). Thus, the definition of the required level of fidelity has to be strongly associated with simulation scenario goals. The scenario goals are generally a measurement of MoP or effectiveness measurement of MoE. However, the focus of the simulation has to concentrate on the performance criteria defined by the MoP which are surrogates for the real world measures of mission success. In military simulations, the measures of performance (MoP) are used to judge how well the system(s) are operating in relation to the operational task and goals.

Fidelity is generally judged using subjective validation by SME, which can lead to incorrect conclusions depending on the interpretation of the goal aspects of the simulation. To simplify this discussion, the less error, the more accuracy and thus is deemed to have more fidelity

(Dunan, 2006), however, what is generally misunderstood is the accuracy of the simulation is directly related to the simulation objectives and goals. Recognition of the object is more important than an precise representation of the object i.e. if it looks like a fern tree on the display it is perceived to be a fern tree which is positioned in this geological coordinate in the real world (this perceived reality is the basis for standardization of data and descriptions in the simulation and is referent for simulation results). The ability to articulate fidelity requirements in addition to the estimation and measurement of fidelity relationships is becoming and will become increasingly important as VRTEs become more prevalent in training applications. There is currently a lack of agreement on the definition of fidelity, however, there is a consensus on the aspects that are important to consider (Hess & Marches, 2009)), these are:

- A measure of realism of a simulation
- The degree of similarity, both physical and functional between simulator and referent.
- The perceived degree of representation within the simulator to a real world object, feature or condition in a measureable manner, this has to be in relation to simulator training goals

Realism has been the defining goal of developing VRTEs to the extent of reproducing, with high fidelity, any real world artefact such as terrain, equipment failures, weather, motion, etc. (Bamodu, 2013) without giving consideration as to what the objectives of such VRTEs are. Technology designers have given little consideration of the use of such technology for acquisition of K&S. Thus, there is a distinct development gap to the continued development of VRTEs for use in training (Gerlach & Durak, 2015). Most flight training programmes (including military programmes) have not evolved to keep in touch with technology, as such there is a gap in knowledge of how best to develop training programmes utilising the technology and even assisting in developing the technology to suit training better (ICF, 2013). This notwithstanding, it is often overlooked that a VRTE does not train; it is a glorified gaming system. The manner in which the VRTE is used provides training quality and needs; the transfer of training (ToT) objectives and instructional aims to the VRTE are the important attribute values for training in such virtual environments.

The decision of how much fidelity representing simulation characteristics is highly dependent of the chosen VRTE specifications and the speed required for rendering; Jean (2008), states the military have no standards to measure the performance or the benefits of simulation-based training. This problem is accentuated by the ‘lack’ of knowledge to fully understand the

relationship between VRTE fidelity and organisational goals leading to misinterpretations of critical aspects of the application domain. The technical usefulness of acquiring a VRTE and deciding on blending mix is bequeathed to the consultations between decision makers who have little or no understanding of VRTE fidelity related to technical usefulness for the intended application and VRTE suppliers who want to sell high fidelity simulators for maximum profit (Reweti, 2014). Hence, an organisation that has multiple VRTEs might discover complications with the simulation appropriateness for the organisation's intended purpose or suitability for combined use with other training devices/tools. It has become important to evaluate and compare the performance between VRTE sessions and VRTE configurations; the purpose of this evaluation is to provide a quality of training that translates into enhanced performance of real world tasks (Goldberg, 2005). The current MoP (subjective ratings) may be considered 'derivatives' of true information that discloses diminutive about the dynamic development the student pilots experiences during training programmes.

Some VRTEs can provide the facility to assess the behaviour of the pilot by providing adequate feedback to be used for analysis of organisational goals. The associated feedback from the VRTE indicates the degree of discrepancy between actual performances and the expected outcome of the mission goal. However, it has long since been thought that the training value is assessed through the degree of technical fidelity, latency times and motion systems (Borgvall et. al., 2008). Furthermore, there is an argument to be made about the choice of training technology based on cost and learning objectives rather than on the basis of technical or fidelity criteria, (Nyssen et. al., 2002).

4.2 Evaluating FTS Family of Systems

The main question that needs to be asked is 'How to evaluate the effectiveness of simulation for air combat training using a blended mix programme of VRTE tools with modern advanced live aircraft with synthetic capability?' To answer such a question key focus on what the goals and objectives are for the evaluation criteria, is needed. Addition questions include: Was the VRTE training effective?; How frequently is it needed?; Is the chosen blending mix used cost effective? The answers to these questions pose a further problem, one training tool might train a pilot quicker than others, but another provides a higher level of proficiency and performance. To compound the problem further, it is clear that certain 'users of technology' prefer (are more comfortable with) disparate layouts and visual cues than

others, also people learn better in different ways, which in a classroom is difficult to achieve (Bell & Kozlowski, 2007). This difference in comfort and learning styles could affect the performance in different ways (Lee, 2005).

Training of humans is a behavioural and cognitive event (Salas & Cannon-Bowers, 2001), and must be a systematic approach consisting on a number of tasks and learning objectives mapped to the design of the learning environment (for: practice (Lee, 2005);(Kennedy et. al., 2007), feedback (Hughes, 2004);(Sottolare, 2004), and performance measurements (Henderson et. al., 2000);(Cannon-Bowers, 2012)). A key goal for blended training is in gaining competencies for mission readiness (Chapter 4.2.1.1) thus it stands to reason that concentration in active learning with a VRTE should be the prime focus of the training system as active learning enhances the development of metacognitive (executive-level processes entailing knowledge, awareness, and control of activity involved in goal attainment (Ward et. al., 2012) and self-regulatory skills. Flight training revolves around the following techniques but not all of them are used in concurrent training programme designs (NAVMC, 2011):-

- Individualism of training – rate of learning of each student is considered and this is fundamental to organising instructional training as they progress through the pipeline.
- Functional Context – course is set around mission scenario modules pertaining to aircraft control, manoeuvres and training objectives.
- Instruction sequencing – to ensure prerequisite K&S before training is acquired.
- Objective measures – training goals are stated that are measureable to give indication of student pilot progress or attainment.
- Minimise costs – substitution of more cost effective technology to meet training tasks.

Most evaluation techniques come from parameters of the training technology rather than the student pilot's performance. For training to be effective, a correlation between the two measurement techniques is needed (Jacko, 2012). More specifically the VRTE should assist in evaluating reaction, learning, behaviour, and goal outcomes as can be seen in Figure 11.

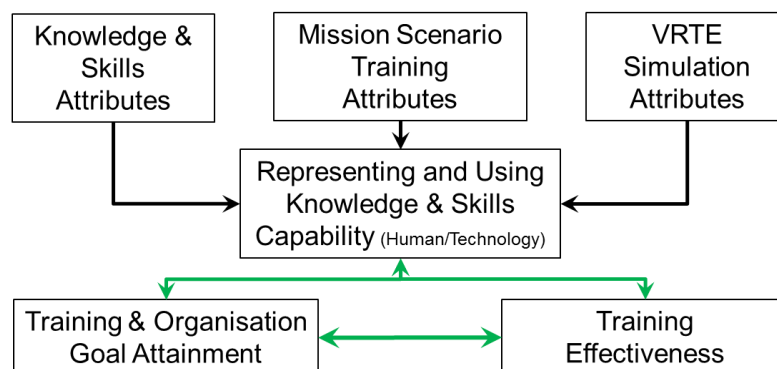


Figure 11 Holistic View of Learning With VRTEs

Reaction can be appraised about the performance comparability using the chosen blending mix with a baseline scenario; learning can be directed to the achievement of minimum objectives (ICF, 2013). Behaviour can be categorised as the correct decision making ability to react to a learned process during training on the chosen blending mix and the results are directed to the impact of the training on organisational objectives (generally, the main focus is on cost effectiveness and the transfer of learning (K&S) onto the student pilot – generally referred to as MoP). Knowledge in this context can be described as instructional objectives to learn about something or to learn how to do something (declarative or procedural) and a skill how to action something using the training technology. Low fidelity and commercially available flight VRTEs provide performance measurement capability, which can be mapped onto cognitive measurement techniques for respective students.

The decisions on the training effectiveness with the blending mix should be judged on the performance of the tasks within the mission scenario and evaluated in term of the overall task performance: safety, efficiency, and effectiveness (satisfying multiple goals) (McAllister, 2013). Unfortunately, relating decision quality to overall task performance is extremely difficult to achieve: to manage information, cognitive work, communication, and actions that must be accomplished within a fixed time or event window. The understanding of which tasks take priority over others and prioritizing these must be clearly understood (IATA, 2013);(Thomas & Lee, 2015).

There is clear indication that a flight simulator that successfully imitates the aircraft goes some way to effectively train a pilot, which intern reflects a pilot's proficiency in operating the aircraft (Neville, 2011). Evaluation of performance conducted in this way constitutes a determination on the readiness of the pilot to perform his duties in a live environment. Assessing student's progress, the criticality of the task, its difficulty and the frequency with which the task is encountered are needed for each mission scenario (ICF, 2013). The output of which is used to assist decision makers to train to the level of 'awareness'. The assessment of how well the student pilots achieve the objectives using the chosen blending mix is used to validate the training

The main metrics for progress through the training pipeline include attributes such as: hours flown, accuracy of manoeuvres (from post mission briefing) and K&S levels, which are considered to be dynamic variables that give an indication of changes in performance (Everson, 2013). For example, student pilots who are performing well in all chosen blending

mixes should show improvements in the attributes in the right direction i.e. decrease in accuracy deviation. Hence, effort needs to be more focussed on monitoring and measuring performance of student pilots using the blending mixes on real world tactical problems that require the use of K&S and the capability of the training technology to provide the correct training levels to successfully complete the mission objectives in order for an objective measure of performance to be calculated. Benchmarking will set a standard for the performance measurement attributes and performance will be compared to this; students whose performance is improving, therefore, will gradually approach the benchmark standard. Evaluation of training systems is multi-faceted and includes reactions to training, attitude, K&S of student pilots and capability of the blending mix and the impact of performance measures to the next planned training mission (ICF, 2013).

In order to establish quantifiable outputs from the FTS, gaining a sense of what inputs to consider for analysis is needed. One of the obvious inputs to consider is the quality of the student pilots who are going to be processed through the training pipeline, however, other inputs to consider include: teaching activities, the instructors, types of aircraft and VRTE used in the training programme and the goals and objectives which accompany them. Thus, any resource within the training programme is required to meet a quality standard (for this thesis concentration is given to a subset of these aforementioned attributes). The outputs for consideration include the performance outcome of the mission scenarios or more importantly the training scheme as a whole (the training programme is a scheme of instruction); furthermore, attention will be given to the blending mix environments and configurations as they are integral to the training activity. The outputs should be able to assist in estimating the most effective and efficient quantifiable choice of which training system/technology and configuration environment (aircraft or VRTE) to use for which training mission scenario and associated goals for each student pilot at a specific point in the training pipeline. The outputs are calculated by considering the attributes of the inputs into the FTS and identifying relationships between them in order to obtain metrics to calculate which training environment to use, as per Figure 11, to successfully achieve the main goals and objectives of the current training scenario (see Chapter 6 & 7 for allocation of attributes for MoP and MoE).

4.2.1 Pilot Education & Performance Considerations for Blending Mix Choice

'In fighter flying, a panic message is the greatest of all crimes. Practice on the ground the exact words you will use to cover any situation in the air. Say it over and over again until it becomes automatic' - Group Captain Reade Tilley, RAF.

Overall effective performance of a FTS includes two categories containing both human factors and technical systems that must be synergistically integrated operationally (Mavin et al., 2013);(Yang et. al., 2014), as in Figure 12.

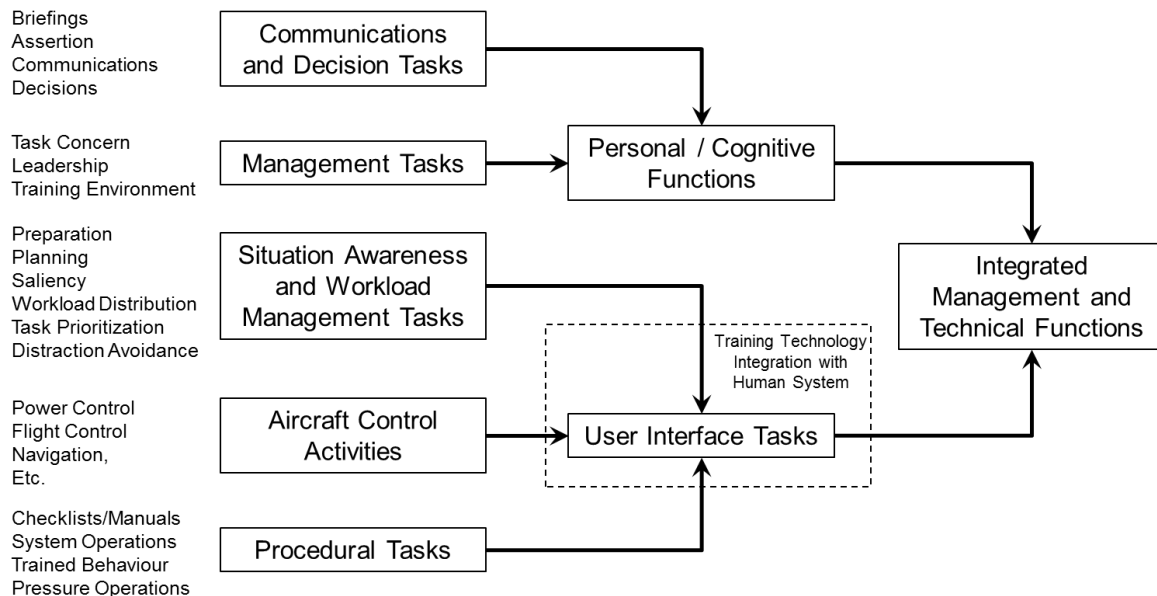


Figure 12 Integrated Performance Model for FTS

The process of information transfer from instructor to student pilot is crucial to effectively improve pilot performance. Higher levels of communication (briefings) need to address mission objectives and any technical issues in performing the mission (See Volume II, Table II); open communication between student pilot and instructor can lead to a greater understanding of mission goals and blending mix constraints: communication can determine the success or otherwise in achieving goals. In flight, performing a tactical mission scenario where stakes are high, communication effectiveness is essential and is one of the primary means to enable student pilots to develop and coordinate activities in order to achieve mission objectives (Wihl, 2015). One of the key factors is the ability of the decision maker to have an indication of the quality of interpersonal relationships with the student pilots being trained and how they will relate training in the chosen blending mix. This inevitably leads to the encouragement and exchanges of information relevant to the tasks at hand.

The management of student pilot's workload and Situation Awareness (SA) can dramatically affect performance (Gundert et. al., 2012). Pre-flight preparation and planning is fundamental in relieving some of the pressures of workload and SA (Bell & Kozlowski, 2007), especially with disparate technology capabilities and configurations. The controls and procedural tasks finish the representation of the user interface. Clear understanding of the two categories

should provide insights into management of the training system and assist in identifying any weaknesses within it.

‘The picture we get of good Captains who are playful, anticipate difficulties, use time during the normal phase to prepare for higher workload periods, and who are ahead of the curve....I suggest that the good Captains, by articulating plans and strategies...helps build a shared mental model for the situation. It enables to make suggestions, coordinate actions, and offer information that contributes to solving the problem and making the decision’ – (Orasanu, pp.13-15, 1990)

Markers associated with situation awareness and workload management categories are as follows (Flin et. al., 2003);(CAA, 2013);(Fernandes & Braarud, 2015):

- Being aware of stress factors that can reduce attention
- Actively monitors environmental system and instruments for relevant information
- ‘Look ahead’ – Prepares for expected or unforeseen (possible) situations
- Workload condition is clearly understood, communicated and acknowledged
- Ensures the additional secondary tasks are prioritized
- Recognises when limits in own human capabilities is reached and reports
- Plans before manoeuvres to avoid high workload conditions
- Ensures relevant parties are aware a status of flight and self.
- Recognised potential distractions and takes appropriate preventative actions.

One of the difficulties in training pilots is some students possess excellent ‘stick and rudder’ control, but management of information input especially in modern cockpits is sometimes of greater importance and difficult to train than pressure tasks. The ability for a pilot to direct attention to relevant information in a cockpit full of touchscreens and illuminating indicators is an important factor, especially in disparate blending mix layouts, as familiarity in high workload situations will reduce risk.

“When considering the acquisition of some high specific knowledge and skill, certain laws of skill acquisition always apply. The first of these is the ‘power of practice’ – acquiring skill takes time, often requiring iteration practice in retrieving a piece of information or executing a procedure.”

– (Higgins and Howell, 2004)

In training individuals (especially adults), personality and character differences has to be taken into account, some of the more prevalent differences between individuals include:

- Aptitude (ability to learn a certain task), (West, 2011)
- Personality and learning style (learning styles), (Xu, 2011);(Yanardener et. al., 2014)
- Prior experience (context for learning), (Perrota, 2011)
- Coping with stress (correlation between stress and performance), (Delahangi, 2011)
- Attitude and motivation (attainment), (Cummings et. al., 2012)

- Cognitive processing skills (memory span, processing speed), (Grossman, 2011)

To correctly operate a complex advanced cockpit, the development of SA moreover hand-eye coordination and the mental capacity to maintain spatial orientation in a time critical operational task has become one of the biggest challenges in modern flight training as cognitive skills are harder to train and these psychomotor skills are desirable but difficult to train with adult pilots (ATSB, 2007);(Li & Harris, 2013). During the learning process aspects of skill become more automated (via muscle memory) and the demand on working memory decreases whilst at the same time the importance of perceptual speed increases. Thus, the success of the training system is not just down to the organisation and decision making processes but also the available blending mix, and the character and personality of the student pilot(s). Understanding the relationships between these attributes is significant for a successful flight training system, moreover than the type of VRTE to use.

A training programme should concentrate on competency (demonstrated ability) and readiness (Vermeulen et. al., 2014), which can be described as the ability to use skills taught in training, in the ‘heat of battle’ (CASA, 2009). Hence, the high level goal of the training programme is to teach a level of competency with the ability to attain readiness (CAAP, 2009). The training system is also required to be flexible and adaptable to allow standardised training for every student pilot, but also adapted tailored training that is student specific (Smit, 2012). Humans inherently have the characteristic of learning at different rates and optimisation of these rates may require different training tools/blending mixes (Stevens et. al., 2015). The process for choosing the blending mix should allow for identification of which training technologies best suits each student pilot; additionally, the decision maker must have the ability and guidance to personalize the training for each student to achieve a common standard of training.

Scenario based training emphasises the development of critical thinking, flight management, and flying skills, a skilled pilot might experience during operations (Ayers, 2006). The main goal of training in the VRTE is to accelerate the acquisition of higher level decision-making skills and airmanship, by using the mission tasks developed by SMEs to strengthen the required levels of training attributes. The use of the VRTEs is perceived to allow for recurrent training to achieve a level of readiness, which can be validated in a mission scenario involving real flight (Thatcher, 2007).

There are three states or outcomes to transfer of training (ToT): positive (improvement of real world performance), nil (no effect), or negative (degraded real world performance) (Alexander et. al., 2005). Training concentrates on the positive aspects of the ToT and it is deemed that the percentage of ToT in the first hour is higher than in the second, and so on: consequently the effectiveness diminishes over time (Krausert, 2015). Negative ToT occurs when an individual applies incorrect methods and techniques learned in one environment in another environment and this is often a real hazard in configuration settings for individuals; and often used as an argument that VRTEs must have high fidelity characteristics to assure that no bad habits are learnt that are subsequently transferred to a real-world situation (Bell, 2007). Thus, the practical implementation of VRTEs used in flight training is that pilots train themselves using the tools provided in direct control, support and guidance of instructors (Nisansala et. al., 2015). This notwithstanding according to Persing and Bellish (2005), the cost of development of simulation models increases exponentially as fidelity increases which puts a cost burden to training budgets, which is contrary to the government objectives for military training.

Identification of appropriate blending mixes for efficient ToT is both dependent on the stages of learning of the student pilot and the selection and feedback of prior training mission scenario executions using various systems (technology) of the FoS (Rantz et al., 2009). The data gathered from VRTE's along with subjective assessments in associating blending mixes to mission scenarios and student pilots continue to evolve during the lifecycle of the FTS to gain knowledge of the most cost and time effective blending solution for progression and identification of appropriate levels of FoS characteristics needed for current training (K&S) levels. The link between ToT and fidelity is a crucial but a highly contextual issue with the levels of fidelity required for training tasks to be determined for achieving desired levels of training without over investment. The argument of high fidelity training that can closely emulate real world conditions can effectively train and aid in the transfer of learning into the real world might be misleading without considering the stage of the learner (Astwood et. al., 2010). It has also become recognised that the more familiar a pilot is with a simulation or training system/technology design for the transfer of learning, the greater amount of fidelity attributes they needed to sustain adequate transfer of learning rates (Bilotta, 2013). Thus, it is important to distinguish the roles of fidelity for training and assessment, with the goals of the training tool and the stages of the learners. As a result, the concept of learning and assessment must be viewed as complementary to training.

4.2.1.1 Pilot Readiness

“Experience is that marvellous thing that enables you to recognise a mistake when you make it again.”
- F. P. Jones

Readiness is one of four components for preparedness, namely force structure (resource numbers), modernization of forces, and sustainability (Dunn, 2013). Readiness can be obtained from the ability to perform a manoeuvre correctly after a number of repetitions, without error, over time (Chapman & Colegrove, 2013),(Rostker, 2014). Thus, a small force could not be considered prepared and likewise a more sophisticated and better equipped force could not be considered prepared if inadequate training was prevalent. Therefore, a fully manned unit with modern equipment in perfect working order would be classified as not ready, if it trained for only a brief period of time. With resource restrictions it has become clear to military decision makers that fixing readiness than modernizing and enlarging their forces is vital for being prepared. Readiness is maintained by the successful repetition of appropriate mission scenarios periodically as a function of pilot experience (Yang et. al., 1996).

One of the key issues with readiness is a term called the ‘startle’ factor (Martin et. al., 2012). This describes the ability of the pilot to manoeuvre an aircraft in the training environment correctly with the occurrence of an unexpected event, however, once in real flight the same unexpected occurrence causes a startled feeling in the pilot: this is perceived as the difference in feeling safe on the ground (i.e. no danger whilst given a high degree of danger in the air). As a result, training in the VRTE must include the ability to suppress the startle response, confirm the situation, and then apply measured and proportional corrective inputs during a realistic training mission scenario (Martin et. al., 2013);(EASA, 2015).

Currently, a level of proficiency across a range of tasks is used to measure readiness levels; once a participant reaches a minimum level of proficiency in a range of tasks, they are declared as ready (Levy, 2006);(GAO, 2015). Therefore, it seems natural to identify the relationship between blending mixes and proficiency for the different levels of fidelity. The analysis must be robust enough to indicate whether it would be more cost effective to increase retention and retain experienced pilots or fly more mission repetitions in a VRTE. The analysis must make clear:

- Whether the mission scenario must be flown or be executed in a VRTE, or can be either. (this will be based on tasks and level of flight goal accuracy required)

- The importance of the tasks within the mission scenario (for choices in times of reduced availability of resources)
- Periodicity – how frequently a certain task must be performed.

The analysis should consider the nature of training in increasingly complex exercises intended to improve proficiency and hence readiness (Marken et. al., 2007). For the student pilot to qualify in any mission scenario, they must meet the minimum standards prescribed in the associated MoPs and MoEs. (MoPs identify knowledge, skills, and abilities related to the process, MoEs identify quantifiable mission outcomes). Performance can be defined as completing a task to some set standard. This standard includes rating performance in time, distance, quantity, accuracy, or even an objective evaluation (Franks et. al., 2014). The assessment of training must balance the need for realism against the expected threat; pilot and blending mix capabilities; and safety. (See Chapter 6.5 for the readiness metamodel used in the methodology).

4.3 Pilot Attributes in Relation to the Systems of the FoS

‘The winner (of an air battle) may have been determined by the amount of time, energy, thought and training an individual has previously accomplished in an effort to increase his ability as a fighter pilot’

– Commander Randy Cunningham, USN

Situational Awareness Factors

Situation awareness (SA) is crucial for good decision making in dynamic systems, such as aviation and is critical when considering capability and suitability of blending mixes for acquiring levels of K&S for the student pilots. SA is defined by (Endsley, 1995) as *‘the perception of the elements in the environment within a volume of time and space, the comprehension of their meaning and the projection of their status in the near future’*. SA can also be defined *‘as the continuous extraction of environmental information, the integration of this information with previous knowledge to form a coherent mental picture, and the use of that picture in directing further perception and anticipating future events’* (Jeannot, et.al., 2002). SA is defined by Hjelmfelt & Pokrant (1998) as *‘Goal oriented and not-task oriented’*. Tasks are performed through a mission to accomplish goals, but the goals remain relatively constant for the duration of the mission. Based on SA, one makes decisions to do certain tasks to accomplish the high level goals’, (Nofi, 2000). Despite considerable research of SA, currently there still is no universally accepted definition or model of the concept (Endsley, 2003);(Salmon et. al., 2006);(Panteli & Kirschen, 2015).

SA considers dynamic events and situations and has the effect of allowing the HITL to give meaning to the situation. SA is viewed as a localised parameter to the individual but is influenced and maintained by the training technology and simulations; but is not the ability or knowledge of how to act. This knowledge is gained from the mission goals that are translated, to a degree of understanding, by the pilot of what processes and information is deemed relevant for the mission using the chosen blending mix. In evaluating SA for a mission scenario, current and past events need to be considered when concerning decision rules. These decision rules are based on behavioural models and trained actions. These decision rules are relevant to the current value of SA, as future actions depend on the current decision that is actioned (Mantovani & Casterl-Nouvel, 2003). It is important that relevant information presented by the training technology during mission execution is maintained in working memory (Burgon, 2014), important rapidly changing information can be sampled when needed and then once actioned upon, can be forgotten, and rules for training should be compliant with ignoring irrelevant information (Strybel et. al., 2008), which could cause issues in disparate layouts of technology.

‘Know and use all the capabilities in your plane. If you don’t, sooner or later some guy, who does use them all, will kick your ass’

- Lt. Dave Pace, USN, US Navy Fighter-Weapons School Instructor

Measures of SA provide an index of how well pilots are able to acquire and integrate information in blending mixes where there is competition from a number of indicators, displays and the environment for their attention (Bares et. al., 2010). The FoS is concerned with the manner of displaying important feedback information to enable acquisition of data (based on the mission goals) under operational conditions (Skaffin, 2010). Of interest is what ‘costs’ the pilot faces in terms of improved situation understanding when used in context of all other displays in the systems of the FoS (blending mixes). Designing systems that assist in improving SA for pilots is vital for improving training efficiency and increasing performance outcomes (Bullemer, 2013).

Analysis of SA should include pilots information gathering, interpretation of current situation, including tasks relating to in-flight decision making with regards to mission planning, navigation and tactical flight (Kozuba & Bondaruk, 2014). Conceptual modelling of the scenario can give an abstract delineation of situations that are of importance, for example current position of aircraft is known with a possible future situation if state of system is unchanged. The action of the pilot can reveal how the current situation is perceived, what

information the pilot has attended to and what the pilot finds important for action to occur. An undesirable outcome is when the pilot acts with incomplete SA when the pilot believes it to be complete e.g. confusion on symbology on a display causes incorrect pilot decision making and associated actions on erroneous awareness factors.

Some of the difficulties of SA to be considered for each system of the FoS is summarised by Endsley, Bolte, & Jones (2003) as:

- Attention Narrowing – System needs to support multitasking activities across multiple goals and decisions. Focussing on one activity for prolonged period of time can lead to high risk situations.
- Workload – Increased workload can decrease information acquisition, increase fatigue and other stressors.
- Data Overload – Too much information, rate of change of data can lead to an overloaded situation in the pilot.
- Out-of-place Salience – Too much visual features, bright lights and colours can overwhelm and distract pilot's attention.
- Wayward Mental Models – The events causing certain system behaviour needs to be known by pilots. Comprehension and projection relies on the pilot understanding how a system operates being triggered by events. This understanding forms the pilot's mental model of the system.

SA fashions critical input, but is separate from pilot decision making that is the source of all subsequent pilot actions. For SA, the pilot perceives with his senses some object within their visual field / environment along with its properties (colour, size, location, etc.). Based upon knowledge and experience gained along with other recognised objects, the pilot gains a holistic representation of the environment, grasping the significance of objects and events and generates a world model; which leads to a projection of possible future events. This projection allows the pilot adequate knowledge (and time) essential to decide on favourable courses of action to meet mission objectives; and uses the training technology to exercise the actions. The abstract view of the world for decision making is illustrated in Figure 13 (Meystel, 2001).

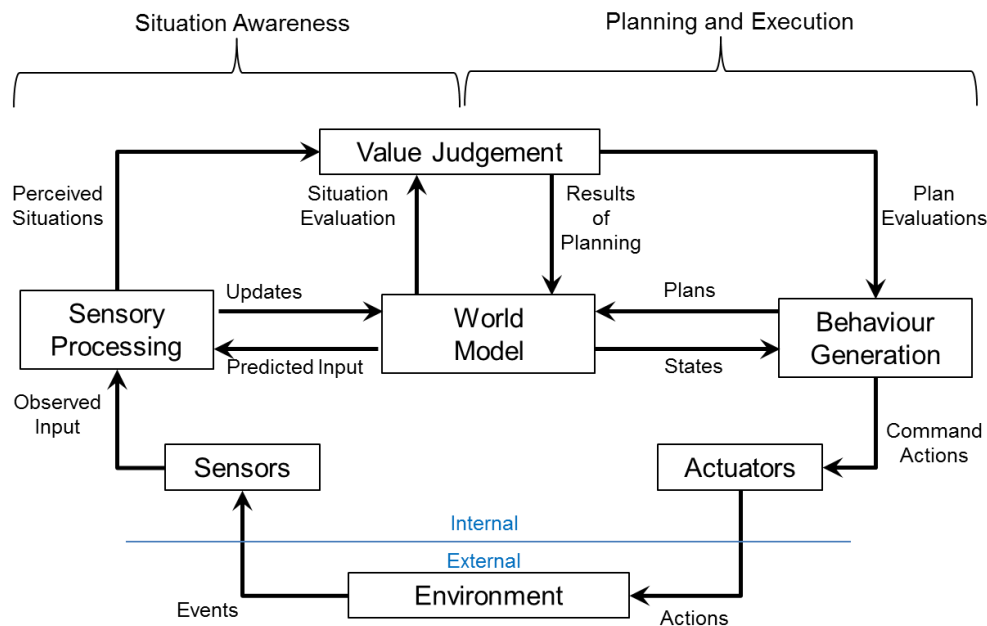


Figure 13 Functional Relationship between SA and Decision Making Factors (Meystel, 2001)

From Figure 13, the model can be extrapolated in relation to planning, estimating mission success and student pilot performance as planning the mission requires *a priori* knowledge of the general behaviour required by the pilot using the chosen blending mix to action the tasks for mission success; this requires a world view of the flight path and the associated environment that might be encountered with the estimation of the number of activities / actions required by the pilot to control the aircraft course. Knowledge of the blending mix capabilities with respect to visual, physical and functional fidelity is needed to assess the constraints of feedback information available for the pilot to make efficient and accurate decisions on based the new received world model. The update of SA, from sensory inputs, is needed to make a value judgement on the next state of behaviour that pilot is required to action. In execution of the mission scenario, this process of updating SA based on feedback information continues as the pilot progresses through the mission tasks, where the student will compare current flight state with the mission requirements. SA, therefore, is a critical component for decision making (Endsley, 1991).

With technology (VRTEs and aircraft) being developed in a technology-centric view has created a condition where pilots need to search for information they really need, leading to poor SA, high workload, and multiple opportunities for errors to occur (Chialastri, 2011). In training, these factors need to be managed. Visual perception stimulates the pilot's information acquisition process from the cockpit display and out-of-the-window (OOTW) observations (within a simulation environment the OOTW is a function of the virtual system).

The pilots activity comprises of both periodical sampling of information associated with the demand and the particular task being conducted at a specific time. Certain demands require the pilot to intentionally focus on a specific task and thus a specific type of information to support confirmation of the task is required to be processed by the training technology (Corker & Guneratne, 2003). The pilot behaves within certain constraints and thus certain assumptions have to be made by the pilot, as an example:

- The pilot should apply a different scan pattern according to the availability of OOTW information provided by the blending mix.
- Overlapping areas of interest should signify a conflict and thus an inattention to the instruments. Decision making ability of the pilot would permit attention to the appropriate information outlet at the correct time within the flight.
- In order to account for disparate functionality and display layouts of the cockpit, the pilot should gain a generic fixation on control settings to gain familiarisation to commonly used information outlets within the cockpit or simulation.

The multitasking capabilities of pilots involve the primary role of vision. Each visual activity can have a priority associated with it; the procedure for the pilot to scan the OTTW and the user interface (UI) can be separate procedures within the full task of activities (Shy et. al., 2002). The product of these procedures should determine the likelihood of a successful mission and may lead to the identification of emergent behaviour within the pilots visual scan pattern using the constraints of the chosen blending mix. These scenario events can lead to the identification of possible human error based on current knowledge about the systems capability and performance strengths and weaknesses. In seeking to minimise performance deviations to expectations some key human performance capabilities have to be considered to develop a better understanding of the sources of performance creep (CAA, 2013). Capabilities include (Lee et. al.,2005);(IFALPA, 2012):

- Perception/discrimination – Issues can include but not limited to misinterpretation of object recognition and communication misunderstanding using the technology.
- Situational Awareness.
- Intention Formation - Goals are models in terms of events and actions (tasks), Priorities drive the process and in a multi-goal/task environment priorities may dynamically reorder and lead to incorrect intended actions when processing the change.
- Attention – Attention weights can change dynamically over sub-elements of a complex and demanding task. Thus, attention strategies may be adopted and change in the course of performing a complex task using disparate layouts / configurations of technology.

- Memory
- Expectancies – Successful performance depends on the anticipation of the probable future events. The pilot's ability to perform routinized tasks under high workload conditions involve multiple trained tasks each competing for processing resource.
- Execution – Actions executing motor movements. Task execution has a time limited attribute, executing a task outside the constraints represents a performance deviation.
- Workload – Competition for tasks to be executed in a short time frame calls for strategic workload management for pilots. Error occurs when too many tasks require attention and cause pilots to skip less important tasks and cause performance creep.
- Effects of stress.

As the functionality of training technology is increasing pilot decision making and performance is pushed further towards the limits of human capability, which increases the potential for confusion, additional workload and loss of situational awareness for the pilot (Casner et. al., 2014). Research is continuing to investigate into factors that affect performance attributes of the pilot (Mayer et. al., 2012). These are categorised into headers task variables (pilot interfaces with aircraft controls), environment variable (temperature, noise, lighting etc.), procedural variables (trained SoP for the pilot in scenarios), and pilot-centred variables (pilot personal characteristics) as per Jirgl et. al. (2014), each of which can influence and affect pilot's SA. It has been found that most losses of SA have been attributed to needless complexity, coupling, autonomy and inadequate feedback of training technology (Woods, 1996);(Gray, 2012). (See Chapter 6 & 7 for integration of SA measures in the methodology).

Workload Factors

Workload has been defined by Garner and Murphy (1979) as a set of task demands, and as an activity or accomplishment. The goals of the task, time allowed, and the performance levels are the task demands, Wang et. al. (2013); availability of technology, information required and stress are factors in consideration. Performance measures give a good indication of workload, as it is perceived that as workload increases, performance scores decrease; furthermore, motivation and learning are factors that could influence the performance metric. Laudeman and Palmer (1995), considers the workload measure to be based on a logical connection between workload and task constraints, and not on theory.

Pilot workload can be described as the cohesion of mental and physical effort required to fulfil the perceived demands of a specified flight task (Cummings et. al., 2009). Mental workload effects are directly associated with the work task in relation to the characteristics

and experience of the individual performing the task. If task demands are too high or too low increases the likelihood of degradation in human performance. With the invention of the Heads Up Display (HUD) and VRTE systems, technology matured and with it the operational requirements for training pilots needed to be upgraded to keep up with new technology insertion and aircraft system functionality. This emerging technology presents a new challenge with pilot workload. With the number of displays, switches and buttons as well as different cockpit orientations there is a danger of the pilot becoming saturated to the point that flying the aircraft was a secondary concern (Petitt, 2012). However, with the VRTE systems it is now possible to seek familiarization with layout, displays, symbology, etc. of the cockpit before implementation.

Workload can also be defined as the human resource expended when performing a specified task (Sang et. al., 2011) or characterized by the interaction between interfaces of a system and task and the operator's resource capability, motivation and current state of mind (Mehler et. al., 2012). Workload assessments can be evaluated focussing of technology characteristics of the blending mix and integrated into a SA evaluations. This presents important information to pilots and decision makers of mission scenarios to enable more intuitive approaches of presenting training goals and objectives using appropriate training technology / blending mix with concentration given to situational awareness and workload factors (Olson, 2007). (See Chapter 6 & 7 for integration of Workload measures in the methodology).

4.4 FTS Decision Making Strategies

The primary decision making strategy is predominantly developed by consultation with SMEs where the probability of selection of the correct decision on a time line basis for each task is decided (Beaubien et. al., 2015). The goal for the pilot is to develop situation awareness by perceiving future events and moreover hazards, which are present in events that could lead to undesirable future events (Mayer et. al., 2012). The pilot should systematically identify and list any hazards associated with any flight plan using the chosen blending mix, including external influences caused by the environment that could affect the objectives on the scenario and the level of situation awareness of the pilot; the risk elements are described in Figure 14.

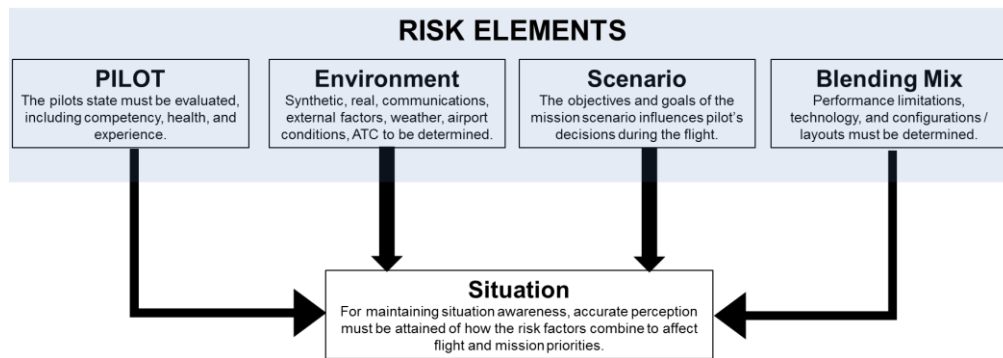


Figure 14 Risk Elements Affecting Situation Awareness

Task-specific information required to construct the mission scenario attributes was obtained by studying various task analysis and through consultation with SME's from the RAF. The main tasks under observation are:

1. Sustaining speed of the aircraft to mission plan
2. Maintaining heading of the aircraft to mission plan
3. Being aware of aircraft location at all times
4. Control of altitude of aircraft to mission plan
5. Management of time constraints during mission
6. Following appropriate SoP, especially during poor visibility conditions

These tasks have a direct impact on navigation, they do take time to execute and time, especially for quick decision making under pressure, is of limited resource in making navigational decisions. The mission scenarios should be designed to reduce the time available to the pilot to make navigational decisions with the time constraints becoming stricter as pilots continue through to further stages of the pipeline, (Damas et. al., 2011).

In more advanced complex automatic systems the humans role has changed from being less involved in active control of the system and more pre-occupied with activities such as: perception, monitoring, evaluating, communication, and problem solving (Donmez et. al., 2010). Two primary phases of flight has the tasks consisting of monitoring the state of the aircraft and maintaining up-to-date representation of that state i.e. updating the belief state to match the real state. Skilled pilots actively check for a number of events and conditions that do not occur in the scenarios, such as: late changes in wind direction that are direct dependencies of the mission scenario task (Hohmann & Orlick, 2014). In order to obtain a realistic workload level inclusion of these sorts of checks in the model are required.

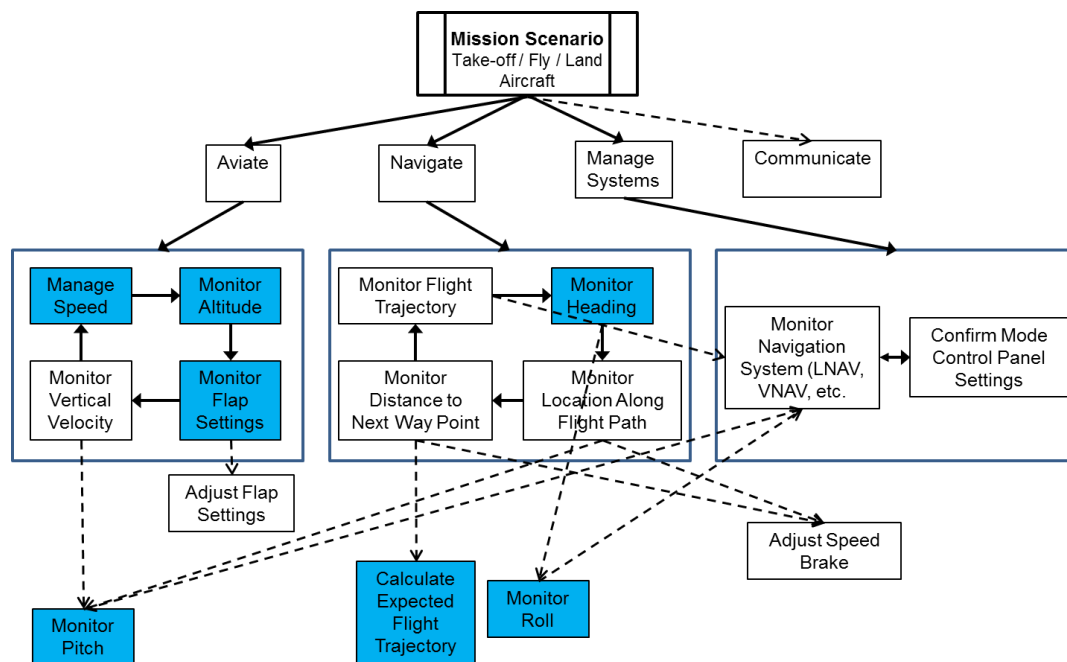


Figure 15 Goal Oriented Flow of Control Resulting from Task Analysis (adapted from Keller, 2003)

A typical architecture to formalise the flow of control from task analysis is shown in Figure 15 where each block describes an activity (Act_i) the pilot has to action in the tasks, which is a modified version of the task analysis structure described in Keller et. al. (2003). The dashed lines represent optional sub-goals which may not occur on every scan pattern of the pilot, the shaded boxes indicate some information may be required to perform the action. (See mission planning sections of Chapter 7 for task analysis in the methodology).

4.4.1 Human Performance Factors Affecting Choice Of Blending Mix

'Human behaviour, either cognitive or psychometric, is too diverse to model unless it is sufficiently constrained by the situation or environment; however, when these environmental constraints exist, to model behaviour adequately, one must include a model for that environment.'

- (Baron, p.6, 1984)

Measuring performance is especially relevant when the requirement is to develop methods and metrics to assess the effectiveness of blending mixes for training purposes and for predicting mission readiness of pilots, i.e. transfer of skills to the real world (Harrison, 2014);(Evans, 2015). Individual behaviour, performance levels and strategies explain how or why a particular outcome to a mission scenario occurs. Sample output measures include accuracy of performance to a baseline, timeliness of actions, and number of errors. The delineation of behaviours can assist in developing performance measures: from various literature reviews these behaviours can be placed into 5 categories (Carretta, 1992);(Rodgers, 1993);(Staal, 2004):

- 1) Demonstration awareness of surrounding environment
- 2) Recognising problems
- 3) Anticipating a need for action
- 4) Demonstrating knowledge of tasks
- 5) Demonstrating awareness of important information.

The specification of training objectives needs to be identified in the mission planning phase (Jenkins et. al., 2008); these include critical tasks and conditions and the specific learning objectives and with it the specific MECs embraced by the mission scenario using the specified system of the FoS (chosen blending mix). The learning objectives of the mission include behaviours which are deficit in the student pilot(s), those which are subject to skill decay, or those which are difficult to perform that need frequent practice. Thus, mission planning has to consider the importance of obtaining an estimated performance measure for respective student pilot *a priori* to undertaking the mission (Gehr, 2004).

‘Performance means initiative – the most valuable moral and practical asset in any form of war’

- Major Sholto Douglas, RAF

The examination of expected to actual behaviour can be used as measurement data to support feedback and future training needs for particular student pilots (McAllister, 2013). Mission training advocates a subjective assessment of performance during all phases of training, these are a sequence of plan, brief, mission execution and after action reviews (commonly known as post-mission brief) (Smith, 2000). Objective measures can be used as a basis of comparison with subjective measures to reflect whether attitudes reveal what actually happened during execution of the mission scenario (Foster et. al., 2009). The generation of a pilot model to predict behaviour based on subjective data is deemed to be inaccurate for the complex HITL system due to ‘Startle’ factor that could affect the pilot at any point in the flight. According to Parasuraman (2002), human factor inputs into the system functionality requirements of the blending mix, should reduce errors within the real-world operational field and thus assist in created error-tolerant designs.

‘A mismatch between the functionality as specified by the designer, the operating environment (i.e. procedures) and the user’s requirements for the system or his/her mental model of system functionality...[resulting in] inefficient system performance, errors, and possible adverse performance including accidents.’

- Parasuraman et. al. (2002, p.7)

A Human-centred design and evaluation process method is shown in Figure 16, which is amended from Hooey (2002). This method can be used to identify system and display

designs and associated procedures, which may lead to errors due to incompatible design to task and behaviours required to achieve system goals with HITL operations.

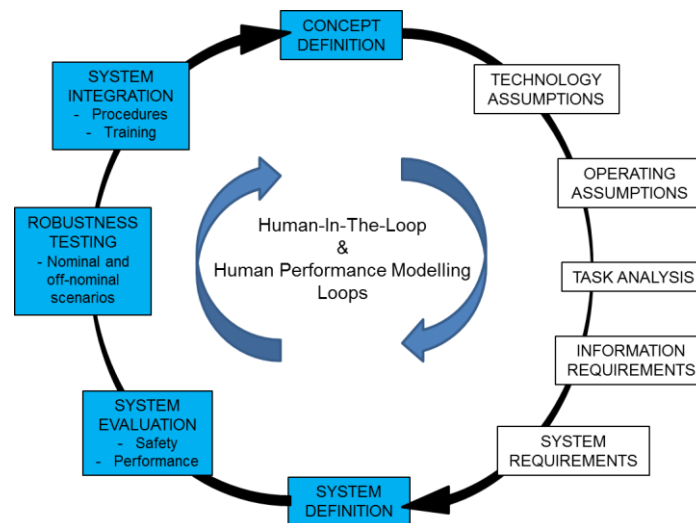


Figure 16 Human-Centred Design and Evaluation Process (amended from Hooey, 2002)

The process integrates task analysis, technology and operational assumptions into system requirements, which are then instantiated as the defined system. The iteration loop within this process allows for performance evaluation, definition and integration of procedures. The HITL and HPM can be applied to investigate the human decision process and behaviours; the process is evolved into the process workflow design for the FTS for scenario precision estimation with the ability of recursion and multiple technology (blending mix) assessments. (See Chapter 6 for design of Workflow process).

Levels of minimal performance must be identified and the performance level of the student must be tracked (Franks et. al., 2014). Once the acceptable performance is known, it should be used to make decisions regarding the student progression, whether to repeat the mission scenario, change blending mix choice configuration or remediation / intervention, is required. This criterion is important as it allows an assessment of training attributes that is more helpful than a ‘check-box’ system (Hardison et. al., 2015). Thus, the success of the training programme is not just down to the organisation and SME decision making processes but also the character and personality of the student pilot(s) and the capability and suitability of blending mix choice(s). Understanding the relationship between these attributes is crucial, moreover, the type of training technology/tool to use for a successful training system.

4.4.2 FTS Effectiveness Measures

In systems, the FTS can be described in terms of an integrated network of people, technology and other resources that accomplish some objectives and/or goals (Brannen, 2014);(Tamarskar et. al., 2014). The introduction of new technology with the increase in automation will require the pilot to gain an additional understanding of how the system functions. This should increase the technology confidence of the pilot and give an understanding of what the limits are.

The progression of student pilots through the flight training system should concentrate on the acquisition of K&S for mission readiness. A skill is described as a task performed to a specific level of competence or proficiency that generally requires the manipulation of controls and instruments; competency is derived as a skill performed to a specific standard under controlled conditions (Bullet, 2010);(Mavin & Roth, 2014). To encompass these factors and in an attempt to advance this type of competency-based training (Beaubien, 2015), the training methods include elements such as:

- Pre- and post-mission briefing sessions to inform the pilots what is expected of them and allow an opportunity for questions to be asked. The post-mission brief is used to feedback to pilots their performance with the chosen blending mix, with advice on how to improve in future training missions.
- Classroom based instruction includes both media-based and discussion forums to impart knowledge and techniques from instructor to student. This classroom time has the unique property to be tailored to individual pilot's needs.
- Training technology are used to become comfortable with instrument and control layouts with approaches, and to become familiar with the concept of switchology (master of checklists and flows) that can be done without the added distraction of the motion of a live aircraft.
- Live aircraft is essential for a pilot to gain appreciation of how a real aircraft handles. Training with live aircraft gives the pilot a chance to explore the boundaries of controlled flight and how to recover the aircraft from troubled states (stalls, spins, etc.).

A formal method of predicting operability and performance issues during the training programme has been the goal in modern flight training for a number of years (Levy, 2006);(GAO, 2015). Flight training operability must be defined in terms that are relevant to flight operations but also to training and technology developers (Petford & Frank, 2009);(SDI, 2015). Operability goals and objective measures should be established to determine compliance or satisfaction with the requirements. The operability assessment must assist in

identifying operational system drivers (training attributes), critical requirements (organisational objectives) that are significant influence on operational costs, performance factors, training schedules and risk. The training attributes have a direct correlation to main objectives and goals of training (Fletcher & Ward, 2014). Two methods of assessment have been developed to assess student pilots: the first is to use normative results that compare student's performance to peers and rank where the student is in relation to others; the second method uses a comparison to a benchmark where students are graded to how accurately they match the benchmark solution (ICAO, 2013). Most training systems have a robust method in place for controlling and tracking input attributes to keep track of progress and K&S levels attained for each pilot (Smaili et. al., 2013). The management of a structured feedback system should consider the process as a standard activity and have the ability to detect any deviations from the standard, thus, be able to put in place any remedial actions to correct the deviation (Keller et. al., 2003b) (i.e. deviation: student's slow progress in attaining a specific skill, remedial action – additional practice and skill specific concentration). Schnieder (1985), discusses some fallacies relating to skills training; there is a risk of boredom which accompanies repeated concentration because the task is a familiar one, and perhaps the primary goal of skills training is not for accurate performance but for consistent improvement, not detriment, in performance for each student pilot.

To model the performance to planned goals of the tasks in the mission scenario it needs to be possible to evaluate positive and negative behavioural outcomes and its expected outcomes. This includes the elements of the perceived ease and difficulty with which an individual is able to perform the trained behaviour (Sparko et. al., 2010) on the specific training technology / blending mix used for execution of the mission scenario. To strengthen this concept, practicing what they have learned in real flight is essential for monitoring pilot performance as they progress through the training pipeline. Some theories of behaviour and behaviour change emphasise the impact of technology on behaviour; technology can perpetuate unsustainable behaviours, some of which may replace instinct and decision making with innovation and change the behaviour of pilots (Shove, 2010).

The proposed methodology in the thesis suggests a multiple set of quantitative metrics and qualitative data to be included for qualifying scenario performance within the workflow process (baseline performance level, self-assessment questionnaires); the evaluation on blending mix configurations are based upon mission scenarios, which are designed to enhance

pilot performance and situation awareness, the pilot's decision making effectiveness, and the overall performance-based outcome of the scenario. Therefore, is it important to decompose the mission tasks and evaluate the effect of OOTW view and decide whether this view is a critical component of the K&S to be trained and whether any possible negative Transfer of Training (ToT) will occur.

4.4.3 Transfer of Learning

Transfer is defined as 'the change in the performance of a task as a result of the prior performance of a different task' (Cormier & Hagman, 1987). The transfer effectiveness ratio concentrates on the time or trial to reach a performance criterion after the transfer to a live environment and thus gives a long term indication of on the job performance rather than just proficiency in the training school (Grossman, 2011). The minimum acceptable performance of the pilot should be based on three simple categories (Burke, 2007);(Grossman, 2011):

- (1) The observed behaviour during flight (virtual and live)
- (2) The range of acceptable performance of mission goal(s) – the range of acceptable performance should be based to reflect real-world operational requirements consistent with safety.
- (3) The applicable conditions – the type of training technology the mission is being performed on as well as the surrounding environment which may affect the behaviour of the pilot (e.g. weather, aircraft configuration, personal / emotional conditions of the student, etc.). A further condition which is sometimes overlook is the condition which the proficiency checks are taken, i.e. has the range of accuracy measurements taken account the direction and force of wind, pre-mission state of pilot or other technology/system or environmental conditions.

Learning and performance are highly coupled and task dependent (Tubsree & Tubsree, 2012);(Gibbons, 2014), i.e. the ability for K&S transfer depends on the similarity of the practice and transfer tasks (Chiaburu, 2005). Factors to consider for these transfers to occur are (Cheng, 2008):

- Cognitive outcomes – declarative knowledge, knowledge structure, and cognitive strategies.
- Skill-based outcomes – skill accumulation and automaticity
- Affective outcomes – motivation and attitudes.

These factors indicate different propositions for performance as well as approaches to measurement (Velandia et. al., 2007). Within the blended mix of training technology, performance must be decomposed to its constituent elements to provide a more accurate

feedback on student's performance (Alipour et. al., 2011). There are actual performance measures from various data sources and those derived from decision maker (SME) judgments. Kilpatrick (1976) schema classifies training effectiveness evaluations by abstract functions or questions, however, the schema is easily adopted for this proposed methodology by the suggested usage of data collected by training surveys (questionnaires and interviews), training analysis (forecasting training effectiveness for chosen mission scenario in comparative study), and performance-based research (effects of training in task execution), which are set by a number of training objectives. Summers (2012) and Stevens et. al., (2015), provides further insight in research designs for assessing the effectiveness of military training devices. The effectiveness of training can transpire by a number of activities for evaluation (IATA, 2012):

- Identification of specific measurements – what to measure.
- Assessment of performance measures – collection and use of measures.
- Use valid and reliable performance measures.
- Work within the constraints of the evaluation – measurements in relation to the technology available or use a surrogate model to estimate the measures based on empirical data.
- Use baseline scenarios to control the research situation – obtain a greater degree of control over any experimental situation.
- Use of analytical models at early stages, permit evolution of models as more data is available – cost of producing model will reduce after each successive execution during its lifecycle. The model is required to be updated and revised to reflect performance-based results to determine test conditions and estimate circumstances that are currently unknown.
- Include judgmental measures to supplement performance or analytical data – provision of the 'human' judgment is needed to account for individuality in the evaluation of training effectiveness.

The methodology proposes methods for the extent to which data from performance measures can be attributed to the use of training technology/systems. Performance measures can be grouped into two broad categories:

1. Subjective measures based on human judgement
2. Objective measures based on actual feedback from technology/observations after execution of training mission

The proposed objective measures are based on how effective the training directly enhances the performance of individuals; an advantage of quantifiable indexes of training effectiveness can be mathematically manipulated to predict effectiveness and used to develop trade studies between technology suitability, capability, performance and cost attribute values. User

reaction to the training system / blending mix is an important measure for training effectiveness. Gaining a view of the impressions of the system, especially collating negative reactions to training technology can give an indication of how the individual or groups will interact with the technology for training purposes (Nisansala et. al., 2015). Reactions on training objectives using the technology can provide vital insight of the face validity (useful for the training scenario) of the training technology. Mavin & Roth (2015) provides some examples of a formal assessment of student pilot's reactions.

4.5 FTS Current Trade-off Options

Training in the UK progresses from initial flying training basic to advanced fighter training, successful candidates will then undertake operational conversion unit. RAF flying hours for jet pilots is between 180 and 240 hours per year (18.5 per month on average). Of these hours, 150 hours (12.5 hours per month) are believed to be minimum hours to ensure safety. An additional military element is added for tactical manoeuvres adding up to minimum 3 hours per month for a total of est. 186 annual hours on average (180-240 hours). Flight-hour-to-simulator ratios vary by stages in the pipeline and who is managing the training syllabus; for initial Hawk training there is about a 5:1 (live:VRTE) ratio using a legacy syllabus (140 flying hours and 28 VRTE hours). It is envisaged a new syllabus under development will have the ratio est. 1.8:1 with a reduction of flight hours from 140 to 106, and an increase in flight simulator hours from 28 to 60 (MOD, 2013).

Practice on tasks could occur in VRTE and evaluated in a live exercise; verification of proficiency should remain firmly in a flying event with much of the training occurring in the VRTE (NDCAF, 2012). Both live and VRTE training concentrates on the acquisition and improvement of a skill(s) by giving more practice to respective skill(s).

Research in the domain area have drawn some conclusions, which are largely based on those who have aviation experience (Plat et. al., 1991);(Haar, 2005);(Blumel & Haase, 2010);(VRS, 2014):

- Live training is definitively needed for perceptual-motor skills
- VRTE is extremely useful for switchology, introduction, practice, procedures and rehearsals.
- VRTEs are not a direct replacement of flight hours, but are complementary to for reduction in training deficiencies (bought on by task load and increased complexity).

- Tactical functions can be learned in a VRTE and then evaluated in live flight (distributed simulations are generally more cost effective to execute with VRTE).
- VRTEs, when not supervised correctly, can incur inappropriate responses; they can give a false sense of accomplishment and achievement.
- There is a gap between the introduction of the VRTE used for training and the ability to translate its use to readiness and proficiency improvement.

A number of trade-off options between live and virtual training are available (Schank et. al., 2002);(Kirby et. al., 2011), these are:

1. Add VRTE hours to flight hours to ‘theoretically’ gain greater pilot proficiency; costs increase and this type of trade-off is questionable if pilots are already meeting minimum readiness with flight hours.
2. Substitution of flight hours to VRTE hours; Pilot readiness currently with this trade-off is based on a number of assumptions and is deemed to be a cost saving in real terms. Issues like availability of suitable fidelity VRTE to match task difficulty and evolution of syllabus for pilot qualification has to be overcome. (Qualification is time limited (has periodicity) and will vary by task complexity).
3. Modification of training programme (event, periodicity) to achieve readiness for available flight and VRTE hours with current minimum readiness levels. This trade-off concentrates on the training programme when budgets are tight and number of available resource decline.
4. Modify the minimum readiness standard to account for greater VRTE hours at the expense of flight hours. This is to ensure the number of qualified pilots remain at current levels in time of budget and resource constraints.
5. Modify training system to achieve maximum pilot proficiency with available budget and resources (i.e. equipment, personnel). Additional variables are needed to be analysed within the training system (e.g. simulator fidelity, aircraft availability, human factors, etc.) to form the core of progression and mission scenario assignment to resource.

These trade-off options are based not on quantitative assessments, but on subjective (qualitative) opinions without considering capability of the training technology / blending mix to provide necessary levels of K&S to enable the student pilot to attain a level of proficiency.

4.5.1 Understanding Trade-off Relationships

It is likely that there will be a number of emergent behaviours (e.g. resources will not be saved or readiness not maintained (or increased)) once trade-offs are investigated and implemented and the complexity will have hidden constraints. This complexity results from the relationships between the attributes, and expenditure on higher fidelity VRTEs or synthetic mixes within live aircraft may not improve efficiency of training if the relationships between relevant training variables constrain the improvement. The attribute values under consideration include: individual performance, tasks, technology configurations, personal characteristics, availability, aptitude, cognitive retrieval, capability, K&S, scenario planning and management.

It stands to reason the more training functions a VRTE can do the more complex events can take place, which can affect the objective of the pilot in attaining greater levels of proficiency and therefore readiness (Kennedy et. al. 2014). It becomes clear that a formalised understanding of the relationships (mathematical representations) representing the constraints could lead decision makers in gaining knowledge of the interactions between the attributes and the cost and benefits of trade-offs between available technology resources. However, the organisational objectives need to be clearly understood in formal relationships that can be mathematically manipulated to determine what factors are important in the analysis i.e. is it important to reduce cost?, maximise efficiency?, or a multi-objective goal where trade-offs between conflicting objectives have to be considered.

For the basis of the research it is deemed that classroom training does not play a significant role in any trade-off decision, as the classroom training clearly plays an important role in the training lifecycle (Scott, 2010). Furthermore, classroom training is a precursor or adjunct to both virtual and live training, as a result it is excluded from the trade-off analysis (Malmin & Reibling, 1995a;b);(Katajavuori et. al., 2006), but is an essential feedback measure to evolve training styles and methods.

4.6 Chapter Summary

Current flight simulator training is not viewed as being as valuable as real flight training, and key components of pilot experience using VRTEs is still missing (realism, feedback upon change of state of aircraft), but there is a general consensus that they can provide complementary training to reduce deficiencies. Additional data is needed to analyse the effect on proficiency with repetition and the degree of acceptance of the simulator as a

training tool and not as a glorified gaming machine. It is clear that economic and resource constraints can affect readiness of pilots as such, decisions need to be made about where the trade-off needs to occur. Additional pressures to training managers will be undesirable to the training system if technology and economic resources are not available to train the next generation of fighter pilots.

Significant technological advances in fidelity have improved productivity and realism in modelling, simulation, and distributed training, however, there is still little change in the trading-off / balancing between live and virtual training, and choosing the correct blending mix. One of the serious problems yet to be overcome is the poor fidelity of the simulators compared with the actual aircraft, especially regarding the physical fidelity characteristics. Detailed understanding of human factor issues hold the key to the success of blended training as without taking factors such as, resolution, stability, FoV, distortion, etc. may ultimately result in a VRTE system that is difficult and awkward to use. Currently, metrics to understand the relationship between technology and human factors are in the infancy and little understanding of trade-offs seem to exist. A user centric approach relating to human factors should drive technology development as will a more detailed understanding of the human visual, auditory and haptic/kinaesthetic systems. The training and prior experiences should allow the student pilot to apply K&S to act on unforeseen or planned occurrences during a mission scenario using the chosen blending mix (Dismukes et. al. 2007). Additional information including questionnaires on K&S, self-efficacy, situation awareness, stress, and motivation with using the chosen blending mix may provide additional information related to human factors and performance to enhance training technology evaluation techniques to consider more of the end user of the system, which directly affects readiness in relation to training goals. Mathematical modelling, such as Monte-Carlo processing is not a suitable substitute for realistic pilot-in-the-loop evaluations as user feedback gathers more useful data than a distribution analysis based on possible future events.

Without any performing metric to match the human operator to the training technologies, designing and choosing the correct FoS system and FoS Configuration will be merely subject to opinion for a varied selection of SME's all with different views (i.e. the windows or MAC argument). Thus, if suitable metrics were to exist, then quantifying the benefits of a particular FoS system for an operational task would be beneficial to decision makers. Once a set of metrics have been established, the real benefits of all systems can be realised and subsequent costs of acquiring such technology can be reduced to suit a particular need.

A ROSETTA FRAMEWORK FOR MODELLING AND SIMULATION

“On two occasions I have been asked (by members of parliament), ‘Pray, Mr. Babbage, if you put into the machine the wrong figures, will the right answers come out?’ I am not able rightly to apprehend the kind of confusion of ideas that could provoke such a question.”

- Charles Babbage

Outline of Chapter

This chapter describes the advantages of using the strengths of both qualitative and quantitative data within one framework structure proposed by an implementation of the theory of ROSETTA within the methodology to assist in the development of surrogate models that can be used for modelling and simulation of a complex multi-attribute problems to aid in the decision making task by providing measures relating to design solutions. The concepts of fidelity characteristics of technology are discussed and how they relate to the effects of training in relation to trade-studies between available training technologies (blending mix); with further discussions regarding the objectives of the analysis with respect to providing blending mix solutions within the flight training system and the potential benefits of using the ROSETTA framework for decision support. The chapter concludes with a synopsis of architecting a mission scenario to ensure traceability of performance and goal success of the mission along with the perceptions of human related characteristics and behaviours with respect to interaction with technology that will affect performance outcomes.

Mixed Method Research Integration within the ROSETTA Framework

The compatibility of integrating both quantitative and qualitative data into one model is dependent upon the fundamental values between the two and their relationship to the research method (Venkatesh et. al., 2013). Evaluation of the data using analysis techniques can then be used to check the importance of understanding and assist in making uniformed decisions, with the belief ‘that the world is complex and stratified and often difficult to understand’ (Reichardt & Rallis, 1994, p.88); thus, often requires iteration of the analysis method to check understanding and can lead to acquisition of new knowledge.

The qualitative approach concentrates on induction to theory and data analysis and incurs a subjective relationship to research process, whereas the quantitative approach concentrates on deduction in the theory and data analysis and objectivity in its relationship to the research

project (Eickmeyer & Gruhe, 2010);(Bendansolli, 2013). It is now thought that using strictly quantitative approach in complex problems involving human interaction within the system would not sufficiently capture the complexity of the problem. However, it is hard to argue against that both approaches incurs some bias on values based on the person conducting the research, thus, mixed method approaches attempts to remove the bias by allowing the researcher to see both types of data and make a subjective opinion based on both (Wall et. al., 2015). Both methods, however, must meet appropriate criteria for rigour and both must be complete i.e. each should stand alone.

The mixed method approach used to gather data for analysis in the prototype workflow process is not designed to replace the more traditional research methods, but works to utilize the strengths of both in the same research project to minimise the weaknesses of the individual approaches (Johnson & Onwuegbuzie, 2004). Using the perception that both individual approaches are seen as real reflections of reality, if each approach delivers contradictory results this could be an indication of different perspectives or a biased development of judgements and thus lead to the knowledge that one or both approaches are in need of redesign for a clearer picture of the real world. The integration of both approaches in a single research project is to challenge the underlying assumptions of the two approaches themselves and question or highlight any biases used in each.

5.1 Introduction to Relational Orientation

The preservation of relationships during the transformation process from requirements to the designed system model is the prevalent task for systems engineering design; nonetheless a FoS/SoS is not ‘designed’ but rather assembled and integrated. Tracing requirements through the FoS entails not just the specification of how individual systems must operate but also how the systems interoperates and communicates with each other to deliver the required capabilities. Thus, tracing of relationships through the definition and decomposition process of systems engineering is an integral part of FoS requirements traceability and is preferable to be accomplished at the architectural level (Salmon et. al., 2015). The benefits of traditional systems engineering approaches such as the Dependence Structure Matrix (DSM) is it can help characterize the dependences / requirements (Johnson, 2010). Dickerson & Valerdi (2010) describe relational transformations as incurring the ability to provide a mathematically based formalism for model transformations that permit precise computation of the transformation of parameters used to model a system. The methodology integrates and

extends the work done by Suh (1998) on axiomatic design. The formalism is established in the theory of models in mathematical logic but has been adapted to the practice of engineering.

5.1.1 Why use a Specialisation of ROSETTA for the FTS

“If we are to achieve results never before accomplished, we must expect to employ methods never before attempted”
– Sir Francis Bacon

In early stages of system design there are typically a number of competing designs; simulation of the alternatives is used to explore variations for designs and process enhancements. The derivation of mathematical models to accurately describe real world problems is, most of the time, overwhelming and an improbable task due to the complexity of multi-attribute problems and inherent ambiguity of characteristics that these problems may possess (Ganguly & Krishnamurk, 2015). Simulation provides systems engineers with the ability to study the SoI under specific scenario conditions. With the models describing both the systems architecture and the scenario representation under computational control, analysis of behavioural parameters for system performance should be possible before the implementation task is performed.

M&S adds to the strength of the QFD approach by creating quantitative transformations between the input and output variables; the difficulty is then identifying the surrogate models which will best describe the relationship or sensitivity between an input parameter (requirement) and an output parameter (design metric) of the framework. The model characteristics can be based on historical knowledge that can constrain the exploration space: physics models that can model physical properties of designs by using previous mathematical function solutions that have proven to accurately describe the relationship, or data gathered in file format from real tests (Meyers et. al., 2009). One of the simplest practices for design space exploration is the response surface methodology (RSM) (Scott et. al., 2011), which is a key enabler for performing rapid trade studies and to explore the effects of system attributes on system responses.

The ROSETTA framework captures the behaviour of the M&S environment in the area of interest by using the RSM to encapsulate a set of surrogate models (response surface equations (RSE)) around the M&S environment. The RSEs enable rapid execution of the simulation and permit trends across the design space that should be easily understood and quantified by users (or decision makers) of the framework and thus assist in optimisation and

influencing the magnitude or strength of the relationships to be represented in the main body of the framework. The RSE provides a direct translation between information contained within the RSE and the information elicited by SMEs. The desired relationship, therefore, concentrate on the design of experiments (settings of each variable in the framework) between the requirements and the design metrics. The mappings describing the relationship, is performed by a transformation frame. These mappings then leverage the mathematical foundations of relational orientation as discussed by Dickerson & Valerdi (2010) to ensure precision. If the sensitivity between a metric and requirement vary as a result of other assumed values of other metrics, then this implies that sensitivities cannot be represented by a single value (as per QFD), but with partial derivatives that are functions of the other metrics. RSEs are polynomial regression of the model; therefore, the behaviour of the M&S environment can be explored through the behaviour of the mathematical functions (Ajao et. al., 2012).

These surrogate models create a prediction profile within the ROSETTA framework that specifies the sensitivity of one metric to one response; the result of which should determine any coupling between metrics. The prediction profile includes hairlines which are used within trade-studies to reflect changes in system relative to a change in system parameters enabling the decision maker to determine the appropriate system output using the RSEs and hence use the M&S framework to determine whether the change still satisfies all customer requirements and objectives and whether additional trade-studies are needed (see ROSETTA 1&2 in Chapter 7 for further details).

ROSETTA uses a subjective approach to develop the relations between framework parameters. This desired relationship can be obtained using a modified version of the M&S approach taken in Technology Identification, Evaluation, and Selection (TIES) method (Mavris & Kirby, 1999), producing the needed information through exploration of the full design space within the ROI so a parametric environment is required to capture the behaviour. These surrogates are typically continuous functions through the design space: it is possible to visually and mathematically examine the shape of the design space. The general form of a second-order RSE is shown below, where the error term (ϵ) is assumed to have a normal ($N(0,1)$) distribution

$$R = b_0 + \sum_{i=1}^k b_i x_i + \sum_{i=1}^k b_{ii} x_i^2 + \sum_{i=1}^{k-1} \sum_{j=i+1}^k b_{ij} x_i x_j + \epsilon$$

Where,

b_0 is the intercept coefficient.

b_i is the regression coefficients for the linear terms.

b_{ii} is the regression coefficients for the quadratic terms.

x_i are variables of each input variable which affect the response.

Surrogates are created by selecting an appropriate Design of Experiments (DoE) and executing this on the actual codes. If the assumed type of surrogate model does not meet the desired accuracy, then a new form is assumed and a new DoE is created; the process is repeated until acceptable surrogates are created. The added advantage of using a mathematical approach in this way is to provide complex knowledge in a M&S environment in a computationally inexpensive way, leading to better allocation of resources and assist in traceability and decision making tasks from the requirements definition and allocation process through to system realization to address the research question (Tashakorri & Teddie, 2006). As a result of this mathematical simplicity, the RSEs in the prediction profiler should permit portability and the ease of understanding; and allow the SMEs to decide on the magnitude and shape of the relationship: this will be highly dependent on the assumptions about the given problem and how much information has been provided about the domain of interest.

5.1.2 Modelling Differential Equations

'One expects a mathematical theorem of a mathematical theory not only to describe and to classify in a simple and elegant way numerous and a priori disparate special cases. One also expects 'elegance' in its 'architectural', structural make-up. Ease in stating the problem, great difficulty in getting hold of it and in all attempts at approaching it, then again some very surprising twist by which the approach, or some part of the approach, becomes easy, etc. ... – all this is much more akin to the atmosphere of art pure and simple than to that of the empirical sciences'.

- The Mathematician by John von Neumann (von Neuman, p.196, 1947)

The sensitivity between requirements (input) and design metrics (output) can be interpreted as the slope of the tangent to a response surface and any point. This two dimensional surface can be a simple straight line and the sensitivity may depend on both the value of X and the amount of change Δx about that point. The objective is to evaluate the sensitivity at some nominal point (X1, X2) such as the point defined by the mean or median of X1 and X2. At any point, the sensitivity of the model output, $Y=f(X1,X2)$, to one of the inputs (x_1 or x_2) is represented by the rate of change in Y per unit change in X. This is the slope of the surface at

that nominal point in the direction of X and is expressed as $\frac{\partial Y}{\partial X}$. If the relationship between Y and all of the inputs is linear, then the response surface is a flat plane and each of the partial derivatives at each point (X_i, Y_i) , will remain constant regardless of where the point is on the surface.

For differential equations, a derivative function is used to specify the change of state variables; thus at any time, a given state and an input variable, the only known parameter is the rate of change of the state; thus, the state at any point within the response curve/data has to be computed, in this case, by virtue of adjusting the hairlines and with it the magnitude of the partial derivative slope. Hence, the state variable 's' is represented by the current change in 't' of the variable representing the state 'x':

$$\frac{ds(t)}{dt} = x(t)$$

And the output 'y' will be equal to the current state:

$$y(t) = s(t)$$

The partial differential gives a powerful mathematical relationship between independent variables, which can be used as a measure of robustness that can be visually examined when overlaid on a graphical display with either the surrogate model(s) or the actual response data measured between the two variables. For a robustness metric the partial derivative value should be minimised, however, trade studies have to be performed to maximise or minimise the output $y(t)$ depending on the requirement objectives of the system. The trade study analysis problem is further exasperated when multiple relationships exist between multiple independent ROSETTA framework parameters, consequently, the problem then involves trade-offs between optimized system outputs and desirable system outcomes.

5.2 FoS Model For Assessment Of Blending Mix

Understanding the relationship between the VRTE capabilities and the required training goals should allow the tool to be used to achieve training objectives (Boosman, 2014). Having a certain standard of VRTE will permit the rostering of the pilots to the most suitable tool to economize on real flight training hours. It is necessary to qualify the training devices (MARFSTD, 2011), the configuration of VRTE, and tasks to enable a checking facility on which tasks can be performed to achieve the current planned training goals, which is comparable to the latest ICAO document 9625 (ICAO, 2009) (see Volume II, Table IV for a brief summary). The tolerance levels are dependent on specific features and are established with SME

consultation on aspects such as: switchology/buttonology, pilot UI, performance (range, accuracy).

Some tasks require a limited number of sensor inputs, i.e. instrument flying generates less motion and visual cues than basic flight manoeuvres (Emre, 2016). The general rule of thumb being used when simulation is possible within human sensory limits is: the FSTD is capable to replace actual flight, thus $F(A) = R$, where A is the simulation model, and R is the actual flight. Where the simulation effect differs considerably from reality, the lower boundary rating is given for the respective task being trained. Thus, it is important that the models used for training in the VRTE should be consistent with the level of training sought. This subjective rating for minimum levels is done within the qualification of the FSTD (JAR-FSTD, 2008) and the associated tasks to be flown are rated by SMEs who should consider the levels of K&S required in relation to the capability of each disparate training system/technology.

The uses of MBSE graphic modelling techniques to model the system in manageable blocks represent greater clarity in the assessment process. The ROSETTA frameworks provide structure against which assessment on suitability of the technology/FoS system can be performed to provide a transparent mechanism for SMEs, managers and students. The suitability assessment should be driven by the mission scenario training attributes, which will describe the intent that the technology should satisfy.

To add to the complexity of choosing the correct blending mix is the advancement of emulation (synthetic) functions on live aircraft to reproduce training elements for sub-systems and weapons. A single element of training may be conducted in the synthetic environment for some of the repetitions and a live re-enforcement activity to ensure ToT. If the competency or task being trained is adequate to be evaluated in a VRTE then the decision is easy, however, time and stress factors are a critical aspect of the specific task then the synthetic environment on a live aircraft maybe the only way to ensure ToT for this competency level. Of key considerations are to ensure that every training mission is well planned and of value, and that the performance is evaluated to clarify that ToT has occurred (Ayers, 2006). To enhance the physical fidelity feedback to the pilot, the mission tasks are performed and evaluated in a live environment and comparisons in the performance levels between ground based and airborne training environments can be completed to ensure competency and that a level of readiness has actually been achieved.

The issues surrounding these aspects is revolving around the amount of fidelity actually needed to satisfy training goal(s), as per brief discussions in Chapter 4.1. The increase in fidelity leads to increase in costs, increase in cost often leads to fidelity compromise (Williams et al., 2004). A number of objectives required for decision support are (De-Braun, 2007):

1. Train pilots with a blending mix to gain experience from strategic to tactical levels.
2. Integrate LVC elements within a common synthetic environment
3. Ability to practice mission scenarios to meet objectives within proficiency tolerances
4. To meet organisational goals to training
5. To allow tailored training for individual pilots.
6. To achieve a level of pilot readiness for operational needs

To link these objectives to the fidelity of VRTEs would be beneficial in improving the MECs of the training system for improvement of pilot readiness for the front line (Colegrove & Alliger, 2002). The facility to assess, track, and compare performance of pilots on each blending mix would provide meaningful information of the choice of technology to meet operational goals. The mission scenarios used for the assessment should represent a spectrum of difficulty. The choice of blending mix/system then should be made on the basis of cost and learning objectives rather than on the basis of technical or fidelity criteria (Nyssen et. al., 2002). Before training can commence in earnest with any training system the pilots must familiarise themselves with the switchology and buttonology (look, shape, and feel, and functionality) of the training technology in the cockpit to aid memory and execution of SoPs. These SoPs have to be memorised by pilots and evaluated before actual flight can progress reflecting motivation, safety and good pilotage. The relationship between fidelity and ToT can be misleading if consideration of the learning aptitude and stages of the pilots is ignored. Prediction of performance in the real world and the measure of capability and suitability of a VRTE tool to ToT is effected by: the learning stages of the student; the ‘comfort’ and ‘confidence’ of VRTEs to adequately train the appropriate levels of K&S; and pilot’s own opinion of use of such technology. Current data about the VRTE success may reflect the instructor effectiveness with the pilots or the pilot’s motivation in real flight rather than the degree of VRTE effectiveness and ToT (Roessingh et. al., 2009).

Currently physical fidelity approaches dominates training design for transfer, and the lesser known psychology fidelity, which concentrates on a training system as a series of experiences that build key skills from basic to strategic to more complex adaptive skills (Kozlowski, 1998), has been generally ignored. The logic of psychology fidelity is based on the use of

theory to design simulation experiences that induce cognitive processes that are critical to performance requirements in the target domain (Graham, 2004). Buttonology refers to how interaction via keyboard, joystick, mouse, etc. and the mapping between the real world and the simulation world becomes a problem due to the inappropriate interface of the VRTE. It is known that more time practicing these functions and familiarity with layouts and interfaces, issues become less evident during mission scenario execution (Kennedy et. al., 2010). During practice, however, if the simulations become easier to execute, repetition needs to be limited so that they do not become ineffective as a training tool (Haslbeck et. al., 2014). Slight variants of configurations or training goals can give new motivation for reaching learning outcomes. It is this process which can assist the decision maker to decide which fidelity of VRTE can be used to effectively evoke the required response from student pilots based on SME declared ToT levels (Boosman, 2014). The fidelity dimensions required for the FoS assessment using the ROSETTA frameworks is kept at a high level of abstraction and the fidelity characteristics of concern for the research project is listed in the Table 2, a modification of the PERFORM project (Estock et. al., 2006).

Table 2 Fidelity Dimensions (advanced from Estock et. al., 2006)

Fidelity Dimension
Visual
<ul style="list-style-type: none"> • Visual Scene Display Field of View • Visual Scene Display Resolution • Visual Scene Display Object to Background Contrast • Visual Scene Display Detectable Lag • Visual Scene Quality of Information
Motion (<i>beyond the scope of this thesis</i>)
<ul style="list-style-type: none"> • Platform Motion • Manoeuvring Motion Cues (e.g. G-cueing) • Disturbance Motion Cues (e.g. Engine Out)
Physical (Cockpit)
<ul style="list-style-type: none"> • Shape and Layout of cockpit Controls and Displays (switchology) • Content and Operation of Cockpit Controls and Displays (Buttonology) • Ergonomic Control
Audio
<ul style="list-style-type: none"> • Aircraft Operating Status Sounds (e.g. Engine Noise) • Aircraft Alerting Sounds (e.g. Viper Spiked Alert) • Communication (e.g. Synthetic AWO)
Task / Psychological Fidelity
<ul style="list-style-type: none"> • Simulation Object Model Accuracy • Student Pilot Feedback Relationship • Performance Feedback (to Baseline)

The effectiveness of training scenarios, and more holistically the training syllabi, used with a specific VRTE is measured by examining pilot's performance to a comparable 'benchmark' (Baseline). The difference in performance is determined by both objective and subjective performance data and from previous completed mission scenarios (Symons et. al., 2003). Mathematical functions are then used to define relationships between K&S elements and the fidelity dimension, which also accounts for the strength, direction (positive or negative correlation) of the relationship. The functions are used to describe the changes in training effectiveness with a given fidelity level. Symons et. al. (2003) developed a rating criterion to encapsulate the relevance of a mission scenario to evaluate the air-to-air K&S elements.

The ability to establish a quantitative relationship between simulator fidelity characteristics and training attributes will provide information to enhance the value of training programmes that uses various blending mixes to train pilots. This measurement system provides additional strength to current feedback systems and is designed to work together synergistically to leverage current learning processes for both student pilot and instructor. This active learning process (Kozlowski et. al., 2001) promotes a mastery and understanding of the task domain and places training objectives, rather than performance objectives at the foremost in the training system; the sequence of practice shifts attention from basic K&S to more complex knowledge and strategic skills using the available blending mixes as the training programme progresses through the pipeline.

5.3 Analysis of Organisational Objectives

The flight training system problem involves a number of non-deterministic systems for which a number of non-commensurable objectives exist. The main objectives for the flight training systems are to reduce cost of training and to improve pilot readiness for the front line using blending mixes. These objectives are multifaceted and negatively correlated. It stands to reason that using the more expensive real flight option for training will lead to greater pilot readiness and more efficient acquisition of the relevant K&S and stress management. However, there is an argument that more advanced technology being used for flight means that workload for the pilot is more directed at system management and monitoring while the technology in the flight management system (FMS) (Wagener & Ison, 2014), notably the fly-by-wire, will take care of the more complicated flight demands. This has led to more interest in training within VRTEs to ease the economic pressures of real live flight for training purposes to familiarize the pilot with cockpit and MFD control and monitoring tasks. With a

number of disparate VRTEs with multiple configurations available and a diverse scope of training attributes the blending mix needs to be capable of satisfying, a combination of qualitative and quantitative (mixed method) data sets should be used to optimize a training programme for efficient progression of students through the pipeline.

The aspect of training should be based on obtaining the correct mix of K&S to produce an effective training system concentrating of the ToT in relation to business and organisational goals. When considering the choice of blending mix to use, attention has to be given to both the goal(s) of the training and the information required to improve effectiveness of K&S levels. The choice, therefore, should be based on a number of factors surrounding instructional effectiveness, e.g. geographical distribution, platform availability, budget, and content stability. However, research in the choice of which technology to use has found no significant difference for learning effectiveness (Miller et. al., 1997);(Higgins et. al., 2012). There is also an argument that good design in instructional process will compensate for less than ideal choice of delivery media (Kirkwood & Price, 2012);(Rajkoomar, 2013).

The systems architecture should give information about the differences between the technology choices in order for a rational decision to be made about which technology to use for which training scenario. Furthermore, it may result that there will be more than one solution that will satisfy the training goals. The choice then must be based on a disparate set of considerations, such as: buttonology/switchology, cost and availability. The art of the designed architecture is to assist in and provide more accurate trade-off decisions. For each training stage, a description of desired performance and rules need to be clearly understood for the student pilot to be successful in gaining the relevant K&S, hence, the blending mix has capability attributes to provide the means of leading them through a learning process and find their own desired learning rate.

Probabilistic analysis along with self-assessment feedback is required to ensure training progress through the pipeline ensues as efficiently as possible with the ability of removing biased data sets from the analysis. The ideal data source for analysis includes technological artefacts, SME opinions, and other sources which can provide statistical significance in the analysis. The difficulty with decision making within a flight training system is evidence is often not available for the following reasons:

- Patterns of technology refresh are difficult to predict and manage as new technology inherently means new features, which needs to be taught to all relevant stakeholders (Gelijns et. al., 2005).
- Data sets which are available are often interpreted differently by SMEs, especially in the presence of high uncertainty; especially about pilot readiness and technology ToT effectiveness (Mohammed, 2001).
- In the case of clear evidence, shown by improved performance and accuracy of execution of mission scenario, SMEs may disagree on the implications due to differing value systems (Beck, 2003).

Statistical and feedback fusion is necessary to measure the extent to which information supplied for the analysis is correct. Decision support then becomes a matter of managing the accuracy of the information supplied into the ROSETTA frameworks for the purposes of simulation and analysis. Each available blending mix will be assessed via optimization algorithms or sensitivities between framework parameters based on the main objectives and K&S level(s) capability. Evaluation of meeting the objective criteria, therefore, is based on SMEs identifying one or more pre-specified termination criteria based on an upper or lower bound on the variables (training objective and MECs) being assessed on the current mission. As the mission is based on goal attainment, the analysis can be used to determine whether a student needs to practice, require training on more advanced VRTEs, whether there are issues with suitability of student on training attribute levels of the chosen blending mix, or is ready to advance difficulty of the training missions.

The primary objective of the ROSETTA framework(s) is to provide decision support for the choice of optimal blended mix for specific mission goals. The ROSETTA framework(s) will assist decision makers (stakeholders) to decide on the most appropriate blended training solution across all military training domains and permit assessment on cost effective training delivery for the operational requirements (Fletcher et. al., 2011). The ROSETTA framework(s) within the workflow should allow for flexibility, and eventually after a number of iterations for evolution of the relationships, produce an agile DSS relevant to operational demands.

5.5 Chapter Summary

In most design processes (training included), there will be more than one right answer to satisfy the requirements. A choice between solutions must be made based organisational objectives, which may include a set of disparate considerations; the art of design is making a decision based on these types of trade-off decisions. Defining the correct choice of training

delivery system to use is based on the goal to provide knowledge, skills, and information with key measures including, cost, time frame and the degree of reliability on acquisition of the key organisational goals. For each training schedule, key parameters including: the learning environment, depth of learning (taxonomies of learning), and training system characteristics (i.e. fidelity levels), are required to be addressed by the decision makers. Focussing on organisation effectiveness could cloud the reality of training effectiveness when a choice of blending mix is being driven by high order requirements such as, cost, time, and platform availability. Instructional processes such as: training application, feedback, and objective evaluation of performance, has to be at the forefront of the decision making strategy. Understanding the intended results of the trade-off choice, with respect to training attributes, should be managed on the front line with the ability to communicate to all stakeholders the outcomes of the training scenario and the suitability of blending mix for the ToT.

METHODOLOGY FOR FTS DESIGN PROCESS

'Following the path of the physical sciences, [engineering design researchers] have come to believe that systems are also precise in nature and can be efficiently analysed by classical mathematics. To be precise, we have attempted to force artificial precision on imprecise phenomena and processes, and in so doing have lost the intrinsic imprecision in human systems in search of precision as a goal. [Design] models which neglect these intrinsic characteristics tend to be over-simplified, too mechanical, and too inflexible to give an adequate description of the complex and elastic real world'

- (Leung, p. vii, 1988)

Outline of Chapter

This chapter commences the development of the methodology used for working towards solving the research problem identified within the requirements narrative. Information from a site visit to a RAF training facility is used for preliminary analysis of current understanding of the problem, which is then mapped to the requirements narrative. The methodology then describes a novel method of requirements analysis using models with system elements incorporated to add new meaning to the interpretation of the requirements. Additional knowledge is then gained with a holistic viewpoint on the flight training enterprise to identify any missing relationships. Functional analysis of the requirement statements is then examined to produce Use Case models followed by a brief introduction to the functionality of the prototype process workflow tool presented by the methodology.

Knowledge Gained from Site Visit

The site visit to the RAF developed an understanding of how the training providers produce an appropriate evaluation plan that integrates with past; but with it an admission that more dependable training method is needed to manage complexity. An abstract method of training produced from this site visit is described as follows:

1. Before execution of missions, pilots are given initial training of how to use the technology including familiarity with switchology and buttonology so as to allow pilots to become comfortable with the controls.
2. Introductory training mission scenarios are planned and executed on the training technology with a gradual increase in difficulty until a reasonable level of efficiency is reached.
3. Supervisory training missions then commence - training pre-mission briefs are done with adequate time for questions and answer sessions.

4. The pilot is then required to execute the mission scenarios with the most convenient available VRTE or aircraft, with an instructor observing the actions of the pilot where possible.
5. If any incident occurs during the mission scenario (i.e. a real or predominantly virtual casualty from a crash or hit), most FSTDs reset the aircraft in the air at some predetermined point to 'carry on' with the mission – this can cause a reduced experience of risk and casualty. The post-mission brief would discuss the actions and events that caused the incident.
 - a. Some mission planning software (i.e. AMPA – used for mission planning and playback) lack the necessary 3-d imagery of the mission scenario tactical manoeuvres, thus, it is difficult for the instructor to use to explain pilots decisions and actions post execution. A legacy post-mission evaluation technique using 'aircraft on sticks' is used to fully post-mission brief the pilot on what they did and what they should have done, etc.
 - b. Most aircraft and FSTD cockpits (if available) lack the viewing of pilots actions (this facility is somewhat supplied in the highest fidelity FSTDs (dome 'FMS' flight training simulators)), this feature enhances the post-mission briefing commanding a clear delineation on how the pilot is reacting to external and internal stimuli.
6. Using a desktop simulator, the pilot(s) have the opportunity to practise the mission scenario and gain confidence and K&S; however, this practice is generally performed unsupervised and could potentially lead to 'bad habits', which could then be demonstrated in the assessed training mission (Campbell & Kuncel, 2001).
7. The pilot would be asked to repeat the mission scenario following the K&S learned from the simulated flight and the post-mission briefing session, for SME evaluation.
8. Once the pilot has reach the required level of competency, judged by SMEs, the mission scenario(s) will be amended for account for greater level of difficulty and the skills learned will be re-enforced in real flight and assessed in the traditional manner.

VRTEs appear to cause concern when it relates to the reduced experience encountered during training accidents, i.e. aircraft collisions and the lack of feedback information. This evidence provides critical information with respect to the design of the methodology for the research project, especially when it relates to feedback data from both the training technology and the stakeholders (e.g. decision maker or instructor, student pilot, training developer, etc.). The

following sub-chapter discusses the methods used to develop the system description to bring precision to ambiguous natural language requirement sentences.

6.1 Logical Modelling of FTS Requirements

“Logic is the art of going wrong with confidence.” - Joseph Wood Krutch

The art of requirement engineering lies in the ability to extrapolate sentences from natural language to produce premises that are meaningful to the system’s design that will add detail to the overarching concern. In logic, these sentences form valid arguments that are said to imply a conclusion or result from the premises to satisfy the systems objectives. These sentences must produce arguments that are always true to ensure that the conclusion of the designed system is valid (Angeli & Murray, 2014). This validity, however, is not the same as truth or with the way systems actually are in the real world. The art of eliciting the truth and drawing conclusions lies in the deductive reasoning process associated with logic and mathematics (Pawar, 2009). Furthermore, this type of deduction is based on assumptions and rules divorced from experience and generally gathered from textual reports and documents that describe operations which are referenced to the real world, but divorced from it; this forms the contract that has to be satisfied by the systems engineers. Nevertheless, for systems design we are concerned with the real world and thus need to insure a kind of scientific reasoning independent of logic, based on the scientific method that is used for inferring inductive arguments, from the particular to the general using a human perspective involving experiences and events that has been observed (Pawar, 2009). Thus, a good and valid system is said to incur both truth in deductive reasoning AND satisfy, with a high degree of confidence, the general argument statements (concerning attributes belonging to actors and/or entities in the system) gathered from individuals with experience in the domain being investigated since there is a possibility that the requirement narrative has been written by non-domain or job specific experts.

Hence, using science and philosophy for discovering the truth behind problems, especially involving humans in the loop, both deductive and inductive reasoning is vital to the investigations to find out exactly how ‘things’ work (Rothchild, 2006). Deduction enables clarity of implications of assumptions without being influenced by perceptions or our immediate experiences and biases, and the truths or conclusions should always be satisfied; induction enables generalisations of experiences to influence the ‘end-product’ to account for

real world experiences and enable versatility of the designed system based on contribution by real actors of the system.

The deductive reasoning process should be used to develop the system initially, but not as an absolute solution to the design problem. The precise language, logic, or mathematics should be used as ideals to validate a system design, but then used as guides in the scheme of gathering performance of activities information of the system and its environment, which are totally dependent upon human sense experience to actually develop a system that will work in the real world.

‘There is a world of ideals that has even more reality to it than the world we experience. It is the world of perfection and absolutes. It is the reality we must continually strive to achieve and according to which we must pattern our lives and our works in this lesser, the imperfect world that we know of as our physical existence.’
– (Plato).

6.2 Requirements Modelling

A model driven design provides precision and comprehension; and some requirements incur intrinsic ambiguity or ‘on the surface’ unrelated to the overarching system functionality (Berry, 2007). To bring precision to requirement sentences, an approach referred to as transformational grammar is used to produce requirement models based on structured well-formed logic from the natural language (Schurr & Selic, 2009). This type of logical modelling aims to extract explicit relationships that describe the logical semantic meaning of the sentence to derive a minimal model of the relations that is complete and captures the intended meaning (Dickerson, 2008). Generally, the ‘nouns’ of the sentence describe an entity of the system and the ‘verbs’ capture the relationships between the entities.

The defining term of each identified sentence gives an indication of the overarching system functionality and is highlighted within the sentence and underlined. The key words give the defining term meaning by searching for descriptions that describe relations between words; these will be modelled in a block definition diagram using SysML. Entities and relationships from the identified word(s) need to be realised within the model appropriately by either a block or an association name, i.e. nouns are blocks, verbs and relationships are on the association lines. If the model of the requirement sentence appears to be incomplete, i.e. the overarching system theme is overlooked, additional noun(s) and relations are inserted into the logical model and the natural language definition amended for concordance and subsequently

circulated to all stakeholders for contract approval before the system design can progress to further stages.

The research project commenced with vague requirements describing the problem of the choice of blending mix to support training capability (re: requirements narrative). Recognition of the ambiguity of words and relationships between words within a requirement involves forming a structured sentence within a model. The model(s) should permit the visualisation of the system architect's interpretation of the requirement within the modelled sentence. These logical models form the detailed structure from which any further derived requirements, which are needed to fully comprehend the true meaning of the natural language requirement(s), becomes perceptible. The generation of these logical models form the CIM within the MDA process (OMG, 2014). The requirements elicitation process used for the research project, modified from Soley (2004), is as follows:

- 1) Elicit Customer Requirements (Industrial Sponsor).
- 2) Capture keywords (nouns), which become blocks in SysML.
- 3) Transform natural language requirements into mathematically precise models using keyword as blocks in SysML.
- 4) Identify semantic relationships between the blocks of the requirements (generally verbs).
- 5) Analyse and derive addition requirements to resolve interpretation issues.
- 6) Form the CIM.
- 7) Seek confirmation from stakeholders about correct interpretation of the natural language requirements.
- 8) Goto (5) if amendments are required.

The advantage of logical models is the determination of the relationships between defined terms within the requirements that describes the minimum detail to extract the real meaning from the natural language counterpart i.e. all intended relationships/semantic relations are captured in mathematically provable concise logical models (Dickerson, 2008). The logical models form the foundation for systems design within the MDA process, as the architect needs both '*formal definition of a concept, for precision, and a natural language definition for comprehensibility*' - (amended from Brooks, p.234, 1995).

Sentence modelling requires one or more 'law-like statements' of contrasting generality followed by a 'symbolism' organised into a conditional expression to enable the models of the statements to be exposed for assessment of their truth value. The model of the sentence must capture the belief about the nature of parts of the world and preserve meaning. As with

any system, states of any part can be seen to change as a result of an activity, thus, the requirements model must describe interactions between parts of the system of which the function is to induce changes in system elements for their benefit or the system's intended outcome.

The model must show all theoretical objects with its own quantifiers and give a strong indication of how these parts are connected together. A static structure is used (recognised by qualified relations as static verbs) to represent a part for the world or a state (e.g. believe or opinions), or a dynamic structure is used (recognised by qualified interactions as dynamic verbs) to represent an activity (e.g. action or event). The symbolism must preserve the processed natural language derived from a story of a scenario which is the general means of representation of a system, communication or even a model (SEP, 2014). The transformation of natural language into its basic constituents of one or more sentences of which complex static or dynamic structures can be constructed into ordered pairs $((a R_1 b)$ for objects and relations and their qualifiers (static state), or predicate logic statements $(P \vee Q \rightarrow Q)$ for objects and interactions and their qualifiers leading to change of state). The flow of requirements modelling, therefore, follows a formal linguistic order as follows: (For all object(s) there exists a function (an interaction or relation) that map to other object(s) which implies a desired outcome.)

$$\forall (O_i) \exists (f) \mapsto O_r \Rightarrow d_o$$

Four invariants are used to describe sentences which make up a scenario:

1. Theoretical Objects (System elements or entities) – {concrete, abstract, symbolic}
2. Relations – structured properties between system elements.
3. Interactions – behavioural properties that affect system elements
4. Qualifiers (adjectives) for selection of class elements – to assist describing actions or functionality of system elements.

When one or more are used together, the invariants form the whole model that describe the scenario and the objective of the system. The logic of requirements modelling is to test the 'soundness' of arguments and make precise statements and conditions that can be mathematically provable and to remove misinterpretation issues between stakeholders and system designers. The ontology used to control and ensure the relevant invariants are identified to describe each requirement sentence, is summarised in Figure 17.

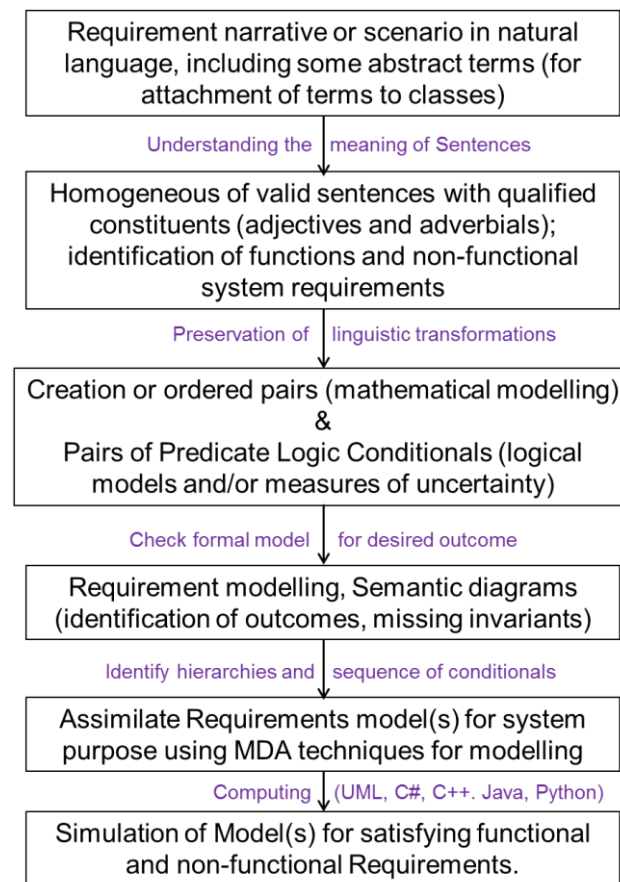


Figure 17 Process and Structure for Linguistic requirements Modelling for Systems Design

The process should allow the establishment of a fundamental view of the empirical systems phenomenon. The structure gives a robust method with which to clarify understanding by the transformation of natural language scenario into a human and machine readable model whose semantic interpretation is strengthened by logic reasoning. This type of linguistic modelling takes advantage of the structure of model artefacts to tell a ‘story’ in a mathematically provable scheme (Ciuni et. al., 2014).

An understanding of the domain in question is needed as some language that describe objects and relationships may be unsuitable to use for requirement engineering. A simple ontology model, illustrated in Figure 18, using No. 1 & 2 of the four invariants, describe at a high level the aspects the model architecture must possess. The syntactic quality is needed to ensure that the model describes the requirement sentence without changing the ‘flow’ of the sentence and hence the meaning of the requirement, i.e. the construction of the model is governed by the rules and principles applied within the natural language sentence. The semantic quality concentrates on the degree of concordance between the model and the domain that is being modelled; it is used as an early indicator that the model is an accurate representation of the domain being modelled and is concentrated on the validity or legitimacy

and completeness of the model to the domain or a set of requirements. This is generally managed by the creation of a common ontology controlled by a set of rules that is used throughout the design process to ensure a high degree of continuity between all stakeholders, i.e. knowledge or word meaning and their relationships to each other are language specific to avoid misinterpretation or misconception. The developed model(s) should be used for the system design process, but also be inclusive of human comprehension (see Chapter 2.3). Furthermore, the model(s) should be assessed for pragmatic quality with all stakeholders, especially with the design and/or software engineers. A model of a high degree of quality should ensure a low degree of misinterpretation for stakeholders and high probability of lifecycle use even after implementation of the designed system model.

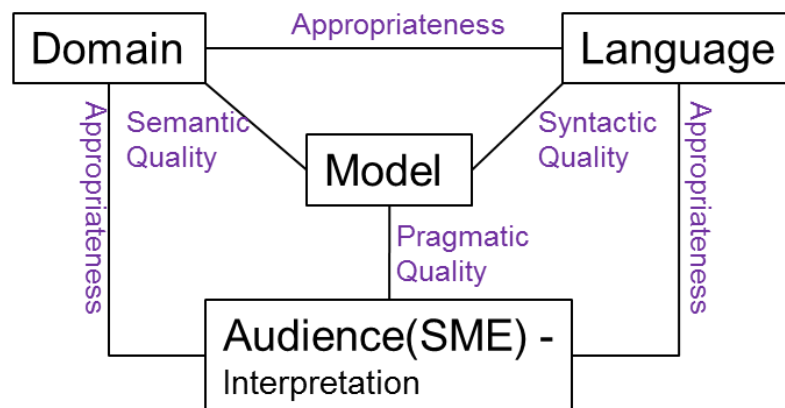


Figure 18 Simple Communication Ontology Model

6.2.1 FTS PhD Requirements Modelling

SysML provides the facility to model non-functional and functional requirements via the requirements diagram. The requirement sentences are gathered from the original narrative in Chapter 1.4, and list below:

- 1) Provide structure for Live/ Synthetic (airborne and ground based) aircrew training and a modelling, simulation and analysis approach for the FoS aviation training problem.
- 2) Support capability based trade-off decisions to select optimal flight training FoS mixes.
- 3) The ROSETTA framework will permit allocation of systems in the mix at the task and activity level for Mission Essential Competency (MEC).
- 4) Will provide automated assessments of the cohesion of the FoS architecture.
- 5) Simulation of the FoS architecture will be used to assess the capability to support a coherent scheme of training for the aircrew under realistic operational conditions.
- 6) Pre-flight mixes can be assessed in trade-off analysis.

- 7) Simulated performance of the selected mix can also be compared to actual performance in the post-mission analysis.

The main objective of the project is to develop a methodology, inclusive of the ROSETTA framework to support trade-off decisions; the requirement sentence stated in (1) is used to produce the ‘super’-requirement (the overarching customer requirement): The decision support system will offer ROSETTA framework structure(s) for Live / Synthetic (airborne and Ground Based) blended aircrew training and a modelling, simulation and analysis approach for the FoS aviation training problem: to remove any technology dependency from the requirement models, as per MDA techniques, the ROSETTA framework will be hence forth known as a framework. To assist in the organisation of the design of the FTS workflow process described in the methodology, the requirement sentences are organised into two main requirement packages and illustrated in Figure 19:

- Framework requirements – To ensure that the design and structure of the ROSETTA framework(s) can integrate relevant data, be usable and permit easy human-readable information within it.
- Analysis requirements – The main aspect of the methodology is to permit a decision maker to analyse the data and perform trade studies; the workflow process will utilise the data within the framework and in some cases simplify the data for quick evaluation.

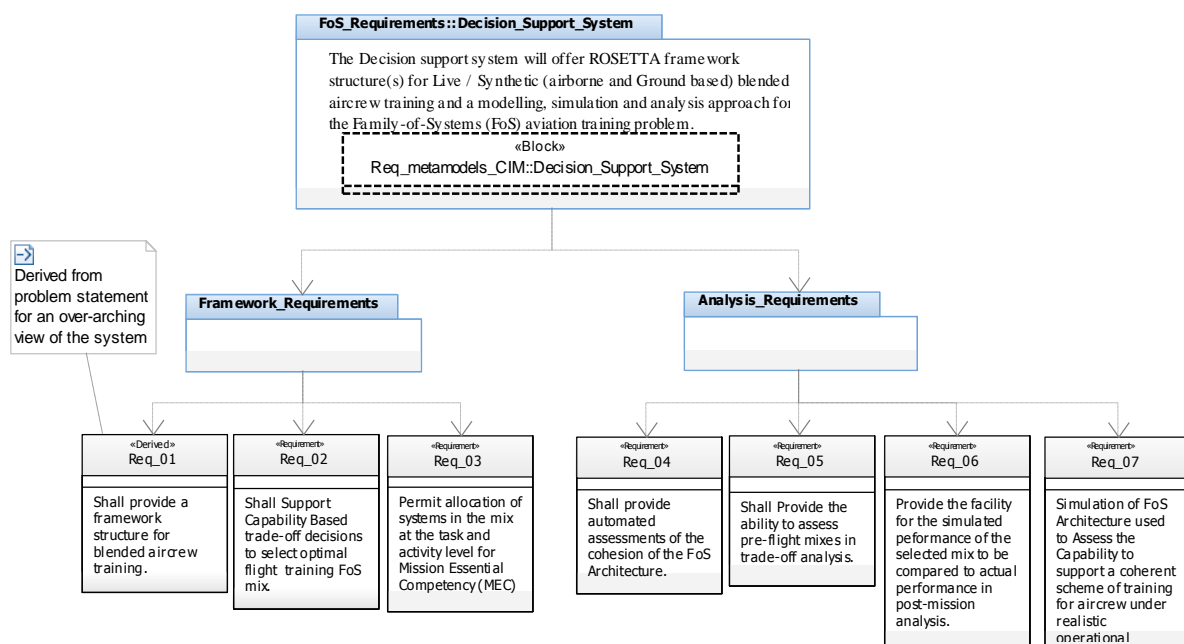


Figure 19 FoS Architecture Stakeholders Requirement Diagram

A further advancement of the linguistic modelling process is needed to account for the bi-directional nature of metamodel relational mapping for robustness. In this case producing requirement models follows the identical flow described in Figure 17, i.e. produce the models as you read the sentence; however, a greater degree of precision is needed to associate the arguments to system design considerations. The MDA technique is based on the notion of a model that gives a representation of the system to be developed, and allow us to reason about it in a simple way.

The most important aspect in modelling is not its semantic aspect, but the kind of systems it can be applied to and the kind of information it can model. These features are typically stored in a meta-model. A meta-model is a description of the concepts the modelling language provides without being sensitive to the actual syntax; the meta-model can be used for abstract grammar. Concretely, meta-models define sets of concepts and relationships between them which can be used as an abstraction filter while modelling (Kruse, 2015). With abstracting away from low-level technical details, the role of meta-models are to facilitate the use of models and moreover the integration of different kinds of models in order to be able to use them jointly.

Thus, system entities and additional roles or qualifiers are required to be identified to permit engineering characteristics to be expressed and allow the facility to communicate back to the stakeholders as to the systems architect interpretation of the NLSs given in the narrative. Therefore, for bi-directional understanding of objects and interactions the linguistic order is evolved as follows: for all elicited and derived objects, there exists a function (interaction or role) and a relation that maps to other objects which implies a syntactic structure with semantic meaning that infers a desired outcome.

$$\forall (O_{EAD}) \exists ((f_i \vee f_{rl}) \wedge f_R) \mapsto O_r \Rightarrow \mathbb{S}_s \wedge \mathbb{S}m_m \Leftrightarrow d_o$$

From Req_01, the requirement sentence can be revised to be more domain specific to read: The DSS shall provide a (ROSETTA) framework structure for blended aircrew training. In consideration of engineering characteristics of the system, the following requirement model can be developed.

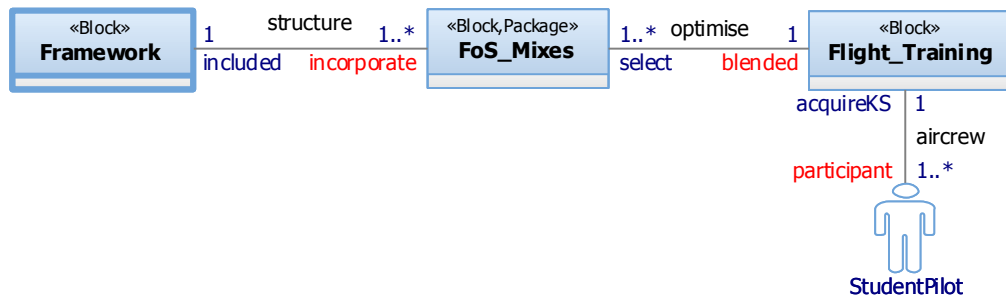


Figure 20 Logical Model of Requirement_01

The logical model in Figure 20, brings precision to the natural language requirement that can be used to communicate system specific details of the requirement to all stakeholders and permit identification of additional system entities and relations needed to satisfy the natural language requirement. From the model the new requirement sentence(s) includes:

- A **Framework** shall *incorporate* **FoS_Mixes** to *optimize* *blended* **Flight_Training** for *participating* **Student pilots**.
- **Pilots** *acquires* *K&S* with **Flight_Training** using *a selection* of **FoS-Mixes** that are *included* in a **Framework**.

This bi-directional navigation is useful for clarifying understanding of the original natural language sentence and gives information to the stakeholder about the development practice of the systems design, i.e. one or more systems in the FoS_Mixes are to be included in the framework. It also give a clear indication of our understanding of the interrelationship of all objects in the system and what aspects of FTFoS is needed to be concentrated on for the solution of the design problem, e.g. $f(\text{structure})$, $f(\text{optimise})$, $f(\text{aircrew})$.

From Req_02, the requirement sentence can be revised to be more domain specific to read: The DSS shall provide a (ROSETTA) framework that shall support capability based trade-off decisions to select optimal flight mix.

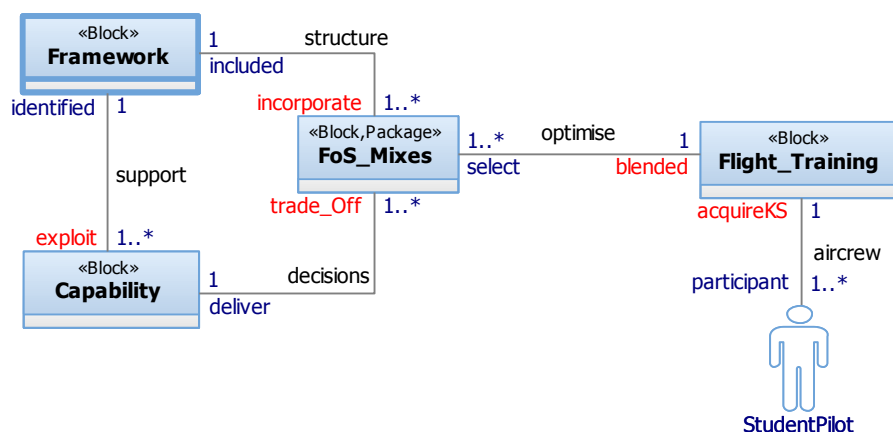


Figure 21 Logical Model of Requirement_02

From the model in Figure 21, the new requirement sentence(s) for (2) includes:

- A **Framework** shall incorporate **FoS_Mixes** to deliver **Capability** identified within the **Framework**.
- A **Framework** shall exploit **Capability** to trade-off **FoS_Mixes** to optimise blended **Flight_Training** for participating **Student pilots**.

Capability in this instance refers to both the capability of the framework and the capability of the systems in the FoS with regard to training pilots. Holistically, capability is considered to be the ability of an organisation or system to achieve overall goals. It is the capability of the FoS-Mixes which are used for trade decisions within the framework. The data within the framework is used to support the decision maker in selection of appropriate blending mixes, moreover, the suitability of a system for flight training at a specific point in the training pipeline for each respective student pilot.

From Req_03, the requirement sentence can be revised to be more domain specific to read: The DSS shall provide a (ROSETTA) framework that will permit allocation of systems in the mix at the task and activity level for Mission Essential Competency (MEC).

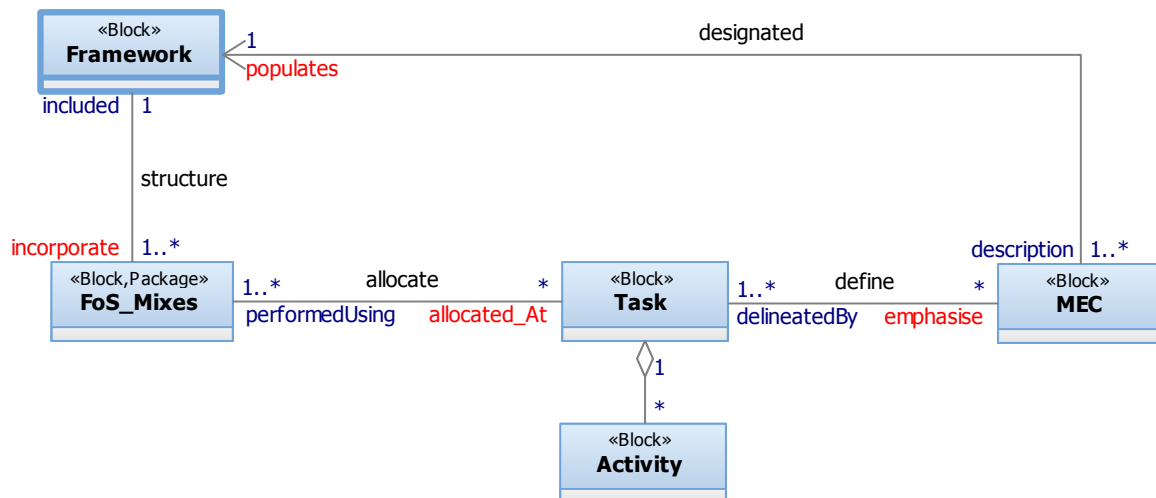


Figure 22 Logical Model of Requirement_03

From the model in Figure 22, the new requirement sentence(s) for (3) includes:

- A **Framework** shall incorporate **FoS_Mixes** that are allocated at the **Task** (and **Activity**) level with emphasis on **MECs** that populates a **Framework**.
- A **Framework** designates a description of **MECs** that are delineated by the **Task** (and **Activity**) levels, which are performed using **FoS-Mixes** that are included in a **Framework**.

From Req_05, the requirements sentence can be revised to be more domain specific to read:
The DSS shall provide a (ROSETTA) framework that will provide the ability to assess pre-flight mixes in trade-off analysis.

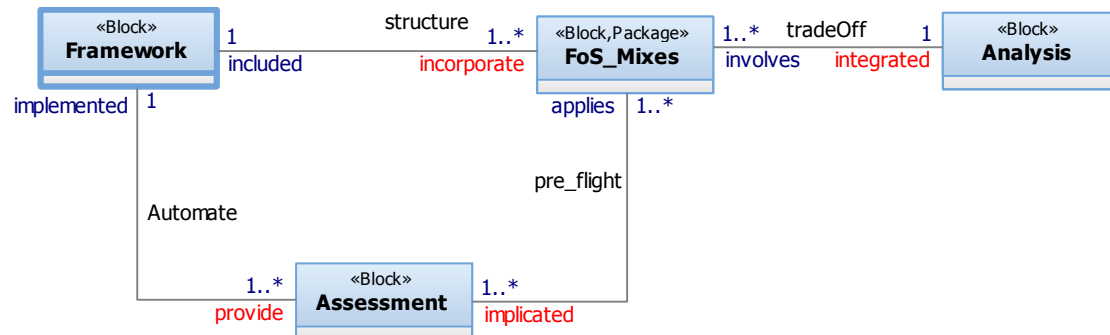


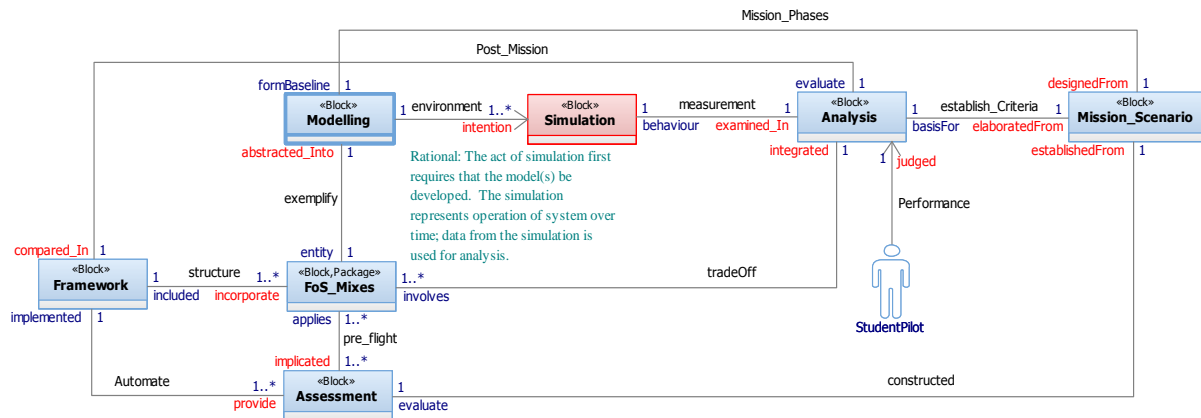
Figure 24 Logical Model of Requirement_05

From the model in Figure 24, the new requirement sentence(s) for (6) includes:

- A **Framework** shall incorporate **FoS_Mixes** that are integrated within a *trade-off Analysis*.
- A **Framework** shall incorporate **FoS_Mixes** that are implicated in **Assessments**, which are implemented in a **Framework**
- A **Framework** provides the facility for **Assessments** that applies to **FoS-Mixes** that are integrated within a *trade-off Analysis*.

The characteristics of the FoS mixes constitute the attributes under consideration within the analysis phase of the FTS. This requires a pre-flight assessment of the suitability of the systems within the FoS, for each student pilot, to eliminate technology that is unsuitable for the training mission. The subjective assessments performed within the framework are used to trade-off technology solutions and identify suitable system(s) within the FoS mixes with which to train.

From Req_06, the requirement sentence can be revised to be more domain specific to read:
The DSS shall provide a (ROSETTA) framework that will provide the facility for the simulated performance of the selected mix can also be compared to actual performance in the post-mission analysis.



Modelling and simulation is about obtaining information / data about how something will behave without actually testing the real prototype.

Figure 25 Logical Model of Requirement_06

From the model in Figure 25, the new requirement sentence(s) for (7) includes:

- A **Framework** shall *incorporate* **FoS_Mixes**, which are *abstracted into* a **Modelling** *environment* for the *intention* of **Simulation** whose results are *examined in* the **Analysis** with *criteria elaborated from* the **Mission Scenario**, which *forms the baseline* for **Modelling** task.
- A **Framework** *provides* the facility for **Assessments**, which are *established from* the **Mission Scenario** that signifies the *basis for* the **Analysis**, which *involves* the *trade-off* of **FoS Mixes**, whose *post mission* results are *compared in* a **Framework**.
- A **Framework** shall *incorporate* **FoS_Mixes** that are *integrated* within a *trade-off* the **Analysis**, which is used to *judge performance* of **Pilots** with *criteria elaborated from* the **Mission Scenario**, which is used to *evaluate* the **Assessment** *implemented* in a **Framework**.

For traceability and consistency, the requirement models already established from the previous requirement sentences have to be integrated into the model in Figure 25. As abstraction techniques is to be used to develop the FoS model, i.e. only model what is needed, only relevant characteristics of the systems within the mix will be included in the system models. The intention is to simulate the suitability of the systems for each student to execute the mission scenario, therefore, before simulation can occur a model needs to be produced. The simulations of the model, which describe the behaviour of the model dependent of attribute values, are used as measurement criteria for analysis: the criteria for the analysis need to be established from the planned mission scenario. The modelling philosophy has to be based on the analysis criteria established from the mission scenario, this philosophy gives

a clear indication of abstraction levels needed, in this instance mission phases e.g. (waypoint flight, task objectives, etc.). The mission scenario also provides the criteria that can be used to design and develop the behavioural aspects of the simulation. It becomes more prevalent in the requirement models that the intention of the system is to perform pre-flight assessments of the systems in the mix, trade-off the FoS mixes in the analysis (used to identify a suitable system with which to perform the mission scenario), then analyse the performance output, using the chosen blending mix, and compare performance in post mission conditions using a framework. It is also made clear in the model that the interpretation of ‘performance’ in the NLS has been identified as the relationship between the student pilot and the analysis tasks, i.e. the interest is in the pilot performance using a system of the FoS.

From Req_07, the requirement sentence can be revised to be more domain specific to read: Simulation of the FoS architecture will be used to assess the capability to support a coherent scheme of training for the aircrew under realistic operational conditions.

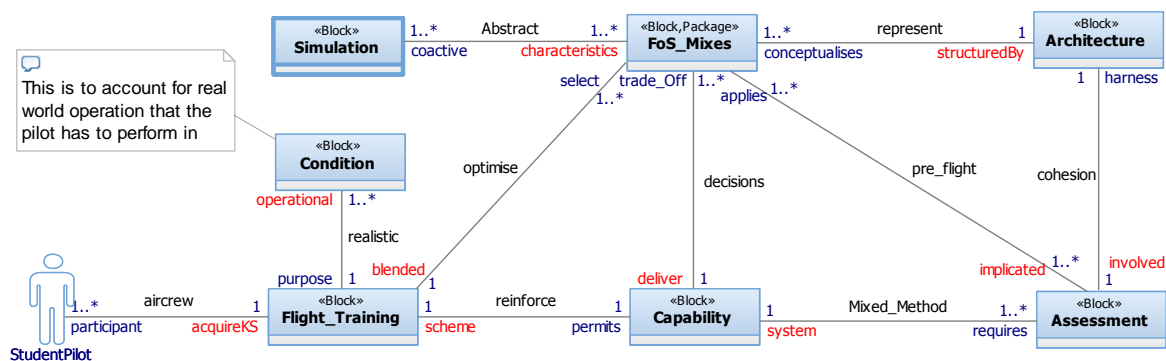


Figure 26 Logical Model of Requirement_07

From the model in Figure 26, the new requirement sentence(s) for (5) include:

- **Pilots** *acquires K&S* with **Flight_Training**, that is executed under *realistic operational Conditions*.
- **Simulation** involves *abstract characteristics* of the **FoS Mixes** that are *structured by* the **Architecture**, which is *involved* in the **Assessments** of the *system Capability* to *reinforce a scheme* of **Flight Training** for *participating Student Pilots*.
- **Simulation** involves *abstract characteristics* of the **FoS Mixes** to *optimise blended Flight Training* and *deliver Capability* that *requires mixed methods Assessments*.

The model signifies that the flight training operation has to be directly associated to real world conditions. The model also gives a strong indication that flight training operations has

The requirements models describe that one or more frameworks are going to be included within the DSS to structure the information gathered from the FoS characteristics to assist in the analysis tasks. The basis for governing the modelling tasks encompasses formalising the decision making process to bring precision to the solution of the flight training problem. The figure also illustrates that live and synthetic mixes are a part of the FoS mixes, however, the specific details of what systems constitute these mixes is absent. As part of the MDA

paradigm, the model includes a stereotype ‘within’ dependency to a FoS_Systems package which is discussed in greater detail in Chapter 6.3.

Holistically, recognition of a greater degree of complexity has been added from the analysis of the NLSs identified from the stakeholder requirements to include engineering characteristics. However, historically this type of transformation of the requirements has been absent for systems design and is only present in the model where substantial work has already been completed. The logical modelling task adds precision to the NLSs and gives a strong indication of the interpretation of certain ambiguities that occur using a rich language with which to specify requirements. To permit for execution of the developed model, the relations, interactions and qualifier have to be strictly diverse but retain semantic and syntactic meaning and precision. As the logical modelling of sentences progresses, relations, qualifiers and interactions between the static structures (object blocks) certifies whether there are any contradictory requirement arguments that need further clarification from the stakeholders. Additionally, iteration (and recursion) of the logical modelling task can identify any missing relations and objects that was previously absent from arguments and thus further analysis can check whether the meaning of the sentence has changed. From the vague requirements given for the project, three logical models of the requirements are found to be absent of identified relations and realised in the following figures.

In Figure 28, a relation between the ‘Capability’ and ‘Flight_Training’ block has been identified as absent in the original model (re: Figure 21). Additional meaning has been found to advance the operation of the ‘Capability’ block: A framework should exploit capability to reinforce a scheme of flight training. However, no contradiction in the meaning of the logical model is found.

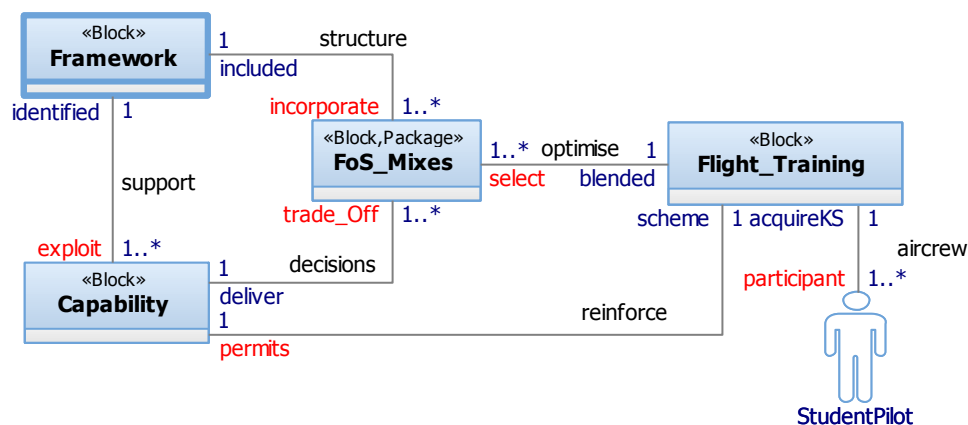


Figure 28 Amended Logical Model of Requirement_02

In Figure 29, the missing relation is identified between the ‘FoS_Mixes’ and ‘Assessment’ block (re: Figure 23). This describes that the FoS systems are to be assessed pre-flight within the framework. This clearly gives additional meaning to the original NLS in recognition that one or more assessments of the FoS mixes are required to be actioned using the framework in addition to assessing the cohesion of the FoS architecture

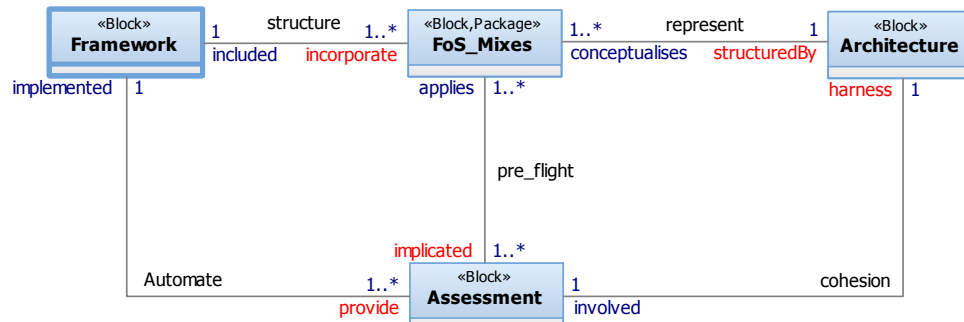


Figure 29 Amended Logical Model of Requirement_04

Additional relations have also been identified in the logical model of Figure 25, which is illustrated in Figure 30. The FoS mixes are involved in coactive simulation and is an integral part of the analysis function of the system. Again, this adds new significance to the sentence but no contradictory information and thus the meaning of the requirement model has evolved, but not altered.

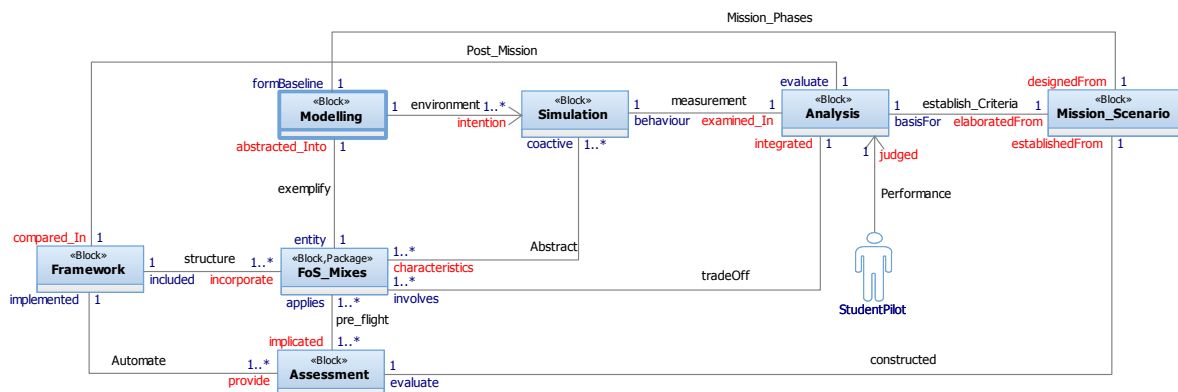


Figure 30 Amended Logical Model of Requirement_06

The models can be used throughout the design process to check that focus hasn't diverged and the abstract operational intent of the blocks and the interrelations between the block, identified in the logical models, are at the forefront of the design intentions. A subset of the identified blocks can be used directly within the system design models with others being elaborated to specific blocks with specific operations, e.g. the 'Simulation' block describes that simulation of the models is required; the 'Analysis' block encompasses a number of FoS systems with different system objectives and goals assigned to them; whereas the 'FoS mixes'

block is by definition a system design block. The general context of logical modelling of NLS requirements is to gain an understanding on what the objectives of the system is and identify key actions required for V&V activities to satisfy the original customer requirements.

6.3 Flight Training Enterprise

With the requirements for the project modelled and agreed upon from the industrial sponsor, further investigation of what constitutes the FTS was undertaken to gain knowledge of the systems of interest (SoI) to be investigated with the research.

The flight training enterprise considers the technology involved in training but also all the other related entities and participants within it. The training system can be seen as an integration of disparate systems into a seamless whole; the FoS mixes represents a real-time training system and the ground based systems administers their usage, organises their maintenance and communicates to other enterprise systems on the training base. Hence, the FTS comprises connectivity from the FoS systems to the enterprise systems (infrastructure services), as seen in Figure 31.

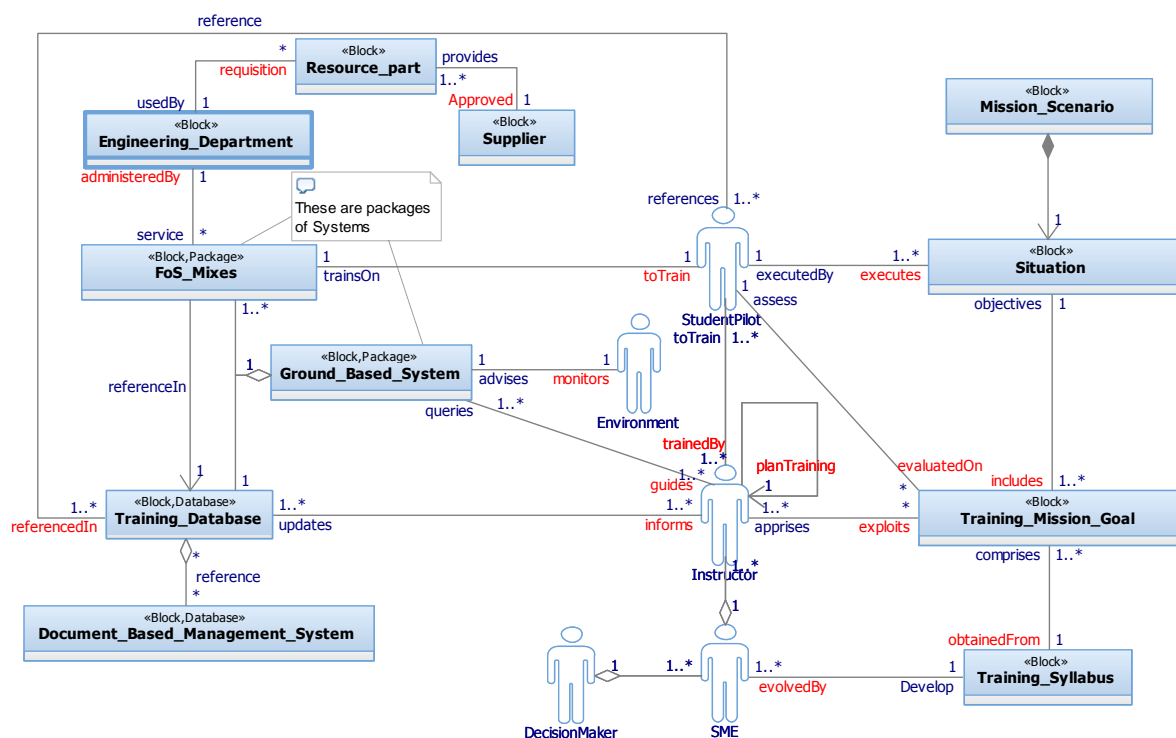


Figure 31 Flight Training Enterprise System Metamodel

From the metamodel, the decision maker 'actor' has been abstracted further to obtain the SME and instructor 'actors' to identify who are involved in training syllabus design and who

will directly interact with the student pilots and the training databases. Argument statements are produced that give clarification to areas of interest within the design process. Arguments gathered from the metamodel include:

- **Instructors** are assigned to train **student pilots** or **student pilots** are trained by **instructors**
- A **Student pilot** executes a **situation(s)**, which is part of a **mission scenario**, to be evaluated on **Training mission goals**_or A **situation**, which is part of a **mission scenario**, is executed by a **student pilot**.
- A **situation** includes **training mission goals** obtained from the **training syllabus**, which is evolved by **SMEs**.
- Etc.

The metamodel has identified additional relationships not prescribed within the logical models, e.g. Student pilot trains on the FoS_Mixes, Instructor updates the training databases where FoS mixes are referenced in, etc. This metamodel can be used along with the logical sentences to bring precision to responsibilities and roles within a complex system and is used to avoid misinterpretation and misunderstandings. The model itself can form the basis for model abstraction in concurrence with decisions based on current understanding from workshops discussing the issue of the FTS.

The metamodel(s) advances the relation structure, from logical modelling, to include traceability involving roles of entities of the system described by Mosse (2002) & Kecea et. al. (2007), for the ability to holistically describe and embed how one object knows another using interaction semantics. The metamodel also provides a visual approach to identify the key system entities which need further investigation and with it the understanding of which entities require additional scrutiny and study to gain knowledge of possible complications to an efficient workflow process that will help manage the enterprise system: the indication is given to the number of relationships each object has to others. It is clear from this metamodel that the objects that need to be understood more are: the student pilot (5 relationships), the instructor (6 relationships), and the training mission goals (4 relationships); these areas identify possible discontinuity of information within the FTS enterprise. Creating a metamodel such as in Figure 31, to inform and gain knowledge of the domain area, no matter where or which entity block you start from is more an art than science.

From the metamodel and the decomposition of the FoS mixes (See Appendix G, Figure G 19), the MDA technique is used to allow modelling of the FTS by domain separation shown in Figure 32. The logical modelling of the requirement narrative forms the CIM of the domain model. Within the PIM domain the users of the system have attributes and behaviours that needs to be managed, they form the ‘actors’ package; the FoS mixes package include the training technology; the enterprise package concentrates on the training organisation objectives required from the FTS, included are the attributes that relate to a measure of readiness and what parameters are needed from decision makers to develop the training system to focus on efficient and effective training scenarios that concentrate on capability and ability of entities in the system.

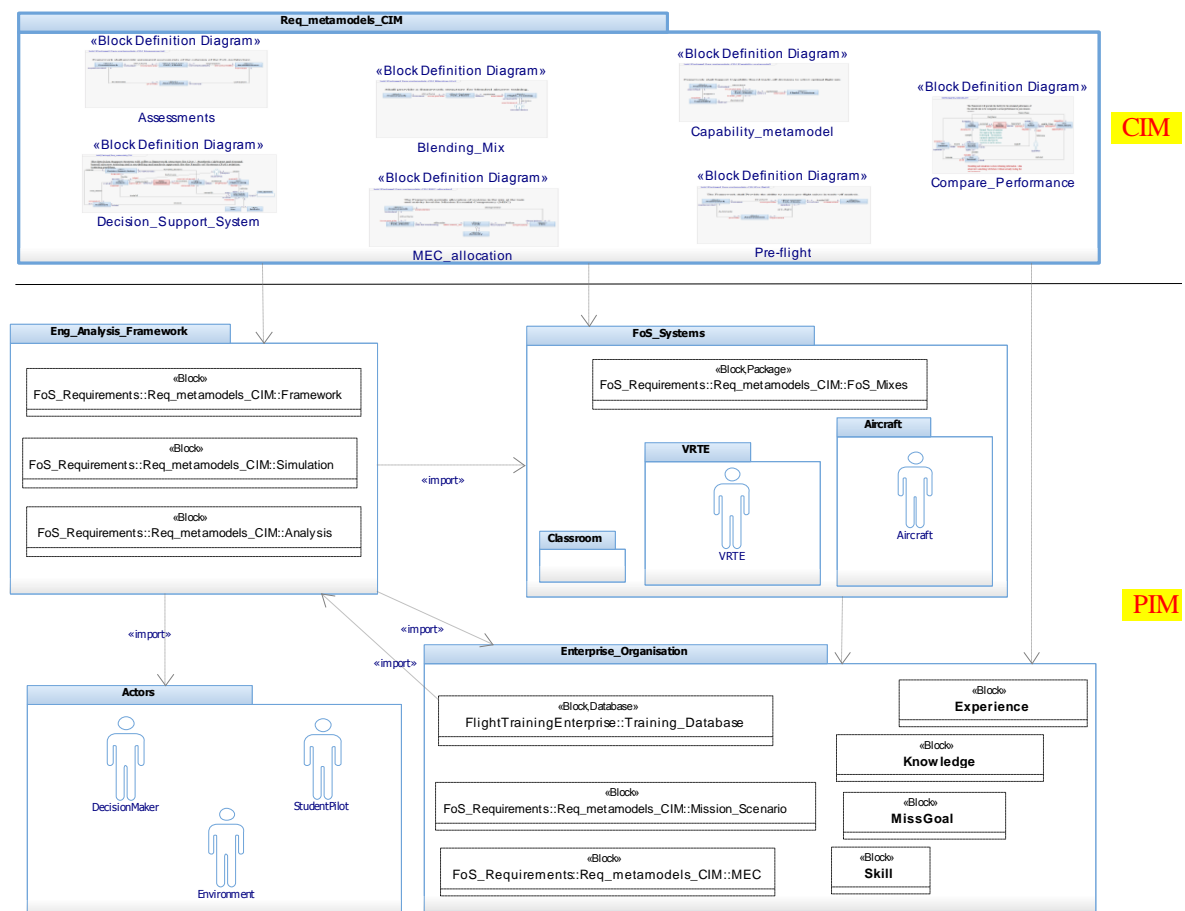


Figure 32 Domain Model using MDA Approach

The requirements narrative focusses on using ROSETTA framework(s) to be able to perform M&S for blended aircrew training for the pre-flight analysis of optimal flight mix. It is deemed that these requirements form part of the engineering analysis framework package, which integrates three system elements as part of the whole analysis, as each stage of the FTS workflow will require a distinctive type of analysis involving disparate data sets within

framework(s). The package needs information and governance from respective entities within the FTS, this data is required to be imported into the framework(s) either automatically from the FoS, databases or from the ‘actors’ of the system; furthermore, the framework is required to be able to modify the training database(s) once the analysis and the decision of which blending mix to use for training, is complete.

The whole PIM domain can be seen as the conceptual model for the DSS (see Figure 4) in the proposed methodology. Within the PIM model, entities from the requirements logical models are directly transformed to form blocks within the packages, this allows direct traceability from the PIM to the requirements narrative and the FTS metamodel. The packages are used for discussion on the level of abstraction needed as the model progresses into further detail for the identification of attributes and behaviours required by both the actors and the blocks of the system.

6.4 Use Case Development

“Use cases are requirements, primarily functional or behavioural requirements that indicate what the system will do. A related viewpoint is that a use case defines a contract of how a system will behave.” - (Cockburn, 2001)

Use cases describe the behavioural view of the system and model the functional requirements; however, they do not capture the entire set of requirements e.g. the non-functional requirements. Nevertheless, some of the non-behavioural aspects manifest themselves in the context of the use cases. The basic method of architecting use case diagrams involve: 1) finding actors; 2) finding use cases; and 3) describing the use case; but in large complex systems this simplistic method has many problems related to size (Armour, 2011). Modelling large systems with use cases involves the use of abstraction techniques to avoid too many use cases within one diagram (Lilly, 2000); hence, abstraction is used to ensure a minimal model of the system’s functional requirements (Adolph, 2003). The correct design process should identify possible gaps in the requirement elicitation process and allow for iteration and recursion back to the use case model for further evolution, which should also be signed-off by stakeholders before returning to the design stage.

For complex systems, the use case model commences with defining the system constraints (that can be updated as further understanding of the system is gained), which will give an early indication of the system and will capture the essence of the requirements that can be

communicated back to stakeholders to clarify their real needs (Probasco, 2000). The system constraints should answer the following issues:

- What is the problem we are trying to solve?
- Who are the actors who have vested interest in the system?
- What are the goals and objectives of the system and the organisation?
- What are the major functional and non-functional requirements?

The next stage is developing a domain model, using MDA techniques, to further define the system constraints and develop a set of common vocabulary for all stakeholders to use (Rosenberg, 2001; Magableh, 2012). This domain model provides a static view of a real world ‘thing’ and conceptualises the problem of the system indicating what problems to solve and what characteristics the system must know about (Eriksson, 2005). The process of requirement gathering for both use cases and the domain model occurs iteratively and recursively as the model progresses through the CIM, PIM and the PSM, as each revisit increases precision and accuracy of the Use Case set. Consequently, concentration is given to the identification of actors of the system, which can give further indication of missing use cases. In the context of larger systems, such as the FTS, the goal is to identify the major actors of the system and not to identify all the actors of the system as this could generate an exhausted list. The logical modelling of requirements and the system metamodel, for example Figure 31, can assist in identification of all key ‘actors’ to be present within the use case modelling stage.

Following on from the domain model is to define the use cases; the use case name reflects interactions of an actor with the system as precisely as possible (Gottesdiener, 2003). It is important that the name incurs a degree of technology neutral phrases or keep within a common ontology to de-clutter information that can obscure the real function and hinder communication between stakeholders.

The model has to have a facility to easily navigate through the model and to be able to understand what the model is portraying. Facilitating different diagrams to different aspects of the development or problem scenario can help understanding and identification of inconsistencies or overlaps (Gabb, 2001). The navigation should provide a story that can be unfolded by the means of decomposition through an hierarchical context to more detailed use case diagrams describing more functionality of the system using abstraction techniques as the basis for model development. Additional textual information can be beneficial to aid understanding, but as we are visual creatures, attaching either an activity or sequence diagram

to the use case not only aids in the verification process, via simulation, but also aids in describing events (with all variations) which explain the use case functionality in a simple way that will be difficult and complex to follow via textual descriptions.

6.4.1 Use Cases for the Decision Support System

The system design process for functional requirement elicitation commences with identification of the main stakeholders in the system followed by the main functional objectives that the system has to satisfy. Figure 33, describe the high level use case diagram of the DSS. The decision maker includes those working in various planning, project and management departments; the 'actor' includes a diverse range of subject matter experts (SMEs), and part of the SME 'actor' are the instructors who are on the front line of training and who are directly responsible for training efficacy. Furthermore, 'actors' who impact the functional objects of the system are the technologies or systems involved in the FTS along with the environment; with the student pilot 'actor' being a fundamental participant within the FTS. A number of assessments are required to be implemented to evaluate characteristics of suitability for filtering of participants before they are registered as a student for pilot training. Part of the FTS is to continually evolve the training mission scenario for inclusion of new technologies into the mix and assist in improving ToT. The DSS main objective is to assist the decision maker in providing blended mix training solutions to achieve training effectiveness, as per discussion in Chapter 4.2; thus 'choose blending mix' Use Case concludes the high level use case diagram that describe the main functionality the system must possess.

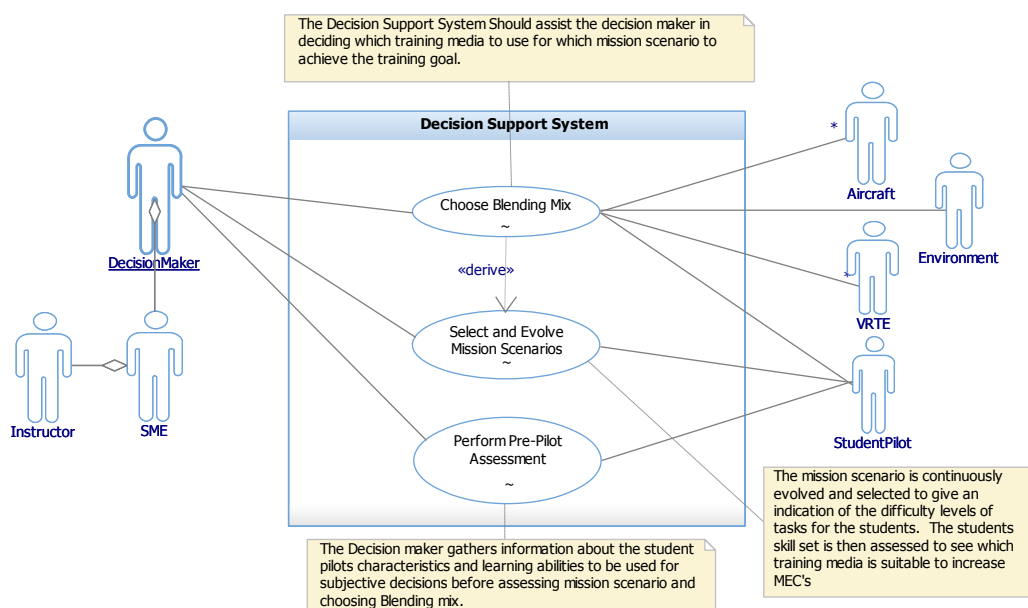


Figure 33 The DSS High Level Use Case Diagram

The high level use case diagram is the first stage of communication back to the stakeholder to clarify understanding of the main functional requirements of the system. Each use case is required to be abstracted further to obtain ‘lower-level’ functionality that the system must execute. Each use case illustrated in Figure 33, is further developed until the full functionality of the system is understood. This type of abstraction technique involves a number of iterations and recursion exercises; further decomposition of the use cases are described in Volume II, Appendix G: all use case diagrams are for final solution of the problem after a number of recursions back to the requirements derivation stage.

6.5 The Readiness Metamodel

‘What struck me was the realization that various government agencies and other organisations place a great deal of emphasis on ‘preparedness’ by collecting stuff, but comparatively little on actually knowing how to use that stuff. - We need to emphasis proficiency as a vital component of preparedness’

- Jim Milner (Ferry County Public Hospital Trust)

The suitability of the candidate training systems are surveyed to determine their capabilities in terms of the same characteristics used to survey training tasks. The results obtained should not be mathematically manipulated and evaluation of the training system should focus on 1) functional deficiency, 2) training task requirements, and 3) can the training technology be modified to resolve training deficiencies. The modelling of training options requires assignment of each training task to a candidate blending mix and effectively building an evolving model of a future training programme. Learning tasks can evolve the training system and incur different requirements on a simulation-based training solution (Lateef, 2010). The training tasks build from simple to the complex and from application of K&S from the normal to the novel as tasks become more difficult. Task analysis of goals and objectives from mission scenarios should examine the required fidelity requirements and the types of instructional support required for practice and evaluation of task performance (generally, the main fidelity characteristics considered are physical, visual and functional fidelity). One way of alleviating concern over the lack of familiarity with simulation is to conduct a comparison analysis of the VRTE used for training and a real live mission system (i.e. the aircraft); this can assist in avoiding conjectures about what a simulation is and is not suitable for (ICAO, 2014). The reasons for unsuitability can then be revealed through the analysis process and the reasons can be formally documented for future use.

Using the requirement models in Chapter 6.2.1, additional knowledge is needed to gain specifics of the information required to be within the DSS to assist the decision maker that will map into integrated flight training management system, as per Figure 12 in Chapter 4.2.1. Of concern for decision makers is to find the right balance between whole and part training using blended training methods with appropriate technology, budget, and time to ensure integrity of training is maintained and carry a high success rate for execution. The ability to direct training to the relevant K&S and pilots own attributes enables learning to proceed efficiently. Competency based training places emphasis on the required proficiency rather than the number of hours training or the number of missions that has been performed. Proficiency can be defined as satisfying a particular standard of performance, for example *“is the pilot capable of performing a manoeuvre and has the skill to apply appropriate control inputs at the appropriate time”*; *“Pilot understands the limitations of the technology regarding the current manoeuvre”*; *“the pilot uses the correct and logical reasoning to put the whole picture together”*. Often, the competencies are defined by SME for a given scenario or job description within a competency framework and are related to the goals and strategies of an organisation or system (IATA, 2015).

Competence is directly related to knowledge and experience gained through training; competence cannot be separated from performance they are uniquely interrelated. Generally, the identification of MECs begin with the analysis of the mission scenarios within a combat environment and are intended to describe training conditions for combat mission readiness (i.e. the MEC approach begins with the mission as performed within the combat environment). MECs such as Detect, Target, Engage, etc., give rise to functions that have a sequential nature from one mission to another, although some MECs are continuous through the training pipeline but vary in difficulty and are scaled to match expected outputs from the student pilot conducting the mission scenario. Colgrove and Alliger (2002) perceive MECs as being *‘Higher-order individual, team and inter-team competencies that a fully prepared pilot, crew or flight requires for successful mission completion under adverse conditions in a non-permissive environment’*.

Furthermore, MECs are divided into specific mission scenarios, which are highly correlated to the type of training technology; and are then decomposed determined by the phase of the mission. There is a consensus that a blended training mix can reduce the distance between continuation training and the proficiencies required in combat, and should be defined in the

MEC. At a lower level of definition are additional supporting competencies such as K&S, are interpreted within the common ontology of the domain to give a common understanding among all actors of the system. The MEC process is intended to fill the gap between K&S and actual task performance (Alliger et. al., 2012).

Knowledge is defined as “Information or facts that can be accessed quickly under stress” and is applied directly to the performance of a task of function; and a skill can be defined as “a compiled sequence of actions that can be carried out successfully under stress” and is a observable competence to perform a learned psychomotor act (a psychomotor act concerns the relationship between cognitive function and physical movement). There are three general types of experiences which go in hand with attaining K&S, which are also identified by SMEs (Schutte et. al., 2012), these are:

- An event that occurs to or a situation encountered by the student pilot – examples include: flying over difficult terrain, fatigue/time on task, operating restrictions, etc.
- An action performed by the student pilot – examples include: live weapons deployment, locating targets, defeating enemy radar, etc.
- An operation for the student pilot that may be helpful to gain the competence required for successful mission completion – examples include: dynamic re-tasking operations, operations against air or ground adversary jamming, etc.

Most MECs use the basis of experience lists to develop and evolve the training programme and enhance the mission scenarios. The key to satisfying the identified MECs is the ability of the pilot to gain experience from the training mission tasks. An experience can be defined as ‘a development event that occurs during training and at various times across the career that facilitates learning a K&S, practicing a MEC or supporting competencies under operational conditions. An experience is an identifiable event that is the facilitator of mission readiness, i.e. being fully prepared for something by consistent high performance. The training given is to bring or maintain pilot’s performance to an agreed standard of competence. Consider the task of piloting, the pilot must simultaneously perform knowledge-based activities such as planning and navigation, make rule-based decisions such as flight path to next waypoint, and exert skill-based control to orient the aircraft. Flaws in any of these mechanisms will degrade overall performance; these flaws result from improper strategies acquired from training and subsequently incorporated into skilled performance. Strategies can be identified by evaluating qualitative data of the control situation and the resultant pilot actions. Partitioning

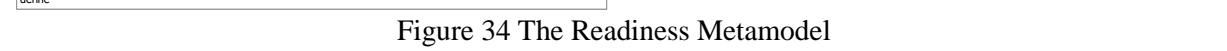


Figure 34 The Readiness Metamodel

6.6 FTS Process Workflow Design for Methodology

For efficient training practices, the specific task-relevant knowledge, skills and experience level of the student are key factors to consider for the successful execution of the training mission. The key component is which blending mix solution to use for the current level of K&S required by the mission and students current ability levels. The task-specific metric considers an individual's current experience for performing a planned task to give a measurement of suitability for completion within specified goal margins. The metric should be based on the current level of experience before executing the task and then reassessed upon completion of the planned task. The metric is a subjective measurement which represents the individuals educed experience of encountering similar tasks in previous flights – a higher metric (rating) is an indicator of the likelihood of success with the planned task.

The performance results from the technology can be used to provide feedback to identify potential problems arising from knowledge or blending mix choice made in the trade-off analysis from the gaps within the management processes of the FTS. A student pilot's own opinion regarding K&S adequacy for their planned task will influence task performance rating. This type of measurement process can be used and applied at any level of task within the training scenario and can give an indication of human and technology compatibility to assess critical task and training performance factors from a human systems perspective.

The subjective personal judgements, elicited by questionnaires, from the actors of the FTS provide the opportunity to quantify the level of K&S and technology suitability, against weighted definable criteria representing the expected grade from the training mission. Subjective assessments rely on perception and opinions of participating actors, thus, motivation and attitudes towards work processes can directly affect the weighted performance factors, consequently, the DSS attempts to account for these complex human factor related measures. The use of subjective measures permits experts opinions to directly affect decisions or complex processes that are too challenging for other techniques to accurately quantify. An overview of the proposed FTS Workflow Process used in the methodology can be seen in Figure 35, which is further detailed throughout Chapter 7.

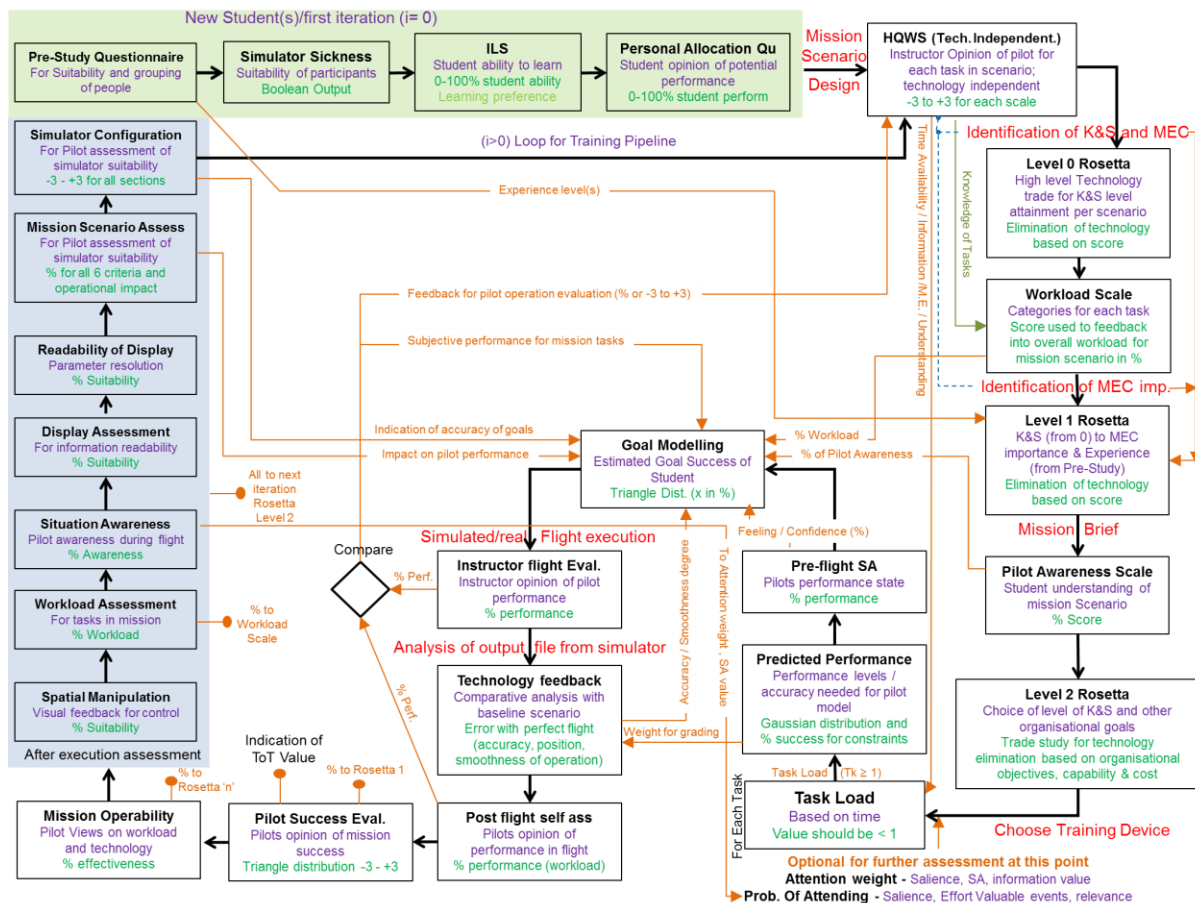


Figure 35 FTS Workflow Process

Each stage of the workflow, indicated by the blocks, can be completed at different times allowing for operational flexibility of the decision maker and allow the same assessment stage to be complete for each pilot before progressing through other stages of the workflow. At the commencement of the process, the decision maker will be asked which stage in the workflow he/she would like to continue the assessment, as per Figure 36. The Decision maker will use the scroll buttons to select the appropriate stage and accept using the ‘Assessment Identified’ control, as per the image on the right of Figure 36, the workflow tool transfers to the assessment stage selected.

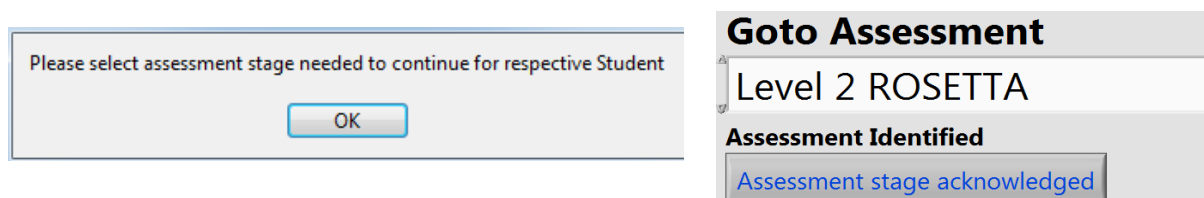


Figure 36 Workflow Process Stage Option

After each completed stage, the decision maker will be prompted whether or not to continue through the workflow, as per Figure 37. The option of ‘Yes’, transfers the assessment criteria

through the process indicated by the 'black' flows in Figure 35, in a natural manner, however, if 'No' is selected the workflow tool exits and the option in Figure 36 is again presented.

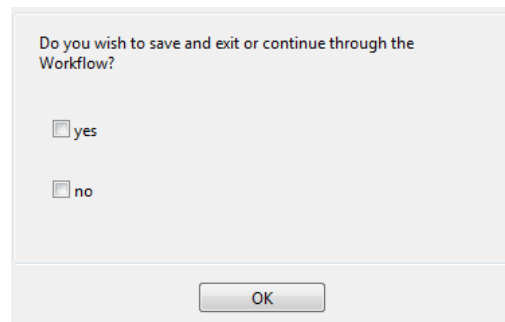


Figure 37 Continuing the Workflow Process Prompt

The workflow process provides the opportunity to integrate subjective, objective and formal assessments to be present in one workflow tool using discrete sets of text files to store the information. The relevant feedback and assessment techniques are needed to assist in evaluating the training exercise and assess suitability of the available blending mix. Performance measures based on the assessment criteria for the stages in the mission scenario and subjective assessments with respect to usefulness of technology, workload and training goals provide statistical data to analyse the success of mission goals. The feedback of performance and subjective data (Magnusson & Berggren, 2002) both from the student pilots and the instructor(s) add to the validity/legitimacy and can give an indication of when additional training or a greater level of difficulty to attain greater levels of relevant K&S and gain further experience is needed. The data gathered from the feedback, indicated by the 'red' arrows in Figure 35, can be used to adapt or modify the training mission scenarios to suit the relevant stages of the student pilot and the comfort level of using a particular training technology (blending mix). Performance feedback to individuals, as highlighted by Fletcher (2000), improves performance by supplying performance data in terms of effectiveness. This leads to a psychological prerequisite for learning that permits individuals to acquire knowledge and confidence on recently exercised tasks, which inherently provides motivation as individuals desire to know the efficacy of their actions and behaviour (Unsworth and West, 2000).

The diverse number of human performance measurements approaches along with technology assessments from the student pilots and decision makers, whilst individually has shortcomings and methodological problems, in combination they can if administered correctly provide a robust and reliable representation of human performance with the respective training technology (Bashir, 2013). It is to be noted that human performance is a

significant parameter to measure for the FTS; the entire process, both human and technology, leading up to the end state of pilot readiness is important to study as a whole.

6.7 Chapter Summary

'The winner (of an air battle) will have been determined by the amount of time, energy, thought and training an individual has previously accomplished in an effort to increase his ability as a fighter pilot' - Randy Cunningham, USN

The objective for the FTS is to reduce the gap between live flight training and that required on the front line by determining the optimal mix of live and virtual sorties. MEC's offer a highly specialised 'blended' task analysis approach that could assist in governing the training regime of pilots based on capability and competence. MECs provide a framework to analyse mission requirements for the design of blending mixes that maximises learning and K&S development, ranging from individual techniques and procedures to complex mission tasks (Bennet et. al., 2007). The move to a blended mix solution requires a level of precision in mission scenario design and execution, and the MEC approach is ideally suited to provide the measurement of both MoP and MoE within the analytical process. The MEC approach supports capability to quantify mission parameters and allow multiple mission scenarios and metrics to be created to support realistic blended training evaluations. The result is the training event remains limited in relation to the most realistic representation of the conditions of performance, therefore, pilot performance ultimately must be measured in relation to the combat environment and as to which it must be trained; hence, be the 'baseline' for the requirement specification.

The ability to assess the effectiveness of the training programme is to provide a robust 'two-way' system which includes subjective data and pilot performance on a comparable 'benchmark' mission scenarios presented throughout the training pipeline. The performance for each individual student using the chosen blending mix will then be compared in addition to instructor and self-assessment criteria feedback. The questionnaires used are developed to gain a subjective viewpoint on how well the training exercise's goals were completed, which are then converted to objective values relating to 'feeling'. The technology feedback in relation to the planned mission is partitioned into relevant stages and each stage has a K&S requirement, which includes the expected attainment levels (estimated from current information the instructor has on the student's abilities) for the current training mission using the blending mixes.

When a system is being initially developed, the best method is to use deductive reasoning to separate intuition and experience, using the bare minimum of reference to the simplest of intuitive concepts; it thereby gains a maximum of general applicability. The minimum number of truths (or assumptions) that are used the greater the possible application of the completed system to the further community. A method of abstraction is necessary to ensure generality that helps forms the basis of axiomatic design theory (Suh, 1998) and permit the design to relate to more areas paradoxically enough to detach the actual system from the underlying experimental knowledge as much as possible. Due to the nature of operational processes within the FTS, the system models are an abstraction of a mathematical process to reduce the potential of bias and to gain further information regarding actors within the system that contribute to variance in performance outcomes or effectiveness. The level of abstraction used for the system models constrain the level of complexity by representing those factors/parameters which are associated with assisting in the decision making process. The system model considers feedback from the decision maker, student pilot and the performance feedback from the chosen blending mix. The interaction between human-and-technology is considered as a prime factor in the success of optimising blending mix solutions. The complex nature of human factors involved will affect the collaborative functioning of working practices, levels of cohesion and possible conflict. The feedback is inherently used to modify the state of the socio-technology system within the FTS and over a long period of time will exercise an influence upon operational performance; this type of process is sometimes referred to as organisational learning and represents shared mental models and evolution of learning (Mullen & Copper, 1994); (Chawla & Joshi, 2011).

ARCHITECTURE APPROACH TO PROCESS WORKFLOW

“An essential point is that it’s a fundamental mistake (so far, endless repeated) to believe there is some special tool or technique in software that will make a dramatic order-of-magnitude difference in productivity, defect reduction, reliability, or simplicity. And tools don’t compensate for design ignorance.”

- Mythical Man month, Fredrick Brooks

Outline of Chapter

This chapter describes the design of the methods used for the methodology along with arguments for the feasibility of using the ROSETTA frameworks for decision support. The chapter commences with a brief introduction that includes an abstract architecture to simplify the workflow process tool to a simple conceptual model relating to the decisions and attributes of concern within the ROSETTA frameworks. The full system architecture for all the stages of the workflow process is discussed in detail along with aspects of the outputs from each stage and how the solutions relate to the traceability of blending mix choice to the success of the organisation and training requirements. Each sub-chapter gives an explanation of the various assessment criteria involved and associates the design to the main literature review in the early chapters; this is followed brief summary of findings from the verification of the various algorithms and GUI information used to develop the workflow process tool.

Abstract Conceptual Model of FTS Workflow Process

Figure 38, which is an abstract version of Figure 35, describes the approach used to develop the DSS tool and the evaluation method. First, the training scenarios are developed from the identified training attributes for the student pilot to obtain. The mission scenarios are designed to establish a sequence of information requirements for the system and conversely collect pre-mission data on perceived SA and workload. The order of categories of workload are ranked and compared to the importance level of the tasks within the mission scenario; a holistic evaluation of workload for the mission is then accomplished. The mission parameters are then used to decide on functional prerequisites required from the blending mix for the level of K&S. The scenarios can be used to gather functional behaviour to analyse and develop a task list to map into the requirements for the appropriate blending mix to use for the scenario and evolve the FTS model.

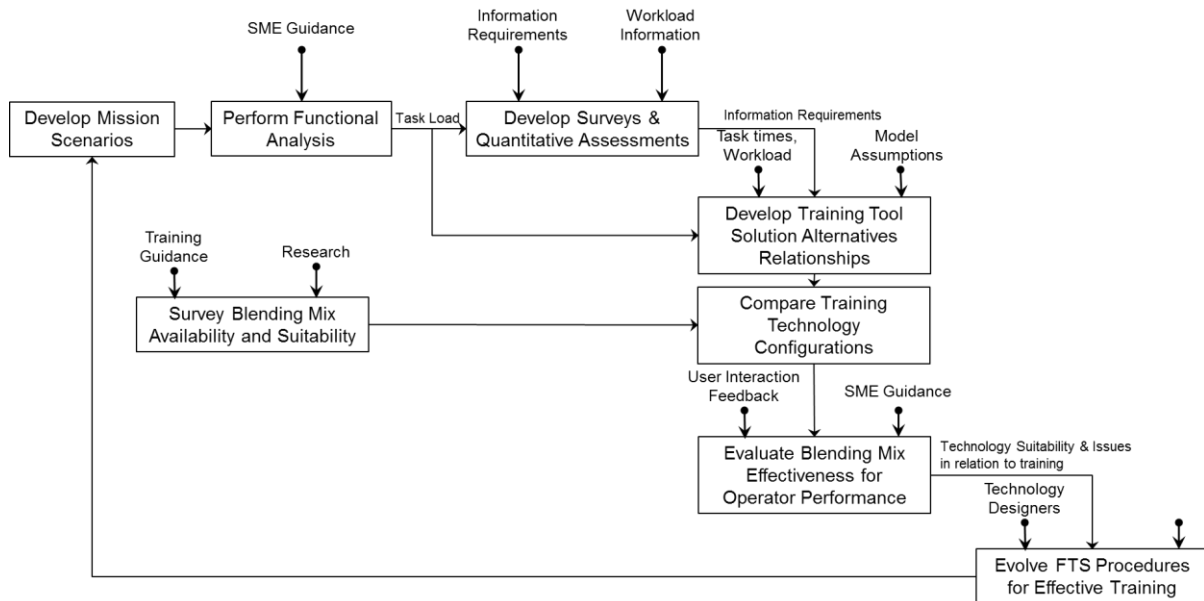


Figure 38 Simplified Architectural Approach

Knowledge regarding the importance and the understanding of the interaction of performance determinants is compiled to give the underpinning performance factors for inclusion within the conceptual model. The designed conceptual model is a practical consideration of the measurability of relevant performance metrics for use within the trade-off analysis defined within the framework(s). The defined relationships between actual measured variables and the supporting constructs concerning relevant attributes are based on the difference between task-oriented and personal relations-oriented attributes. However, one of the most difficult consideration for complex socio-technological systems modelling aspects are cultural aspects of the individual to intangible factors that are difficult to specify objective criteria. Questionnaires are to diagnose cultural aspects based on observable indicators (Grote and Kunzler, 2000) or quantifiable instances of behaviour (Tsui et. al., 2007).

Actual performance dimensions are represented by the relationships between the blocks of attributes that create the ROSETTA frameworks and encompass organisation goals, effectiveness (student and technology), and efficiency (student performance through the training pipeline). Performance goals can be considered as properties of the capability of technology and interactions with it. The clear consideration is to draw upon the conceptual distinction between performance metrics and the outcome of the behaviour of the student pilots to a mission scenario with a specific technology configuration. The framework(s) should distinguishes between the relationship of performance factors and organisational performance; the resulting indicators can be used to determine how effectively a specific

activity (mission scenario with specific technology configuration) produces efficient outcomes for the objectives of the organisation.

However, one of the main difficulties using workflows and the ROSETTA frameworks is the management of large data sets. It is generally custom to use excel style spreadsheet databases i.e. Doors, to hold enormous amount of data that a decision maker has the enviable task to search through, which is inherently time consuming and error prone (Madahar et. al., 2008). For systems, complex management of such processes needs to be avoided; database management for this project has been developed with its own architecture and is illustrated in Figure 39.

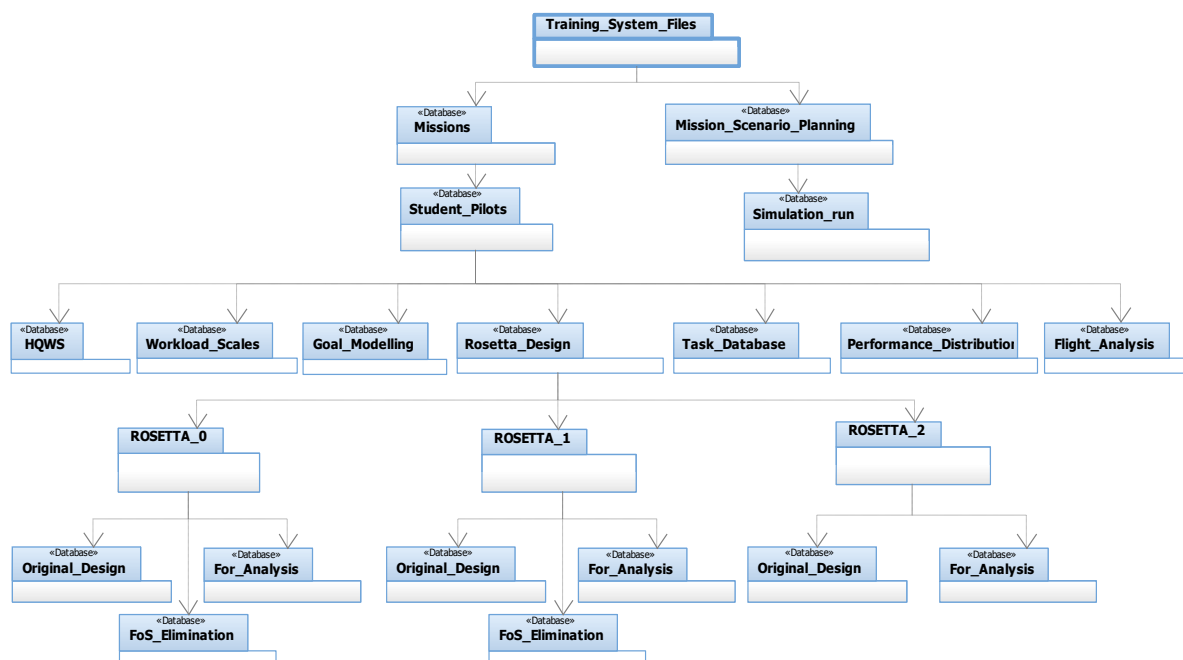


Figure 39 Database Architecture Format

The mission scenario planning is separated from the main database elements to provide central directories that is easy navigable for mission planning and evolution tasks. Each student pilot is given their own unique set of directories which correlate to the assessment and evaluation criteria. The main student files are stored in the parent directory ‘Student_Pilots’, which will provide a summary of the assessment results for each stage of the FTS described in Figure 35. The ROSETTA framework directories store both the design RSEs and the analysis RSEs arrays along with the analysis and evaluation data for technology elimination choice(s). Each database gives simple easy to understand information using mainly text files without the need for substantial searches to attain the information needed for varied decision makers to use and assess. The questionnaires for assessment are based on an

online system that the student pilot completes; the information can then be downloaded as a spreadsheet or textual file, which the workflow tool interrogates by the means of the mission reference and / or student pilot name. The objective sections of the questionnaires are used directly within the algorithms of the workflow tool, whilst, the subjective sections are used for further analysis on student pilots thoughts and feelings about performance aspects of the mission. The information presented to the decision maker by the workflow tool will highlight strengths and weaknesses of FoS mixes and students in a visual way.

The following sections in this chapter discuss methods used to produce the methodology to assist in providing optimised blended training; the examples given are for verification of the algorithms that describe the functionality of the tool. The verification process has been used to determine that the model accurately represents the conceptual description of the problem given in the requirements analysis phase. The scope of the conceptual model includes philosophy and general concepts for conducting assessments when uncertainty and lack of knowledge is obstructing understanding and efficient execution of management system. The concepts described can be used in more general terms than the FTS, and it is perceived that the conceptual model is for long term benefits of an organisation, as the model requires a number of iterations and recursions through the system being used in a practical environment to evolve and optimise the mathematical functions (shapes) that describe the relationships.

7.1 Mission Planning and Baseline Simulation Model(s) (General Overview)

Humans react better when instructions are given in a logical sequence format; information that is scattered and does not integrate in sequence can lead to confusion and error. Within an operational environment this disconnected knowledge can lead to a mental model that will overlook or misinterpret information being presented at a moment in time (Buchelder & Levenson, 2001);(CAA, 2004), thus, the mission scenario should be developed for ease of understanding of task / mission goals and objectives.

7.1.1 Architecting Flight Training Mission Scenarios

Events can be caused by interactions with the environment (users, sensors, switches, etc.) causing a change in state variables and thus change in system state (Burns, 2014), i.e. requires the pilot to action something in the cockpit of the chosen blending mix to change the state of the aircraft. The mission scenario is considered to be constructed out of a sequence of tasks connected in series to form a pipeline, which describes a mission scenario and within the tasks a set of activities that can have difficulty scales associated with them (IATA, 2013). Each task output in the mission scenario is coupled to the input of another and so on; this is identified as a control coupling (CC) and as such identifies a relationship between tasks, illustrated in Figure 40. Since this model is part of an overall training programme, it has an external input coupling (EC_{in}) and an external output coupling (EC_{out}) behaving as interfaces, which receives its value from summation of all the tasks to the last task.

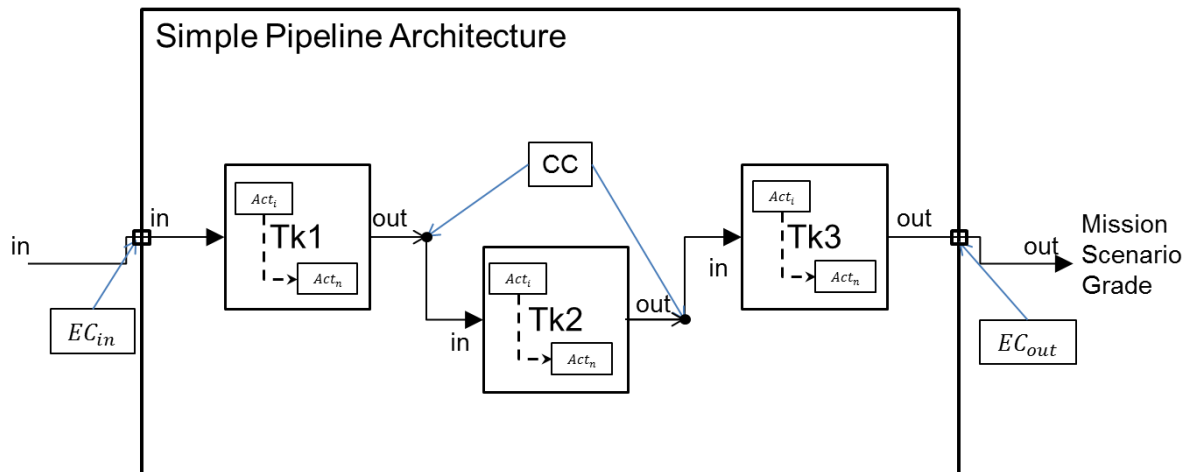


Figure 40 Pipeline Model for Mission Scenario

The specification for the pipeline mission scenario model is:

$$N = (X, Y, D, \{M_d | d \in D\}, EC_{in}, EC_{out}, M_{CC} | (Act_i, Act_n) \in CC)$$

Where,

$$\begin{aligned}
X_{in} &= V \text{ (an Arbitrary)} \\
X &= \{('in', v) | v \in V\} \\
Outports &= \{ 'in' \} \\
Y_{out} &= V \\
Y &= \{('out', v) | v \in V\} \\
D &= \{Tk1, Tk2, Tk3\} \\
M_{Tk3} &= M_{Tk2} = M_{Tk1} = M_{Task} \\
EC_{in} &= \{((N, 'in'), (Tk1, 'in'))\} \\
EC_{out} &= \{((Tk3, 'out'), (N, 'out'))\} \\
CC &= \{((Tk1, 'out'), (Tk2, 'in'), (Tk2, 'out'), (Tk3, 'in'))\}
\end{aligned}$$

Thus, the mission scenario commences with estimated performance tolerances based on precision and goal attainment (X_{in}) of using the blended mix, which are graded as the student pilot progresses through the tasks of the mission with the ability to assess each activity within the task (M_{CC}). The result of performance or goal success can then be summed leaving an overall grade for mission success (EC_{out}). As a result, the mission scenario can be planned and assessed within a series of slots within a ROSETTA frame with the relationship between parameters (tasks) placed in precedence order signifying when a change in state of the aircraft is required for mission and training objectives and to evaluate mission success (Holden and Dickerson, 2013). An external event, x_i (such as flight time, accuracy to waypoints, timings for mission, etc.), arriving at EC_{in} , the first event in the sequence is transferred to the input port 'in' of Tk1 (slot 1), the second to the input port of Tk2 (Slot 2), and so on. Once the training mission has commenced; the student pilot will begin to execute the sequence of events using the training technology/system. As the pilot finishes the first task 'Tk1' and all the actions associated with it, an output goal grade 'out' appears on the output port of Tk1 ('I' Frame) (based on the comparison with an ideal (baseline scenario) or experienced flight output for the task, within the transformational frame) and transferred to the 'in' of Tk2 because of the CC. This transfer process from output port to input port continues until the mission scenario is complete and the 'out' appears on the EC_{out} . The grade for each task is stored in a database and summed for the overall mission accomplishment. This will create a generalised metric grade for progress through iterations of the mission scenario when practiced through a training system.

7.1.2 Training Attributes for Mission Planning Tasks

The attributes of flight relevant to monitoring is stored in a textual data or spreadsheet file (re: Figure 2), which are used within parametric simulations to determine the performance on the

current mission scenario. Of key consideration is the recognition that K&S capability levels of technology will vary, as such the mission scenarios need to be architected with the viewpoint that the framework(s) incurs the ability to be agile enough to recognise what the specific limitations are and what the simulation limitations are and therefore permit the analysis of the success of training to be adaptable for individual student pilots and configuration/types of training technology. The planning has to consider a number of attributes for flight training with a realistic emphasis on risk management skills, decision-making skills including planning and motor control skills. The scenarios require a structured script of real-world experiences to address flight training objectives in an operational environment.

At the heart of planning mission scenarios are MECs, which describe the training attributes requirements and thus places constraints of targeting specific training objectives (Castor et al., 2009). The process involves mapping of any supporting competencies as well as K&S relevant to the current mission scenario. (Volume II, Appendix K Table VI, describe some of the MECs and K&S lists which are under consideration for the methodology). To generate a quantitative link between the MEC and K&S set, the critical step is in the definition of the mission performance requirements, which drive training objectives; the construct of the relationship can be seen in Figure 41.

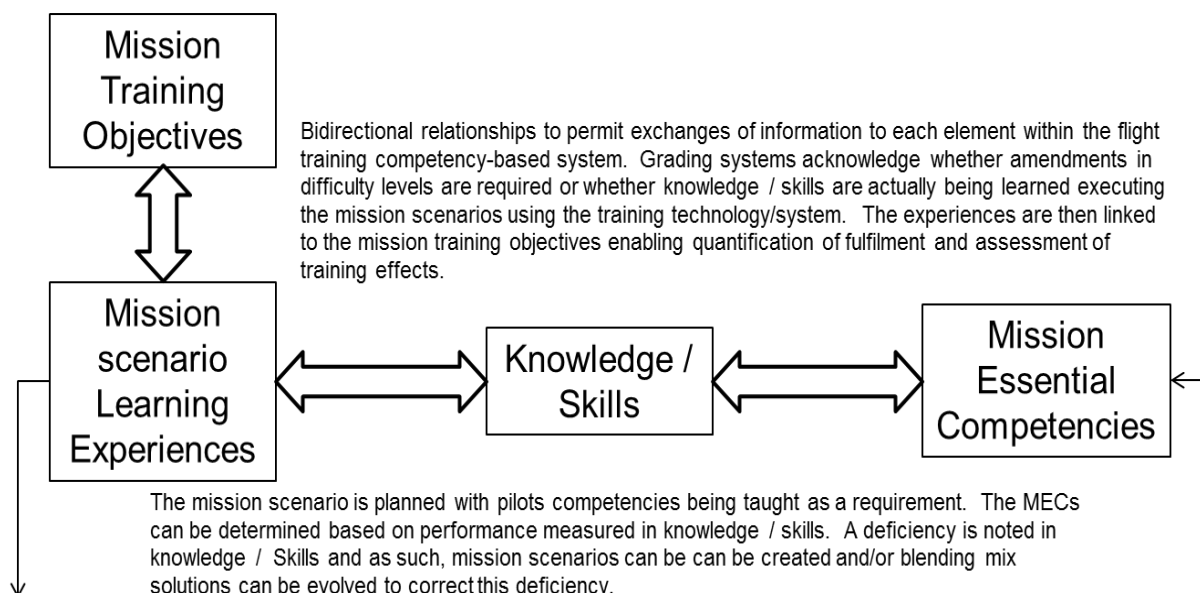


Figure 41 MEC Relationship to the Mission Scenario and K&S

The repetition or rehearsal of the mission scenarios will allow a generalised increase in difficulty until the required MoP has been achieved over a set training timeframe. The repetition is believed to facilitate in the development of skills over the same time frame.

When the instructor grades the success of the mission, a decision is then made whether the pilot advances to a more difficult scenario, is to repeat the same scenario (Denning, Bennet & Crane, 2002), or whether suitability of the blending mix needs investigating.

For detailed description and design of the proposed mission planning system, scenario design and simulation aspects of the workflow tool used to compare student performance with a baseline ‘ideal’, see Volume II, Appendix H, which describes in detail the mission planning and flight simulation behaviour that integrates with the proposed workflow process tool and associated frameworks described in the methodology. The approach to data collection is designed to both capture different types of information and provide a level of triangulation for similar types of data. The methodology proposes that student pilots should be asked to rate, prior to their flight, their perceptions of their capacity to perform the mission scenario using the chosen blending mix. Then, having completed the mission asked to rate the same dimensions in terms of their performance during the simulation/flight, as per Figure 35, to evolve the relationships in the ROSETTA frameworks and subjective judgments of the decision makers (instructors) on student pilots perceived experience and ability using the chosen blending mix for the planned mission scenario.

7.2 Pre-Pilot Assessment Implementation Design

The first assessment concerns personal characteristics, which include details on previous flight experience; the questionnaire can be viewed in Volume II, Appendix I Assessment Form 1. The objective data produced is used to clarify the ability of the pilot to use goal related computer games, produce an age rating grade along with a visual representation of age e.g.:

- A grade of 1 indicates the prospective pilot is over 35years, visual ID is ‘Blue’
- A grade of 2 indicates the prospective pilot is between 31-35years, visual ID is ‘Pink’
- A grade of 3 indicates the prospective pilot is between 26-30years, visual ID is ‘Yellow’
- A grade of 4 indicates the prospective pilot is between 18-25years, visual ID is ‘Green’

The visual representation for Gender is given by Boolean virtual LEDs, one for male, one for female. The objective questions i.e. have you or do you play computer games, are given the score associated with the answer options, thus question numbers: 6,8,9,10,11,13,15,17 are scored with the grade given. However, questions 14 (11) and 16(13) are scored as follows:

- A score of 1 indicates the prospective pilot has less than 1 years' experience, visual ID is 'Blue'
- A score of 2 indicates the prospective pilot has between 1-5years, visual ID is 'Pink'
- A score of 3 indicates the prospective pilot has between 6-10years, visual ID is 'Yellow'
- A score of 4 indicates the prospective pilot has between 10-40years, visual ID is 'Green'

The flight hours is given as an indicator of experience, furthermore, the attribute 'total score for simulator and real world Flight' is based on the average difference between simulator flight and real aircraft flight hours for the undergraduate naval pilot training programme (Teachout et. al., 2013), thus, each hour of simulated flight experience is worth ~45.5% of real live flight experience. All the attributes are stored in the main student file; the UI for the pre-study assessment is illustrated in Figure 42. The completed questionnaire is downloaded into the main database file, the instructor will search for a name from the database and the tool will use the information in the file to give clear indication to the decision maker of candidate characteristics, further information of the behaviour of this stage can be found in Volume II, Appendix I Figure I2. The subjective answers are stored in the database to be further interrogated if needed, in addition the two attributes relating to flight experience are updated automatically following relevant training exercises that can be used to assess levels of performance of student pilots.

Name to Search (SurnameINITIAL)	
trevorH	
Gender	Age Rating
Male <input checked="" type="radio"/>	1
Female <input type="radio"/>	flight hours
	45
Age Category	total score for simulator and real world flight
	52.3636
total score for questionnaire	
66	

Figure 42 Pre-Study Questionnaire Result Window

Pre-Screen Simulator Sickness Questionnaire

The lack and mismatch of physical movement to visual feedback available on most VRTEs cause's symptoms that is similar to motion sickness. It is imperative before acceptance of student pilots to these types of training tools that assessment of any possible health problems are evaluated. The symptoms have strong correlation to age, experience, gender, and postural stability. The questionnaire in Volume II, Appendix I Assessment Form 2 is completed by the prospective student pilot for evaluation. The tool interrogates the information from the

questionnaire to give the decision maker a quick identification of any potential issues as per Figure 43. If any of the questions are answered as 'No', green indicator(s) will illuminate within the 'Issues with Participant' cluster, signifying that further investigation is required before the participant can be accepted into the training system. However, if the 'No issues with participant' cluster illuminates all the green indicators, as above, it is deemed that the participant is suitable to train on VRTEs where symptoms have been known to occur. Please refer to Volume II Figure I 3 for further details of decision maker interaction with this assessment stage.

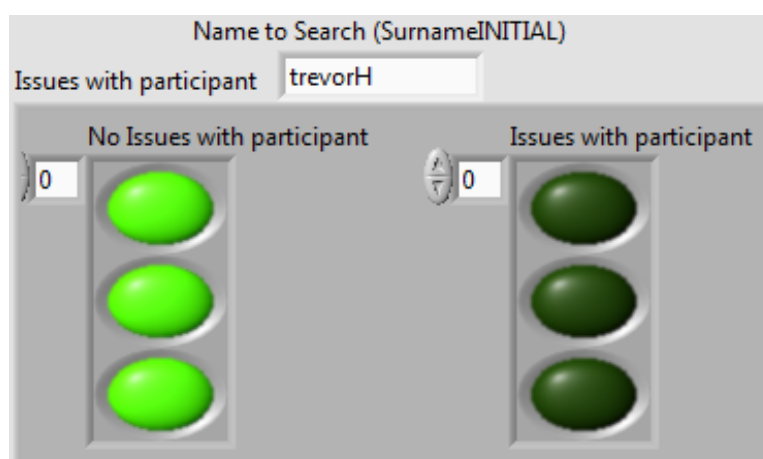


Figure 43 Indication of Simulator Sickness

7.2.1 Index of Learning Style (ILS) for FTS

‘Teaching and learning styles should become one of the greatest interests of the educators particularly their relationship. However, one of the weaknesses of the research into learning styles (LS) is the lack of the investigation into the matching of teaching and learning styles. Theoretically, many variables exist in the educational literature but few researchers dealt with the matching of teaching styles and learning styles’

- (Mulalic, p. 102, 2009)

Learning relates to ‘what’ and ‘why’ people learn (Clarke & Mayer, 2011); moreover the context and methods people learn material in order to achieve goals and objectives that build knowledge and skills (K&S) related to ‘job’ performance. Teaching, no matter what age group, has been executed using an instructional design based on ‘one size fits all’ approach, even though there is a consensus among experts that not all individuals learn through the same experiences or through the same methods of teaching (Stevenson & Dunn, 2001);(Smith, 2002). There is also evidence the different learning styles may help learners achieve better results (Bull & Ma, 2001);(Rayneri et. al., 2006). Bajraktarevic et. al. (2003) confirmed significantly better results in performance of students with providing adaptivity in

the learning material to match their preferred learning styles. Performance and K&S acquisition levels, given the same instructional process and learning material, should be established on information gathered from the suitability of learning styles for the student. It is then perceived that the gaining of K&S will allow the student(s) to follow the process of learning those K&S and adapt (to a degree) in learning style preference over a period of time (Tucker, 2007).

A learning style is defined as the strengths and characteristics in the way people receive and process information and refer to individuals own methods and strategies to learn and use the information (Felder & Silverman, 1988). The students recognise new knowledge, review previous concepts or experiences, reorganise their understanding to match the new one, assimilate and interpret the knowledge and attempt to action the knowledge at an appropriate time. The developed learning styles (Felder & Spurlin, 2005), incurs the benefit of not only giving an indication to the SME of how the student will handle the learning process, but in addition can assist students in fostering a sense on their own learning preferences (Suskie, 2002) and hence their own strengths and weaknesses (Bowles, 2004).

These learning styles relate to how individuals receive, perceive, process, and understand information. Consequently, the Index of Learning Styles (ILS) is chosen to gain insight into how the student pilot(s) will cope with the instructional design and highlight the potential issues with the execution of SoPs during mission scenario execution. The Validity or legitimacy and reliability of ILS has been investigated by Coffield (2004) and Hou (2010) when it relates to identifying individuals specific strengths and weaknesses on specific learning style dimensions. The subjective and intuitive awareness of learning style preferences gives a reflection of individual's diversity (Wagner, 2011).

7.2.2 Knowledge of Strengths and Weaknesses Of Students

Evaluation strategies when it related to the relationship between teaching and knowledge acquisition must be designed in a fashion to encourage student pilots to observe, analyse and express an opinion on the information being presented. The learning styles developed by Felder & Silverman (1988), identifies four dimensions that relate to individuals preferred learning styles, and illustrated in Table 3.

Table 3 Felder Learning Styles Dimensions (adapted from Felder & Silverman, 1988)

Learning Style Dimension	Type	Description
Perception (LSD1)	Sensitive (S)	Prefer to deal with facts, raw data and experiments; are patient with details, but don't like complications.
	Intuitive (I)	Prefer to deal with principles and theories; are easily bored with details and tend to accept complications.
Entry Channel (LSD2)	Visual (Vi)	Easy to remember what they see: images, diagram, time tables, films, etc.
	Verbal (Ve)	Easy to remember what they've heard, read or said.
Processing (LSD3)	Active (A)	Prefer to Learn in groups and work hands-on.
	Reflexive (Re)	Prefer to think and reflect on new information. Enjoy Working alone or with one other colleague at most.
Understanding (LSD4)	Sequential (Seq)	Prefer lineal reasoning process when problem solving; ability to work with specific knowledge once comprehended it partially or superficially.
	Global (G)	Generally, take large intuitive leaps with the information; possibly have a difficulty when explaining how they got a certain result.

The Felder learning taxonomies describe how learners remember, learn and use information. The responsibility is on the student pilot to recognise their own process to gain new knowledge, organise, and assimilate it for use in a real world task. The learning style dimension (LSD) = {LSD1, LSD2, LSD3, LSD4}, where each dimension is defined as a combination of four attribute values according to the LSD and the attribute values, are a combination of type values. Each type value is organised into a series of 44-item questions (Volume II, Appendix I Assessment Form 3), which are answered by the student pilot to evaluate their preferred learning style for each dimension. The preferences chosen by the student pilot is then expressed in metric form to visually indicate if the pilot might have difficulty assimilating knowledge in the classroom.

7.2.3 ILS Evaluation Process

The index of learning styles is a 44-item questionnaire to gain knowledge of individual's personal preference for each dimension and an amended version is described in Volume II, Appendix I Assessment Form 3. To account for the preferred learning style for both classroom and training practice within the FTS a modification to the original scoring method for the ILS is actioned. There are 11 questions that are posed for each type value dimension described in Table 3. Subsets of questions are grouped into their respective learning style dimension, as seen in Table 4. The student is persuaded to answer the questions to the best and most honest they can.

Table 4 ILS Grading Sheet (amended from work from Felder, 2008)

ILS Grading Sheet											
ACT/REF			SNS/INT			VIS/VRB			SEQ/GLO		
Q	a(1)	b(2)	Q	a(1)	b(2)	Q	a(1)	b(2)	Q	a(1)	b(2)
1	*2		2			3	*2		4		
5	*2		6			7	*2		8		
9			10			11	*2		12		
13			14	*2		15			16		
17		*2	18			19			20	*2	
21			22	*2		23	*2		24		
25			26			27			28		*2
29			30	*2		31			32		
33			34			35			36		
37			38	*2		39			40		*2
41			42			43			44		
Total (sum X's in each column)											
ACT/REF		SNS/INT		VIS/VRB		SEQ/GLO		ACT/REF		SNS/INT	
a	b	a	b	a	b	a	b	a	b	a	b
#	#	#	#	#	#	#	#	#	#	#	#
(Larger - Smaller) + Letter of Larger											
#	#	#	#	#	#	#	#	#	#	#	#
(if value is negative, place number and 'b' on the scale)											

When answering a question with an active preference, for instance, a '+1' is added to the value of the active dimension, likewise, if the reflective preference is chosen a '+1' is added to the reflective dimension. Each dimension score is summed together; in addition each type dimension can be assessed individually using linear scales by subtracting the type dimensions in one group, as illustrated in Figure 44. The green arrow indicates the ideal student characteristics for the FTS classroom material, in three groups the ideal student scores '0' between each type dimension; but for decision making especially in the air substantial information is gathered visually using MFDs and warning signals, therefore, a preference of visual over verbal is a suitable trade-off for this problem.

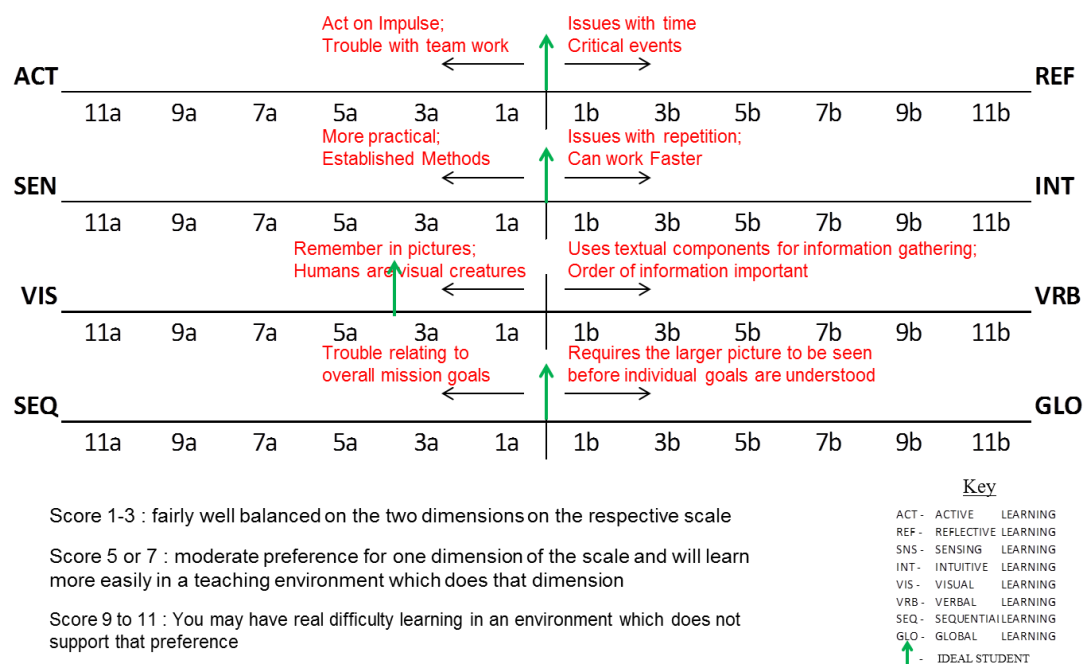


Figure 44 ILS Linear Scale Report Form Example

With careful analysis of the questions, the assessment can be structured to reflect domain considerations; as a result the emphasis of following SoPs, understanding the main goals of missions and using visual information for decision making is seen as important characteristics for student pilots to possess, hence, additional score is added to the relevant type dimension, as seen by the ‘*2’ in relevant boxes of Table 4. Each type dimension is further grouped allowing all the type dimensions on the left of Figure 44 to be summed together, likewise the right; to give distinction and preference between strengths and weaknesses of disparate learning styles as follows:

$$A = \sum_{i=1}^n (ACT, SEN, VIS, SEQ) = 55$$

$$B = \sum_{i=1}^n (REF, INT, VRB, GLO) = 47$$

Scaling the results and placing A and B on a linear axis, a quadratic representation of the ILS evaluation can be created ensuring that the axis of symmetry is in favour of A the following quadratic equation is created:

$$y = (-0.295)x^2 - 0.56x + 50$$

This places the axis of symmetry of the quadratic given the region of interest between -14-to-+12, at -0.95. The score for the prospective student(s) is given by subtracting the highest valued letter from the lowest. For quick diagnosis of suitability and which group of learning style is preferred, the display illustrated in Figure 45 is governed by a number of rules relating to the fill area colour under the cursor:

- If Score from ILS is $< -9.17 \vee > +7.28$, fill is ‘Purple’ signifying student may struggle with course material (Percentage rating of less than ~60% from ideal circumstances).
- If score from ILS is $(\geq -9.17 \wedge < -3) \vee (> 1.09 \wedge \leq 7.28)$, fill is ‘Yellow’ signifying student should be pliable to the fixed instructional process. (Percentage rating of between ~70.7 - 97.49% from ideal circumstances).
- If score from ILS is $> -3 \wedge \leq 1.09$, fill is ‘Green’, signifying the student will be able to cope with the fixed instructional process in the classroom. (Percentage rating of between ~97.5 - 100% from ideal circumstances).

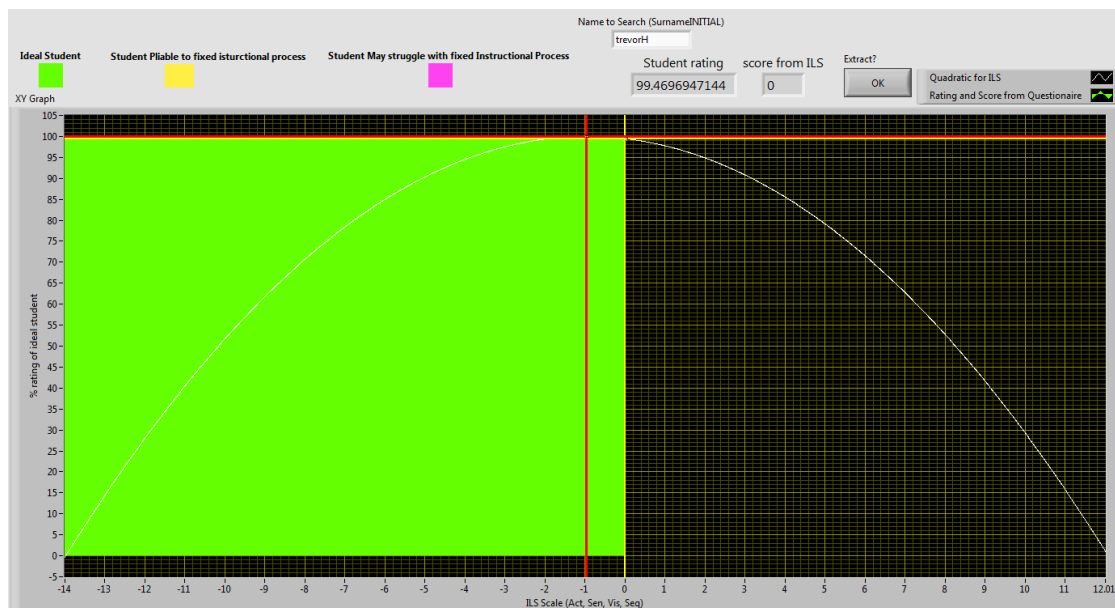


Figure 45 ILS Assessment Result Example

The red cursor denotes the axis of symmetry that designates an ideal student, whereas the yellow cursor specifies the results from the ILS evaluation. In this illustrative example, the score from the ILS is '0' indicating that the participant didn't favour any of the two groups of type dimensions. The fill is green indicating that the student pilot should not have any trouble attaining knowledge from the teaching material and the student rating, gathered from the quadratic plot, given is very high. Depending on what colour the fill is and where the yellow cursor is positioned will signify either that the student pilot will struggle attaining knowledge using the fixed classroom materials 'Purple', or will require intervention involving possible one-to-one instruction 'Yellow with a low student rating'.

This assessment stage is just to give an indication to the decision maker of any potential issues with learning a standard syllabus and is used as a precursor (i.e. a quick reference to learning styles) to any issued relating to understanding the goals and objectives of the mission scenarios, as per discussions in Chapter 4.4.3, that could be the cause of poor performance outcomes observed by the instructor that could affect the transfer of learning; and can be repeated if necessary. The behaviour of the ILS Use Case 'Evaluate Student Pilot Learning Style' can be described in Volume II, Figure I 4.

7.2.4 Personal Allocation Attention Assessment

The socio-cognitive process that may contribute to behaviour in adverse conditions is of utmost concern, especially in military training programmes. The personal characteristics have to be identified to determine whether individual's behaviour can be malleable and be

developed over a period of time. If individuals consistently lose focus in pressurized conditions, it is unlikely that they can maintain control and follow training behaviour when these conditions arise in the operational environment. Likewise, knowledge has to be gained on individuals attributes that are potentialities that can be express themselves in different situations and over time (Maltby et. al., 2010), furthermore, it is preferable to obtain a representation on ability to maintain motivation and deal with inconsistent information. As the decision maker is in a position to judge individuals' success in formal and informal settings, the expectancies of success can affect or colour the way an individual's behaviour is viewed. By focussing on behavioural aspects of the individual, using their own objective judgements about themselves, a better understanding of how attributes are preserved in the face of inconsistent events can be assessed. This enables decisions about acceptance and possible interventions in the future as training exercises become increasingly difficult and complex to understanding for the student pilot.

A simple questionnaire is directed at the student pilots, as described in Volume II, Appendix I Assessment Form 4, that uses a form of Likert scale to obtain psychometric opinions of how the student pilots measure their own strengths in six key attributes: detection times, handling levels, sharing effectiveness, ability levels under marginal conditions (fulfilment of levels of efficiency), maintaining control in fault conditions, and adaptability timescales. The grade is given a scale of 1 – 7 with a semantic description that expresses how much they agree or disagree with a particular statement; so tapping into cognitive and affective components of attitudes. Of interest to the decision maker is the percentage score the student pilot has graded themselves. The percentage is saved within the main student database file(s); with the specific answers available for further analysis if needed. The behaviour of the assessment can be described by the 'Assess Personal Allocation Attention' Use Case in Volume II, Figure I 5.

7.2.5 Pre-Study Summary

The designed architecture for this stage is seen in Figure 46, with the behaviour of the Pre-Student sub-system described throughout Volume II, Appendix I. The objective data from the questionnaires is converted to an integer value, the respected data is then transferred to the responsibility of the relevant block. The accuracy of the assessment is directly dependent of how honest the student pilot has been in completing the questionnaires; the instructor is the actor who uses the workflow tool to interrogate the information within the databases.

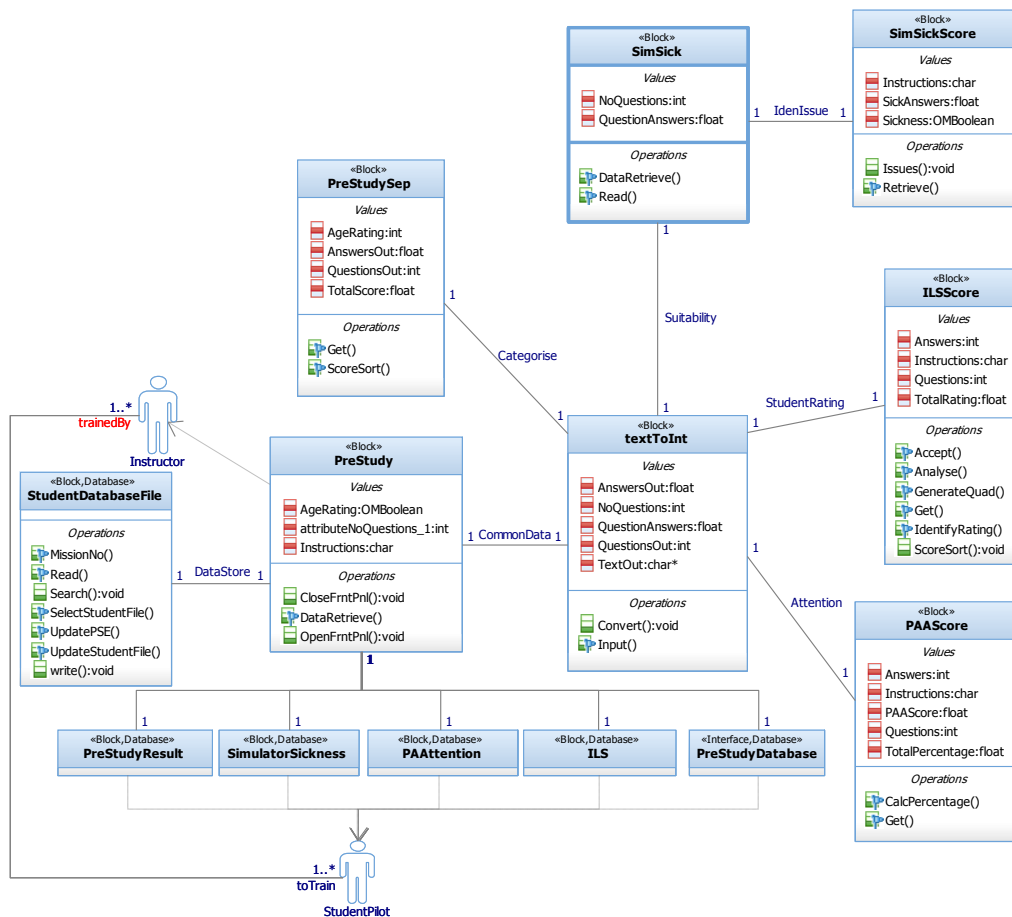


Figure 46 Pre-Study Architecture

The pre-study assessment stage is to gain a view on the student pilot's ability to cope with training for operational readiness and their ability to gain knowledge and follow SoPs using a fixed curriculum. Furthermore, a balanced instructional design should utilize diverse teaching approaches with fluctuating learning and assessment criteria to cater for diversity of individuals learning abilities. Where teaching material touches upon all learning styles, to assist all students to adapt their own learning styles, permitting equilibrium to be achieved across a group of individuals should allow for 'one size fits all' approach. Once this is achieved, true performance in relation to K&S acquired can be evaluated in mission scenario executions and those students who don't 'meet the grade' can be whittled out, based on their true ability and not due to their difficulty in learning on the chosen learning style not best suited to their individual strengths. Using visual clues provided by the workflow tool, gives quick identification of any potential issues with the prospective students, which can orchestrate further investigations and risk assessments for acceptance of the individuals onto the flight training course.

7.3 Handling Qualities Workload Scale (HQWS) Model

In relation to discussions in Chapter 4.3, for ease of communication and understanding, functions and interfaces should be common and consistent, requiring a reasonable number of tasks and methodologies on the part of the pilot. Control measures deal with the degree of difficulty with which the pilot can direct the systems performance during operation. This includes control capabilities to configure the system, but also the level of control that the pilot must exercise. The decision maker is required to plan missions to avoid overload in handling qualities of the pilot and thereby improve performance to accomplish expanded tasks as the pilot progresses through the pipeline. A balance between the amount of stability of the aircraft (virtual or real) and the pilot's ability to control its movement using the blending mix is required to be assessed on a per mission per task basis.

These handling qualities are integral characteristics of the mission, which is directed to ease and precision with which a pilot and aircraft is expected to perform the tasks required to achieve task goals. The problem is associated to the dynamic response attributes of both the aircraft and the pilot acting together to provide harmonised response to task objectives. Since human characteristics are difficult to describe analytically with accuracy; relying on the subjective and objective opinion(s) of the decision maker (instructor) is a viable option to measure capability of the human-technical cohesion for their handling qualities for compensatory events required for completion of tasks. These compensatory events are relatively insensitive to disparate aircraft configurations; simultaneously understanding of the effects of the pilot's situation awareness of the control differences is needed by alteration of difficulty levels required to achieve desired task performance. Various operational aspects of the student pilot required for controlling the aircraft handling qualities are considered as being: navigation, system monitoring, flight path control and command decisions. Various attributes need assessing for pilot specific handling qualities relating to workload qualities, these include:

- The mental effort required to perform the operations – the rating of work for the pilot needed to satisfy required operations;
- The physical difficulty to perform the operations – the rating of perceived physical work needed to action remedial tasks to satisfy required operations;
- Time criticality of operations - the rating of timescale needed before actioning relevant operations fail to satisfy task requirements;

- The understanding of the horizontal position indicator – a rating of the effect with distracting operations for the consideration of aircraft position in relation to the planned flight path;
- The time available to update the pilots current situation awareness and then make command decisions – the rating of time available to identify objects of interest, which requires a change of operational state of the aircraft by the pilot.
- The usefulness of information being presented to the pilot – the rating of pilot's ability to focus on relevant information, which determines the correct decision on what actions to take using trained SoPs.

The emphasis is to evaluate the pilot handling qualities with respect to mission goals in a simple and informative manner and metricise the subjective opinion of the decision maker with respect to the pilot current abilities of successfully accomplishing task goals. This stage is to be used alongside mission planning tasks for concordance with the development of the baseline scenario.

7.3.1 Handling Qualities Workload Scale (HQWS) Implementation Design

The HQWS architecture is dependent on the planned missions and the ability of the instructor to give an unbiased opinion on the ability of the student pilot to control the aircraft in a trained manner (via SoPs), the architecture is visualised in Figure 47. The 'HQWS' block ensures all relevant attributes necessary for the assessment are available to the decision maker. The operation of the sub-system read / writes data to three databases:

1. The 'TaskNoandDetails' database(s) stores an abstract representation of the number and sequence of tasks in the mission followed by the task description, which is needed for further stages in the workflow process. This file is saved in two separate directories: one for the mission database files and one for the student, which additional assessment grades from future stages of the workflow tool will be added.
2. The 'MissionRelatedHQWS' database stores the evaluation data from assessing the handling qualities for all attributes including the task score of each pilot HQW.
3. The student database is to feedback the Pilot Specific Rating (PSR) from previous completed training missions and is used as a compensating factor within the task score calculation in addition to placing average score for the pilot handling qualities in the relevant student pilot database file.



Figure 47 Pilot Qualities Evaluation Scale Architecture

The decision maker is requested to input the mission scenario reference and task details on the first iteration of the software loop, as seen in Figure 48a, with the number of loops highly dependent of the number of tasks in the mission; from the second loop the request is for the task details only, as seen in Figure 48b.

For the Pilot Subjective Evaluation Scale, Please Identify the Mission scenario reference Number and the task number.

Mission Scenario Reference No.

Task Number

Task Details

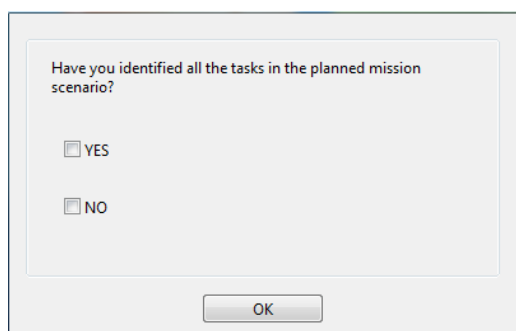
Please identify task details for additional tasks!

Task Details

(a) 1st iteration
(b) from 2nd iteration onwards

Figure 48 Commencement of HQWS

The number of loops is controlled using the workflow tool to prompt the decision maker to decide whether all tasks, which construct the mission scenario, have been evaluated as seen in Figure 49.



Have you identified all the tasks in the planned mission scenario?

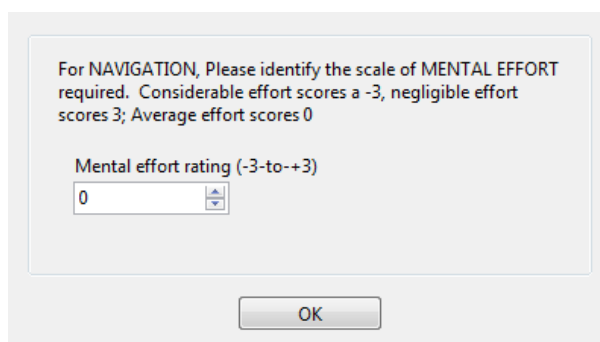
☐ YES

☐ NO

OK

Figure 49 Ending the Mission Handling Qualities Stage

Each pilot handling quality will prompt the decision maker to assess how the pilot will handle the current operation; an example is illustrated in Figure 50. Each handling quality is designed to follow a ‘event sequence’ method relating to the operation quality requirement and not all pilot handling qualities are assessed for each operation; further detail on the behaviour of the architecture can be found in Volume II, Appendix J.



For NAVIGATION, Please identify the scale of MENTAL EFFORT required. Considerable effort scores a -3, negligible effort scores 3; Average effort scores 0

Mental effort rating (-3-to-+3)

0

OK

Figure 50 Inputting assessment scores into the HQWS Sub-system

7.3.2 HQWS Evaluation Process

Evaluating the handling qualities is performed using a 7-point scale ranging from -3 to +3, as per Figure 50, in relation to the training attributes of the task. The assessment is generally completed without reflection of the training technology that is going to be used, however, as with all system design considerations the number of available solutions, in this case blending mixes that can accurately train the relevant K&S, will always be deliberated.

An aircraft operation that requires the specific handling quality of the student that matches their current ability level will be graded a ‘0’ meaning the quality should be at an ability level to adequately compensate for the difficulty of the task; A grade of -3 indicates the quality of the student pilot is not expected to be at a level to efficiently and correctly follow SoPs to compensate for the difficulty of the task operation; A grade of +3 gives a clear indication

that the quality is well within the abilities of the pilot to compensate for any operational difficulty associated with the task.

Each pilot quality is evaluated independently of others to obtain an objective judgement of how the student pilot will cope with each operation of the task. The pilot qualities are averaged per task and the average grade for the mission is calculated leading to a guide indicating how the pilot will perform with task specific knowledge and given appropriate briefing of what activities are required for each task in the mission. The output of the assessment is stored in a text file, where tasks are separated to allow for easy visual examination, an illustrative example using a subset of Oscar123 mission scenario, used as proof of functional concept of the method, can be seen Table 5.

Table 5 MissionRelatedHQWS Assessment Output File (Subset)

Mission Scenario Ref.	Oscar123					
Task Number	1.000000					
Task Details	maintain height and he					
Operations	Mental Effort	Physical Difficulty	Time Criticality	Understanding of Horizon	Time Available	Usefulness of Information
Navigation	0.000000	1.000000	2.000000	3.000000		
System Monitoring	1.000000	2.000000	-3.000000			
Flight Path Control	-2.000000	-1.000000	0.000000			
Command Decisions	2.000000				-2.000000	2.000000
Task Score	0.250000	0.666667	0.250000	3.000000	-2.000000	2.000000
Task Number	2.000000					
Task Details	begin bank turn at 3De					
Operations	Mental Effort	Physical Difficulty	Time Criticality	Understanding of Horizon	Time Available	Usefulness of Information
Navigation	1.000000	2.000000	3.000000	2.000000		
System Monitoring	1.000000	0.000000	-1.000000			
Flight Path Control	-2.000000	-3.000000	1.000000			
Command Decisions	1.000000				2.000000	3.000000
Task Score	0.250000	-0.333333	0.250000	2.000000	2.000000	3.000000

At the top of the file is the identification of the mission scenario reference, followed by the task number and task details. A matrix representation of the assessment is then situated beneath to present the manual rating given by the decision maker. The average task score for each pilot quality placed beneath the last aircraft operation quality 'Command Decisions'.

The average grades for each qualities dimension are used to calculate the pilot specific rating (PSR), which describes the estimated ability of the pilot to manage the qualities contained in the mission tasks. The calculation uses feedback from previous assessed missions to adjust the current rating based on the decisions maker's previous objective assessment, this is used to keep track on the decision makers judgement on student pilot experience and K&S acquisition to cope with the stressors from a similar task experienced previously. As the student pilot progresses through the pipeline it should be expected that the PSR value increases toward 100%. The average grade for the task is placed on the identical rating scale,

which is then converted to a percentage value indicating the percentage PSR, the feedback of previous missions percentage average PSR is then used to automatically amend the rating, as follows:

- If $PSR(Feedback) \geq 65$, then $Task_iPSR = \frac{PSR(Feedback)-65}{2.5} + \%AveTaskRating$
 - If $Task_iPSR \geq 100$, then $TaskPSR = 100$, else, $TaskPSR = Task_iPSR$
- If $PSR(Feedback) \leq 45$, then $TaskPSR = \%AveTaskRating - \frac{45 - PSR(Feedback)}{5}$
- Else, $TaskPSR = \%AveTaskRating$.

Once the TaskPSR has been calculated for all the tasks in the mission, the total average PSR for the planned mission is then saved within the main student database file and each TaskPSR is included in the student specific 'TaskNoandDetails' database file for future assessment stages, an example of subset of Oscar123 is seen in Table 6b. Each task grade value, illustrated in Table 5, is averaged and also saved within the main student database file. An additional 'TaskNoandDetails' file lists the task and task details, an example is illustrated in Table 6a and saved within the Mission_Scenario_Planning directory; this is used as quick reference to obtain information of what specific are involved in the planned mission. The programming of this stage can be easily advanced to include the number of activities that create one task for a more detailed analysis and thereby satisfying the relationship between the 'Task and Activity' blocks described in Figure 34 (The Readiness Metamodel).

Table 6 Subset of TaskNoandDetails Mission File

Mission Scenario		PSR Student Qualities File		
Oscar123		Mission Scenar	Oscar123	
Task Number	Task Description	Task Number	Task Description	Pilot Specific Rating
1.000000	Maintain height and heading and velocity for	1.000000	Maintain height and heading	52.5741
2.000000	begin bank turn at 30Deg	2.000000	begin bank turn at 30Deg	60.9074

(a) Mission Scenario Planning Directory File

(b) Student_Pilot HQWS Qualities File

7.3.3 HQWS Summary

Pre-occupation with personal beliefs about the abilities and personalities of someone generally affect observer ratings. Deficiencies in assessment through bias can be compensated for by the structure of the questions being posed, which enforces the decision maker to assess and evaluate attributes that they may not have considered relevant within the student pilot evaluation process. Assessment of single parameters against a variety of others and documented within the training database can be very helpful for the evolution of the FTS to improve strategies of decision efficacy due to the development and presentation of such

data. The assessment can be extrapolated with actual performance data and a comparison between each should counteract bias with views on pilot's ability to action various operations. The results of the HQWS can be used as feedback to the pilots in context of performance and workload in relation to task requirements. This is an important aspect of communication between instructor and pilot as it gives both the opportunity to comment on the perceived difficulty of performing mission related tasks including stress and possible 'startle' effects on reception of events.

If the decision maker continues to rate the pilot low on specific pilot qualities and the subjective assessment is strengthened by poor performance completing certain tasks and activities with the execution of the mission scenario, it is an indication that an intervention is necessary concentrating on those relevant pilot handling qualities that are evaluated as being weak. Used in combination with assessments in further stages of the workflow process i.e. workload, HQWS should reflect the difficulty that the pilot is experiencing in certain aspects of decision making including actioning relevant compensatory behaviour.

7.4 ROSETTA Level 0 Model

Continuing the brief introduction given in Chapter 5.2, the first ROSETTA stage concerns the selection of relevant mission essential competencies (MEC) including, supporting competencies, knowledge and skills (K&S), and experiences required for the student pilot to successfully complete the tasks of the mission scenario (see Chapter 7.6 for further information on MECs). The experience gained by executing the mission scenario on the chosen blending mix is directly associated with the allocated K&S of the mission and additional competencies can be determined on performance measured in the K&S dimension. In essence, a deficiency that is noted in K&S is also distinguished in the MEC; the particular K&S identified can be practiced by creation of a mission scenario directed at experiencing the K&S thus gaining critical experience that emphasises deficient K&S either in the training technology or the student pilot. This will ultimately cultivate competency in associated MECs, the relationships can be seen in Figure 41.

It is important to classify all MECs and supporting competencies for each task and identify quantifiable relationships between K&S required to be practiced and the suitability of technology for attaining these competencies: the relationship between MECs and the training technology is seen as being too complex to rate subjectively by SMEs. The relationships between K&S and training technology need to have a construct beyond the binary relevant /

not-relevant ratings; the assessment should be surveyed within workshops, where SMEs would determine the quantitative relationships between the non-functional worker oriented requirements (K&S) and the available (or new) training technology systems. With identifying relevant competencies required for completion of tasks and identifying relationships between competencies, mission scenarios can be evolved to maximise their effect on competency development with the training technologies; thereby develop training regimes that focus on K&S in expectation of accelerating the proficiency process.

7.4.1 ROSETTA 0 Implementation Design

As a reminder to the decision maker of the planned mission, the task details are acquired and displayed on the UI. A predetermined list of training attributes are gathered from an already predefined list of standardized training attributes (i.e. MECs and K&S) from a database created from the existing competency framework (see Figure 2) for the purpose of full traceability of training attributes to tasks of the mission, along with a list of systems within the FoS. These are used to scale the ROSETTA framework and enable quantitative results to be stored within the slots of the framework. A number of small databases are then used to save the results of the evaluations, which can be used for further analysis or for future stages of the workflow process. The architecture for ROSETTA 0 is seen in Figure 51, with detailed behaviour described in Volume II, Appendix K.

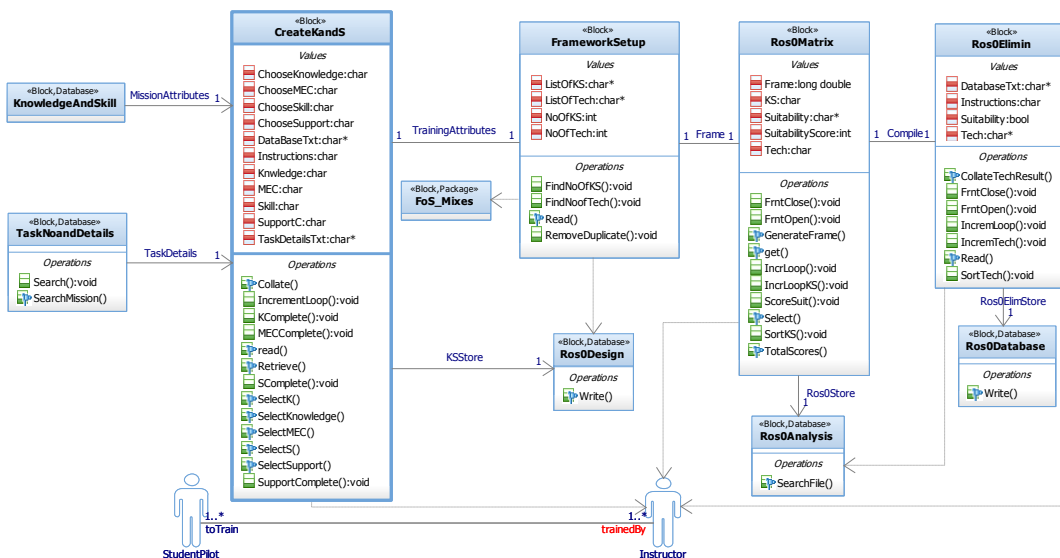


Figure 51 ROSETTA 0 Architecture

Selection of Relevant Training Attributes Per Task

Each task within the planned mission requires relevant training attributes to be assigned to give clarification of the training specifics that the student pilot needs to know before

evaluation of the mission scenario. The tool uses tabs to prompt the decision maker to select, from the list of available training attributes in the database, a list of relevant mission related competencies that need to be trained for each task, as per Figure 52a, with the flow of prompts illustrated in Figure 52b.

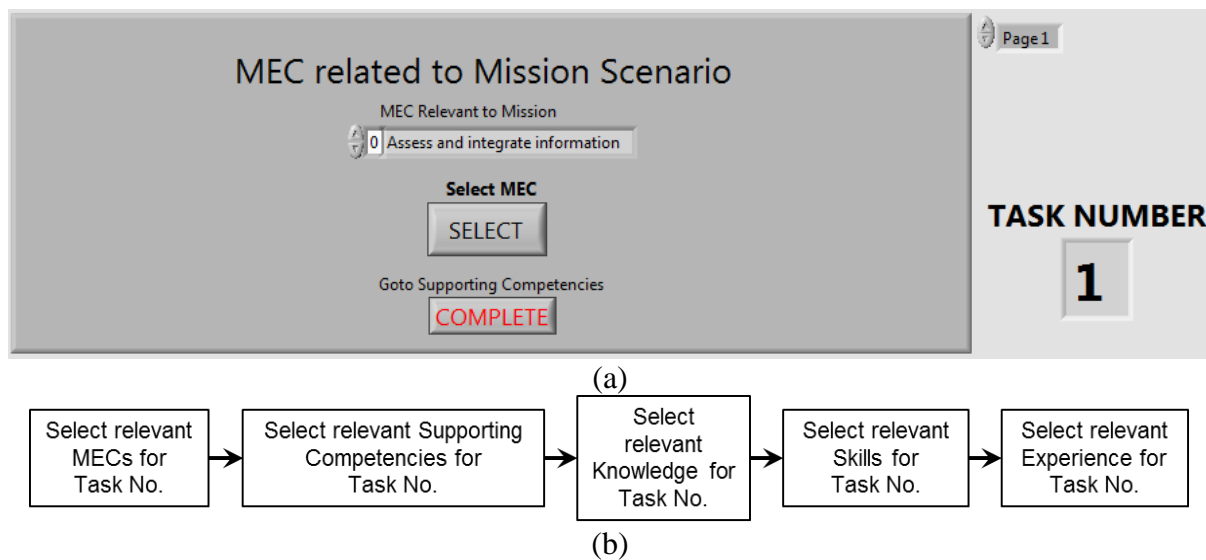


Figure 52 Flow for Selection of Mission Related Competencies

In the MEC example above, the decision maker uses the scroll button to select the relevant MECs for the mission task number; once all the relevant MECs for the task have been selected the decision maker completes this section of the assessment and the tool transfers tabs to permit selection of supporting competencies and so on and so forth. Once all the experiences relevant to the task have been selected the process is repeated for all other tasks in the mission scenario. The output of the design stage produces a database file saved in ROSETTA0 Design directory, which list the training attributes against mission tasks, an example is seen in Table 7. There are repeated activities the pilot will action during different stages of the flight, therefore, it is expected that there will be a duplication of identical competencies in different tasks (for this example this is not the case). The behaviour of ROSETTA 0 design is described by the 'Create K&S for Mission' Use Case in Volume II, Figure K 2.

Table 7 Output File From ROSETTA 0 Design

List of Competencies per task					
Mission Scenario Ref.	Oscar123				
Task number	MECs	Supporting Competencies	Knowledge	Skills	Experiences
1	Assess and integrate information	Information Management	Phase of Mission	manages stress and pressure	Limited time to act/react to situation
1					Task Saturation
2	Process and Analyse information	Identification	Time restrictions	Multi-tasks	Fatigue / Time on task
2	Dynamic execution	Flight Management (Battle)		Sorts information	Mountainous terrain
2					Restrictions to visibility

The database produced permits traceability of training attributes of interest in the planned mission scenario to each task in the mission (or even to each activity, if necessary). This database can be used for reference to identify relevant competencies of any task within the evaluated mission scenario (see Chapter 7.13) that future training missions are required to focus on due to under performance. It also gives a clear sign to the decision maker of what tasks give student pilots issues using a particular technology, by using a comparative study for all sorties flown either by a virtual or real aircraft for all student pilots being evaluated. This information can be used to investigate the training technology characteristics (and student pilot current ability levels) in relation to student pilots actioning objectives required by the mission scenario using the training technology controls and displays.

7.4.2 ROSETTA 0 Evaluation Process

As per the House of Quality (HOQ), ROSETTA maps a set of customer requirements to a set of engineering characteristics (design metrics) to identify key design drivers and trades. This level of ROSETTA concentrates on the non-functional worker oriented requirements (K&S) that map over to the high level training technology configurations within the FoS descriptions. A non-exhaustive list of systems within the FoS can be found in Volume II, Table VII. SMEs are expected to identify what engineering characteristics are required to fulfil the needs of the requirements and describe how strongly each engineering characteristic is related to each requirement. Of interest is assessing the technology based on: size and numbers of monitors, monitor configuration, flight control layouts, ergonomic design, and emulation of cockpit layout to K&S importance levels with training objectives.

The size of the ROSETTA framework is dependent on the number of identified K&S and the number of available training technology/system(s) within the FoS. Redundancy in K&S is eliminated from the output file of Table 7 first and the list of blending mix technologies available is gathered, as described by the behaviour of ‘Set-up Framework’ Use Case in Volume II, Figure K 3. Once this information is imported into the tool, the decision maker is prompted to rate the suitability of the technology configuration on attaining the relevant K&S, as illustrated in Figure 53; this evaluation is accomplished without the need for detailed technical knowledge of the training technology/system.

The Mission scenario, with respect to the technology configuration, is **Adequately Effective** **SELECT**

for evaluating or exercising this applicable Knowledge or skill, **Phase of Mission**

For the ToT to the real world.

Figure 53 Prompt to Decision maker for ROSETTA 0

Each technology configuration is evaluated with respect to each K&S, and the evaluation has to consider whether or not pilots can practice the K&S using the training technology and be applicable for transfer to the real world. The evaluation is based on semantic reasoning, described in Table 8, which maps to a quantitative scale declaring views on suitability. The semantic description is found by using the horizontal scroll bars and accepted using the ‘select’ function. Once the descriptor have been accepted the tool sequences through all other K&S for a particular type of technology configuration before repeating the evaluation for all other technology configurations identified within Volume II, Table VII.

Table 8 Scale descriptors for ROSETTA 0

Semantic Mapping	Quantitative scale
Not all Effective	0
Vaguely Effective	1
Slightly Effective	2
Somewhat Effective	3
Quite Effective	4
Adequately Effective	5
Very Effective	6
Exceptionally Effective	7

The subjective opinion is directed to the layout of the training technology and a sense of suitability for the student pilot to gain and practice relevant K&S levels before proving ToT with real flight. Once the assessment is complete, a matrix describing a ROSETTA frame is created expressing the quantitative relationships between the K&S and the technology configurations, as described in Table 9. Relationships are in the form of an integer number, which describes the strength of suitability to signify which technology configuration carries more weight in successfully satisfying training objectives for the current planned training mission. The quantitative results are summed to give a clear indication of the full mission fitness of the technology. The behaviour is expressed in ‘Assess K&S with Technology Configuration’ Use Case in Volume II, Figure K 4.

Table 9 Example of a ROSETTA 0 Frame

ROSETTA 0 Evaluation Results

Number	reference	Knowledge & Skills		Time restrictions	Multi-tasks	Sorts information	manages stress and pressure	Total from assessment
		Phase of Mission						
1	BTD	4		2	1	4	4	15
2	DTH-24	5		4	4	5	5	23
3	DT-KS	5		6	6	5	4	26
4	DTCC	5		6	5	4	3	23

The results are then transferred to the ROSETTA 0 elimination phase, where the decision maker is asked about which training technology or configurations are unsuitable to carry forward for further evaluation. The elimination GUI is illustrated in Figure 54, where a simple Boolean question is asked about the training tool suitability. Those that are deemed unsuitable are eliminated from consideration in further stages of the workflow process. Once the technology has been eliminated, it plays no further role, therefore, all evaluation data from ROSETTA 0 for that particular technology, is removed from the updated version of Table 9, which carries forward to future elimination stages. Behaviour can be found in ‘Eliminate Technology in ROSETTA 0’ Use Case in Volume II, Figure K 5. This stage is for efficient evaluation of the current technology for rapid elimination of technology that clearly will not satisfy training objectives. The full behaviour of ROSETTA 0 stage can be found in Volume II, Appendix K.

ROSETTA 0 Elimination Selection

Number	reference	Knowledge & Skills		Time restrictions	Multi-tasks	Sorts information	manages stress and pressure	Total from assessment
		Phase of Mission						
1	BTD	4		2	1	4	4	15
2	DTH-24	5		4	4	5	5	23
3	DT-KS	5		6	6	5	4	26
4	DTCC	5		6	5	4	3	23

Do you wish to eliminate the VRTE tool designated **BTD** **From future Analysis**

YES **NO**

Figure 54 ROSETTA 0 Elimination Stage

7.4.3 ROSETTA 0 Summary

The beauty of this type of approach is its simplicity, but care is required in collecting the correct type of relationship. Careful consideration of the question being asked is needed as the decision maker has to acknowledge the overall goal of transfer of training in the assessment. Unlike the QFD approach a series of seven quantitative relationships can be chosen based on the semantic descriptors to identify the strength of the relationship used for

subjective opinion on the suitability of a technology configuration to a particular knowledge or skill. It is important to consider which K&S is being assessed for evaluation in the planned mission scenario as this may affect how the assessment is graded. The overall total(s) in the assessment gives an abstract measure of how suitable a technology configuration is for successful acquisition of the correct levels of K&S to complete the associated mission scenario goals. In addition, the decision maker will always reflect on the availability of technology, consequently, this will influence the decision made in the elimination stage in that it is possible that the lowest overall score maybe a technology that is carried forward to further stages due to greater availability. However, a relationship of rating 0-2 should be a clear indication that this K&S cannot be evaluated using this technology configuration for the goal of ToT and the technology should be eliminated from future assessments, unless the planned mission utilizes this K&S but not specifically planned to evaluate them during execution, i.e. evaluation on other more important K&S. These types of decisions are aimed at maintaining the focus of the decision maker on the specific aspects of the mission objectives in relation to the organisational goals.

7.5 Pilot Workload Evaluation

Since human resources are limited, a system must not overload the HITL, otherwise severe performance deterioration can occur, as per workload Factors discussion in Chapter 4.3. When the task exceeds the pilot's ability to cope, there is no awareness of input information from various sources, so decisions may be made on incomplete information and the probability of error increases. Furthermore, the same levels of performance can impose different levels of workload leading to the necessity of measuring workload as well as performance. Workload data can aid trade-off decisions needed for mission planning and technology allocation to insure that pilot workloads are at acceptable levels across time and optimized for sustainable tasks. Workload can be evaluated by analytical techniques in the design phase of the system. Strengths in learning styles followed by subjective methods using rating scales and pilot's ability levels can be used to provide an indication of student pilot's workload rating for a task.

This thesis will concentrate on integrating SA and workload attributes into the proposed methodology with data from subjective assessments from instructor and student pilot, both pre- and post-mission scenario execution in relation to mission goals and the chosen blend for mission execution

7.5.1 Instructor Evaluation Of Pilot Workload with Scenario Tasks

When missions are decomposed into phases, functions and tasks, the task analysis and workload technique can be used to place importance metrics to tasks from SME consultation. A scheduling mechanism considers both an estimate of the time it takes to complete a task and the availability of visual, auditory, cognitive, and motor channel resources (Miller, 2001) and is described by Volume II, Appendix L Table VIII; the theory assumes the descriptors are independent of each other. Workload values are assigned scale ranges from 0.4 (minimal workload) to 7.9 (High Workload) with associated descriptors that increase in value with an increase processing level of information for the pilot (Wojciechowski, 2006); a total workload value is summed from the all the tasks and is used as the value of average workload for the mission. The advantage of the method is that it quantifies whether a task can be scheduled in a straightforward fashion. The tasks are combined into a scenario timeline and workload estimates are given for each point on the timeline. The proposed method required the SMEs to rate the workload for each task according to: auditory, cognitive, kinaesthetic, psychomotor, and visual dimensions. The architecture for the workload design can be seen in Figure 55; the ‘TaskNoandDetails’ database is used to acquire details of mission tasks; the ‘Workload’ block uses prompts from the ‘WorkloadCat’ block to permit the workload values to be entered, which are then used to calculate average workload that is saved in the main student database file. The behaviour of the architecture can be described throughout Volume II, Appendix L.

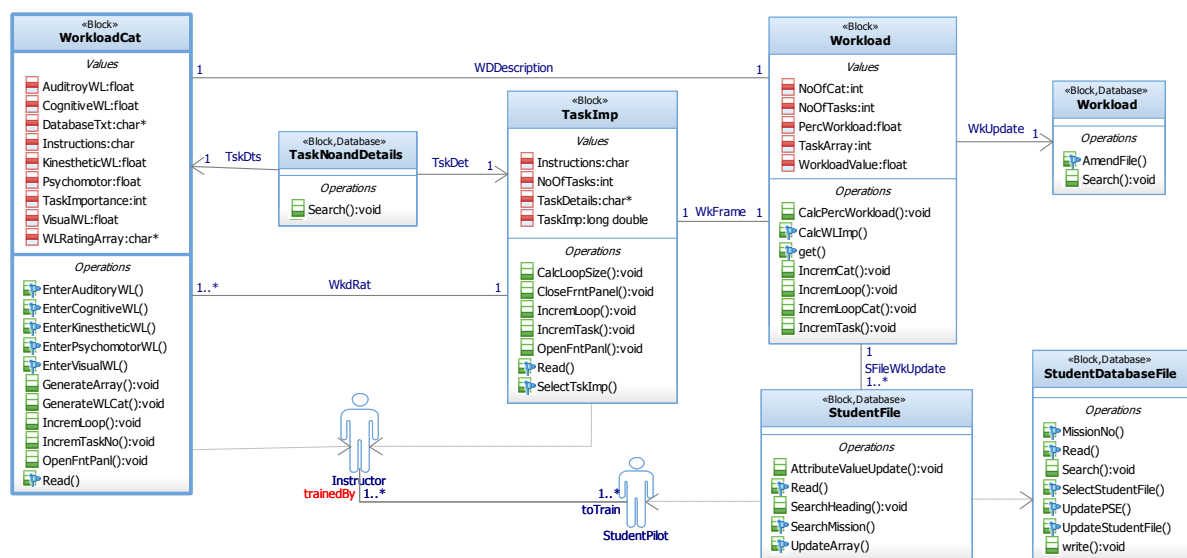


Figure 55 Workload Planning Stage Architecture

The decision maker need to place a concurrent rating of workload for a set number of strategies that the pilot has to action: the information should be based on SME consultation of a typical flight involving changes in state of the aircraft. The assessment can be completed at either the task or activity level, depending on the level of information needed to satisfy the criteria; for student pilot's that are struggling with certain aspects of tasks, concentration is given to assessing activities involved in the completion of the task in an attempt to discover the underlying activities/actions that are causing the student pilot to fail.

7.5.2 Workload Implementation With ROSETTA Design

The task details are retrieved from the 'TaskNoandDetails' database, which determines the number of loops the assessment has to progress through; the verification example is kept at the task level for efficient explanation. When the decision maker is prompted to enter the workload value the task or activity number and the description is displayed on the UI with the rating constrained to 0.4-7.9 and input values controlled by the scroll bar function with 0.1 value increments, as per Figure 56. The activity or task description remains constant until all the workload values for all dimensions have been rated, then the next task or activity and description will be displayed and the process is repeated until all tasks in the mission have been rated with workload values. The behaviour is described by the 'Grade Workload Per Task' Use Case in Volume II, Figure L 2.

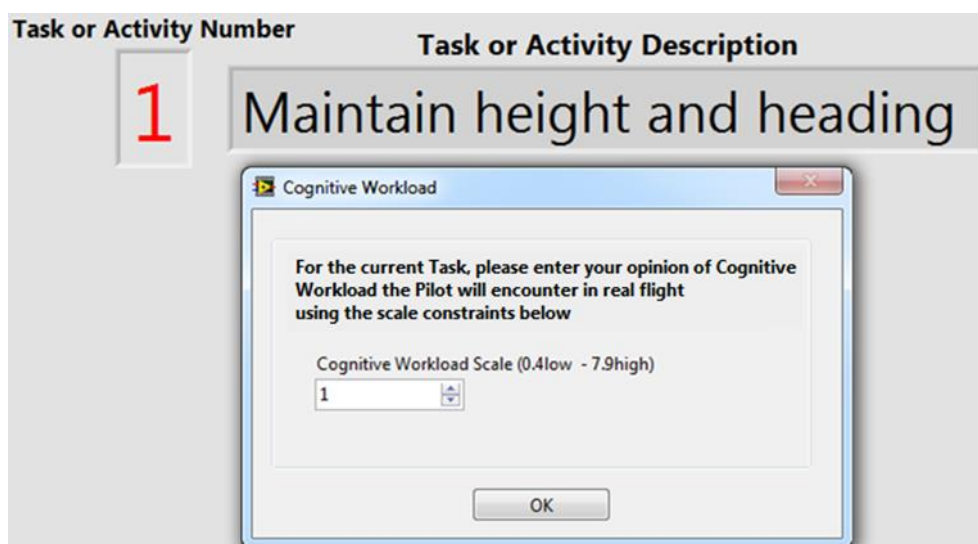


Figure 56 Entering Workload Values

Once the workload values have been accepted, the importance of the task has to be clarified, as per Figure 57. For evaluation of mission criticality of the task or activity involved, the tool uses a 10-point scale for the decision maker to decide the importance to the successful

completion of the mission goal(s). The scroll bar allows the rating to increment by 1 and it is the responsibility of the decision maker to make a factual objective assessment based on knowledge of the task or activity within the planned mission. The behaviour is described by the 'Assign Task Importance Metric' Use Case in Volume II, Figure L 3.

Figure 57 Task Importance Related to Mission Goals

Once all the data has been collected the workload design evaluation results can be seen in tabular form, as per illustrative example in Table 10, which gives as summary of the assessment process.

Table 10 Workload Design Evaluation Results

Workload Design Evaluation Results							
Mission Scenario	Oscar123						
Task Number	Task Description	Task Importance	Cognitive workload	Visual Workload	Auditory Workload	Kinesthetic Workload	Psychomotor Workload
1.000000	Maintain height and heading and velocity for 'n	3	2.0	1.8	1.0	1.5	3.1
2.000000	begin bank turn at 30Deg	7	5.5	4.7	3.1	2.9	5.4

Of interest is to calculate the total average workload for the whole mission, furthermore, this calculation also gives the total workload per workload dimension, as follows:

$$Tw_i = \sum_{i=1}^n T_i \cdot W_v$$

Where,

Tw_i is the total workload per dimension

T_i is the total importance for the task or activity

W_v is the workload value

n is the number of tasks or activities

Once these values are calculated a new array is created which gives the total workload per dimension in consideration of task or activity importance. The total average workload for the mission is calculated by summing the array values and converting the value to a percentage by using the number of tasks in the mission as the basis of calculation.

$$AvW = \sum_{i=1}^w \frac{(Tw_i)}{\left(\frac{n * 70}{100}\right)}$$

Where,

AvW is the percentage average workload for the mission
w is the number of workload dimensions

The total average percentage workload value, in this example 25.6286%, is saved in the main student database file. This value also signifies that the majority of the tasks in the mission do not test the student pilot's ability to cope with stress conditions; whereas a high value (>65%) indicates that the planned mission maybe too difficult for the student to accomplish at their current ability level or the identification of possible weaknesses in pilots acquired K&S to cope with various stressors. The behaviour is described by the 'Calculate Workload Importance' Use Case in Volume II, Figure L 4.

7.5.3 Workload Summary

The separation of workload dimensions along with tasks and activities can give a more detailed analysis of not only the student pilots ability when compared with actual performance results but also any possible bias made by the decision maker when assessing workload values per student. There is behavioural and subjective evidence stored within this type of assessment completed in this way and incurs the ability to identify key areas of strengths and weaknesses of student pilots and can influence further training missions to concentrate on various tasks or processes that are identified as being weak. Repetitive training concentrating in the trade-off on action(s) or processes in simulations of high workload scenarios can alleviate risk during live execution (theory known as muscle memory (Morie et. al., 2011)). This type of training can, in theory, identify to the pilot recognition of low SA, and as result can increase the accuracy of the human monitoring process and evolve the training programme to combat cockpit management issues during high workload tasks (Plioutsias & Karanikas, 2015).

Although the workload assessment should be completed with consideration to real live flight is can be used to identify ToT deficiencies with the chosen blending mix. The workload assessment completed in the design stage can be evaluated with feedback data from performance (Chapter 7.13) and the student pilot subjective feedback can be used to identify in what ways the training programme needs to evolve to compensate for identified weaknesses (i.e. mismatch of student pilot to the chosen blending mix technology). One of the main issues with workload is the lack of familiarity with cockpit layouts which has been criticized by student pilots as the main reason for difficulty (Salaud, 2013). Pilot's familiarity with cockpit layouts and VRTE configurations is critical for improving efficiency in training.

As competency is linked to knowing what to do with the need to think about it, unfamiliarity inevitably increases workload and may contribute to loss of situation awareness. This type of knowledge can influence the workload dimension values and consequently reduce performance outcomes; however, if identified at the design stage the performance expectations of the student pilot can be altered to suit the chosen blending mix used to execute the mission scenario.

7.6 ROSETTA Level 1 Model

For military applications, training can be defined as ‘the ability to enhance the capability to perform specific functions and tasks in order to improve individual abilities to accomplish mission objectives’ (OUP, 2013). This definition includes an explanation of the beneficiary from training, which includes improvement of abilities for overall mission accomplishment. Establishing quantifiable links, using modelling techniques, to analyse the effects of training should allow flexibility to include all types of K&S required to improve the ability of individuals and enhance learning techniques to enable the individual to successfully complete the mission, and incur a degree of ToT, as in discussions of Chapter 4.4. Training effectiveness is the general study of the individual, technology/systems and organisational characteristics that influence the training programme, thus, it is natural to study the causal relationships that influence training outcomes before training occurs.

Pilots tasks are more associated with decision making and monitoring functions using advanced technology integrated into the cockpit, as a result it is important to understand and enhance systematic training missions that promote the acquisition of K&S using the training technology/system and promote sustainable positive changes in behaviour and cognition to achieve overall mission objectives and obtain MECs. To successfully predict the effectiveness of training, the design of the training system and the technology used for interaction must be examined to assess which learning dimensions (K&S) are used in the design, and if they are appropriately applied (Wickens & Holland, 2012). Gaining an understanding about the relationships between effective learning and the training environment is needed to enable scientific analysis of the flight training problem (Colegrove & Alliger, 2002).

7.6.1 K&S Relationship To MEC Identification For Technology Suitability

Some MECs may be temporarily parallel, while others are continuously live throughout the course of the performance of the mission tasks and/or throughout the training programme.

The supporting competencies underlie the successful development and performance of the MEC and include: decision making, adaptability, and situation awareness, which are generally highly contextualised job functions that draw their character from specific tasks. K&S are a low level competencies which are used to identify the relevant ‘actions’ required by the student pilot to accomplish the mission tasks and are both ‘worker-oriented’ detailing human characteristics required for successful job performance. Training for the frontline requires that stress be the key emphasis in K&S; the baseline for the success of training is dependent of the elicitation of the level of K&S acquired for the specific training task(s) using the blending mixes.

Mathematical functions (response surfaces) define the relationships between K&S elements to both the MEC’s and the fidelity dimensions. The subjective and objective feedback will either strengthen or weaken the relationship between the K&S/MEC’s and the technology fidelity dimensions based on previous performance using the blending mix. A further consideration is then given to the experience level of the student pilot, i.e. every training mission increases the level of experience irrespective of goal success. This will then give an indication of how well the student pilot should perform, especially in the cognitive nature of tasks, on the next mission given the identical K&S acquisition requirements; can give an early indication of possible ToT into the real world.

With the distinct absence of experience within flight training operational environment and hence insufficient knowledge with which to make judgements about these relationships (see Figure 8 – Recognition Primed Decision Model), the RSEs used in the frameworks are for proof of functional concept of the ROSETTA methodology only, as SMEs are required to transform qualitative opinions to quantitative relationships.

7.6.2 ROSETTA 1 Implementation Design

The design of the ROSETTA framework consists of two stages, one to scale the framework according the number of parameters needed and the other for analysis of the data within the slots of the framework. The framework consists of retrieving data from previous assessment stages and asking for the decision maker to input importance ratings for both the MECs and K&S that determines the region of interest (ROI) for the generation of the RSEs used in describing the relationship between them for each remaining technology in the mix. The design of the framework also includes functions to output database files to add importance ratings to existing databases including writing the design RSE and the Analysis RSE arrays to

separate databases for further analysis and foreseeable workshops for discussions on the shape of the RSEs. The design architecture can be seen in Figure 58, with the behaviour described by ‘Produce ROSETTA 1 Framework’ Use Case in Volume II, Figure M 1.

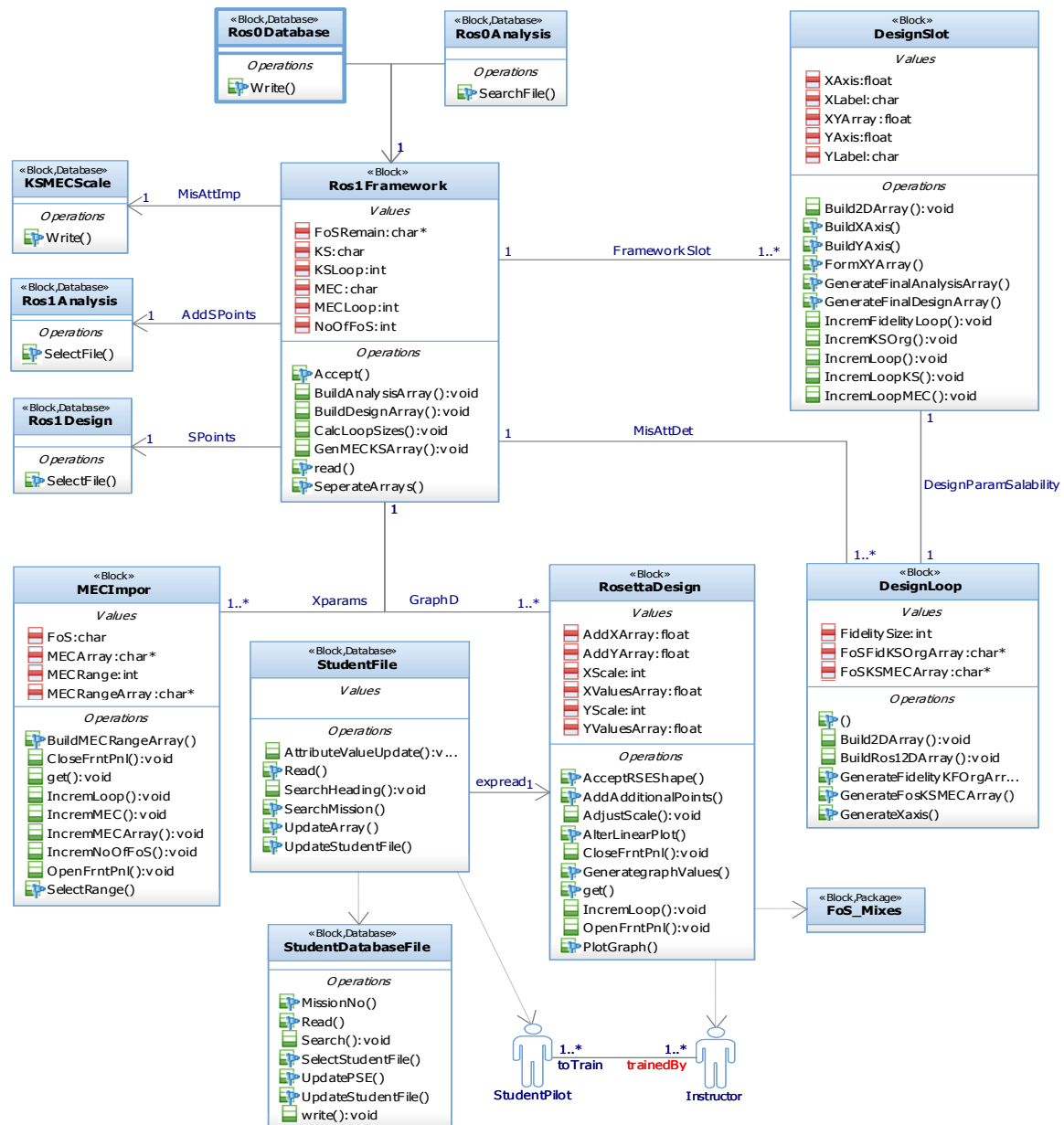


Figure 58 ROSETTA 1 Design Architecture

The design commences with retrieving the list of remaining technologies/mixes from ROSETTA 0 elimination file, as per Figure 59, which states the quantitative rating of K&S importance to training objectives. (Note: In this illustrative example the technology that has been eliminated was the laptop desktop training using keys for specific control of the aircraft; the remaining technologies include the use of HOTAS and rudder pedals with multiple screens and display sizes.)

ROSETTA 0 Elimination

Number	reference	Knowledge & Skills Phase of Mission	Time restrictions	Multi-tasks	Sorts information	manages stress and pressure	Total from assessment
2	DTH-24	5	4	4	5	5	23
3	DT-KS	5	6	6	5	4	26
4	DTCC	5	6	5	4	3	23

Figure 59 ROSETTA 0 Elimination Database File for OSCAR123

The decision maker will then be asked to retrieve the list of training attributes per task from ROSETTA 0 design stage. The list of MECs will then be used to prompt the decision maker to rate the MEC by a level of preparedness using each technology in sequence with each identified MEC, as per Figure 60a&b. The MECs identified by the assessment in Level 0 example are: Assess and Integrate Information, Process and Analyse Information, and Dynamic Execution; therefore, the prompts will ask for the maximum MEC range the training technology/system is capable of providing the student pilot at their stage in the pipeline by assessing one remaining system of the FoS at a time. The behaviour can be described by the 'Allocate MEC Grade' Use Case in Volume II, Figure M 2.

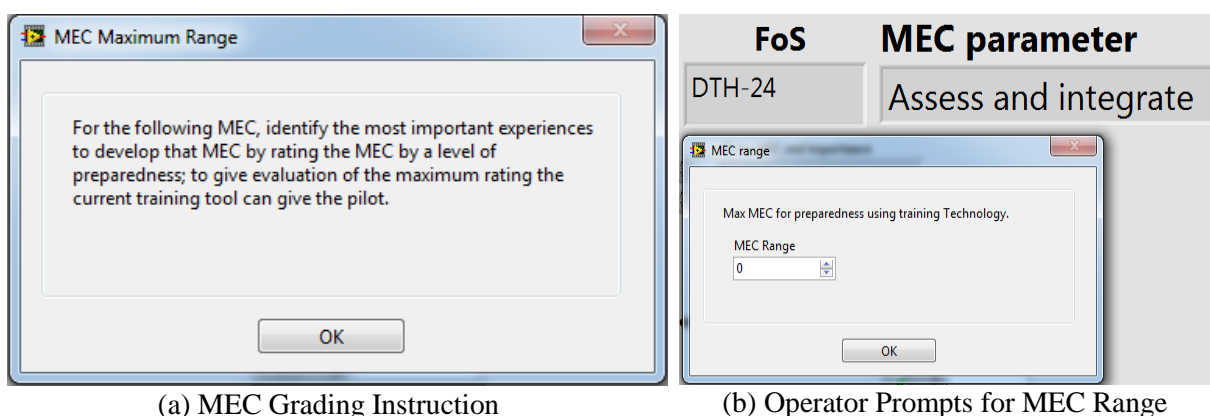


Figure 60 Identification of Level of Preparedness

The level of preparedness is based on a scale to estimate the pilot's maximum proficiency level and thus provide information about areas of MEC capability (strengths and weaknesses) for the respective training technology. The rating is based on a 5-point scale with semantics descriptors to assist the decision maker in grading, as seen in Table 11. These rating are used to bind the problem by providing a ROI with which to perform sensitivity analysis

Table 11 Rating of Each MEC

Rating Scale and Description of MEC	
Rating	Description
1	The student pilot is not ready to perform this MEC in a non-permissive environment or on this training technology.
2	The student pilot is ready to perform this MEC, however, he/she still needs to gain substantial amount of additional experience in this particular MEC using this training technology.
3	The student pilot is ready to perform this MEC, however, he/she still need additional experience in this particular MEC using this training technology.
4	The student pilot is ready to perform this MEC, but requires additional confidence to be gained in this particular MEC using this training technology
5	The student pilot is ready to perform this MEC and needs no additional training for this training technology.

The RSE can be either designed specifically to match the decision maker's subjective opinion or if existing response curves exist for the student using the respective training technology for the identified MEC then the RSE can be uploaded from an existing database file. The maximum boundary conditions for both K&S and MECs are saved in a database file, an example is seen in Table 12. The table clearly shows the headers and the technology references along with the maximum ranges for K&S and MEC; the digit underneath the technology reference describes the number of rows that separates the technologies, which is used for quick reference for other stages of analysis by the workflow process tool. This database exists as a summary of ratings to avoid decision makers searching through complex database files or cluttered spreadsheets and is saved within the ROSETTA 1 analysis directory.

Table 12 K&S and MEC Ranges for Mission per FoS System

K&S and MEC Ranges for Mission per FoS System per Student						
FoS	K&S Range	K&S max Ranges	MEC	MEC Importance		
DTH-24						
5	Phase of Mission	5	Assess and integrate	4		
	Time restrictions	4	Dynamic execution	3		
	Multi-tasks	4	Process and Analyse	4		
	Sorts information	5				
DT-KS						
5	Phase of Mission	5	Assess and integrate	4		
	Time restrictions	6	Dynamic execution	3		
	Multi-tasks	6	Process and Analyse	4		
	Sorts information	5				
DTCC						
5	Phase of Mission	5	Assess and integrate	3		
	Time restrictions	6	Dynamic execution	4		
	Multi-tasks	5	Process and Analyse	2		
	Sorts information	4				

Once the rating of the MEC for a particular technology has been accomplished or data has been uploaded from an existing file, the decision maker has an opportunity to amend the

shape of the response surface that describes the relationship between K&S and MEC. The rating given to K&S and MEC become the 'x' and 'y' axis respectively for the boundary of the mathematical curve, as per behaviour of 'Prepare Sample Points for RSEs ROSETTA 1' Use Case in Volume II, Figure M3. If designing a new relationship curve, by default the tool will display a linear plot with 'n' number of sample points that can be manipulated by the decision maker, which in turn will change the shape of the mathematical relationship between parameters, as seen in Figure 61, which illustrates the default relationship between the first K&S (Phase of Mission) and the first MEC (Assess and integrate Information). Obviously, the shapes of the RSE should be evolved over a period of time using experimental data, generated from SMEs, that gives a clear indication of the graphical shape the best describes the change in the training effectiveness that relates to training attributes. It is foreseeable that the optimised shape of the relationship will take a number of recursions over a number of training missions before the relationship is accepted as a true / legitimate representation between the two respective parameters.

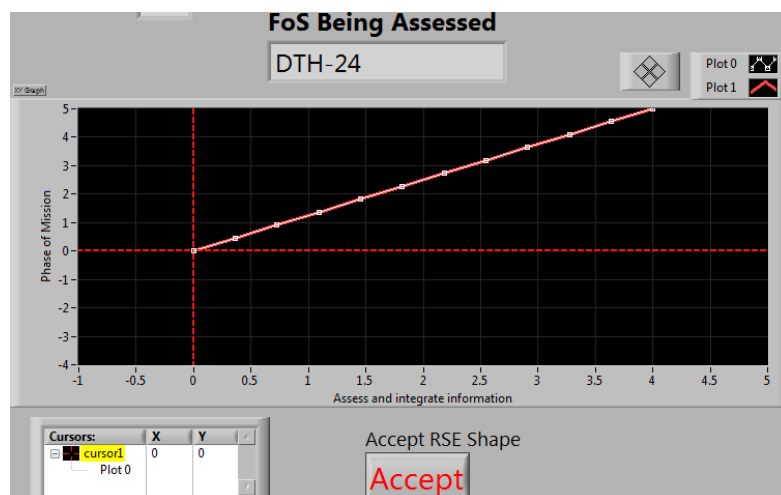
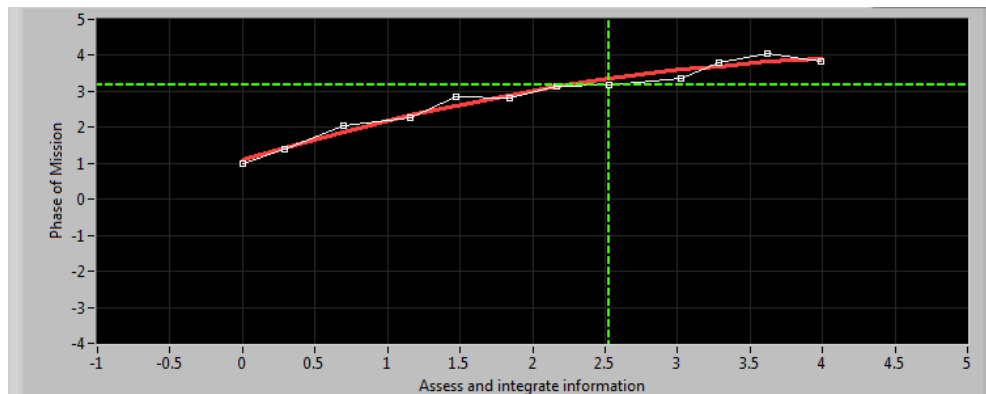


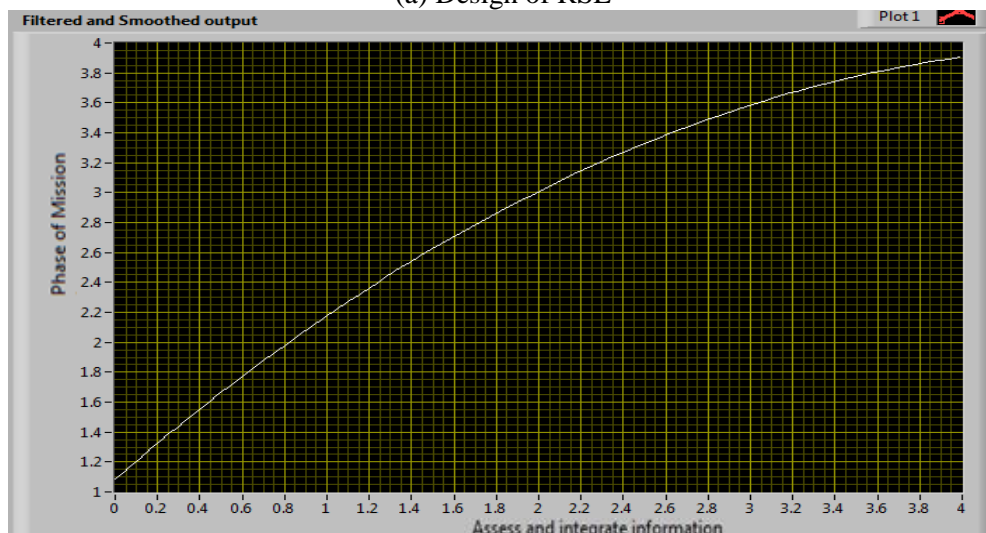
Figure 61 RSE Shape between K&S and MEC

As the shape of the relationship is based on subjective assumptions using SMEs with substantial experience within a FTS, the current method of developing the mathematical shape involves the manual manipulation of data points and then using mathematical regression, as per Chapter 5.1, to smooth an additional curve on the same graph (indicated by the red line on Figure 62a); alternative approaches to generate the RSE include: linear and non-linear optimization/fit to available data sets. Thus, from manipulation of the default data points additional curves can be formed that requires a decision of the correct shape that accurately describes the relationship between the parameters, which is necessary to be accepted and thus generates a surrogate model. To assist in this process the red line curve,

Figure 62a, is further filtered/smoothed and additional points added to give clarification of the RSE shape, as illustrated in Figure 62b. Each changed data points in the graph will automatically update the actual RSE shape in plot b; once the shape of the curve is accepted the tool will reset the RSE shape for the next MEC in the sequence and assessments of the appropriate relationship curve continues. Once all MECs have been completed for a K&S, the next K&S is displayed and the sequence is repeated; as described by the behaviour of the ‘Generate ROSETTA 1 RSE Arrays’ Use Case in Volume II, Figure M 4.



(a) Design of RSE



(b) RSE Shape for the Analysis phase

Figure 62 Creation of RSE Shape

Once all the curves have been decided for one training technology, the tool automatically transfers over to the next technology reference in the list provided by Table 12 and the identical process continues until all remain technologies have surrogate models associated with them with respect to the relationships between the K&S and MECs. The advantage of having the ability to manipulate the response curve in this way is that negative relationships or null relationship can be identified (i.e. some aspects of the MEC does not affect aspects of K&S being trained: $y=n$, where n is the level of K&S). Although the decision maker has

already identified an objective discrete level of the maximum limits of both K&S and MEC, it is more realistic to evaluate the trade-off between the two parameters to gain a viewpoint of all aspects needed for the correct level of K&S for the planned mission that the technology is capable of providing. The separation of concern with identification of surrogate models between each mission training parameter focusses the decision maker into deciding on how the K&S levels affects successfully attaining the MEC levels (required) using the blending mixes. The knowledge gained from these surrogate models can help fill the gap in understanding of how the capabilities of the technology/system can affect the student pilot achieving mission goals, furthermore, it can focus training effort in aspects of learning that are considered to be weak with respect to a particular training technology, assisted by the strength of the decided relationship between parameters.

The amount of prior experience and thus knowledge that the student pilot commands to a training event varies (Martin et. al., 2014); experienced learners can deal with complex instructional material, whereas novice learners require simplification of complex contexts to prevent information overload while learning. The average experience hours for a student pilot to obtain their wings, considering blended training, is ~305.5 hours and there is a strong believe that the level of K&S (knowledge especially) is easier to acquire as experience levels improve. For the purposes of the research, experience is seen as having an exponential effect on the ability of a pilot to attain relevant MECs, therefore, the following function is used to vectorise experience levels and automatically modify the shape the RSE based on the number of hours experience, which is gathered from the pre-study assessments.

$$E_y = \frac{e^{x * \left(1 - \frac{\left(\frac{100}{\left(\frac{305.5}{Exp} \right)} \right)}{100} \right)}}{\max_x e^{x * \left(1 - \frac{\left(\frac{100}{\left(\frac{305.5}{Exp} \right)} \right)}{100} \right)}} * 10$$

Where,

E_y is the experience vector values (limited to maximum of 10 on the scale)

x is the MEC axis values

Exp is the number of hours blended training experience the student pilot has.

The function permits the surrogate to change shape depending on the number of hours blended experience, hence, more experienced the student pilot is reduces the rate of change of the curve. In addition, as the number of experience hours increases the experience vector value constraints decrease to ensure the identification of relevant experience gained from evaluating the MEC in trade off analysis, as seen in Figure 63. If there is a significant rate of change in the experience vector to the MEC level, then it could be an indication that the training level is too difficult for the student pilot to succeed in the mission objectives using the training technology.

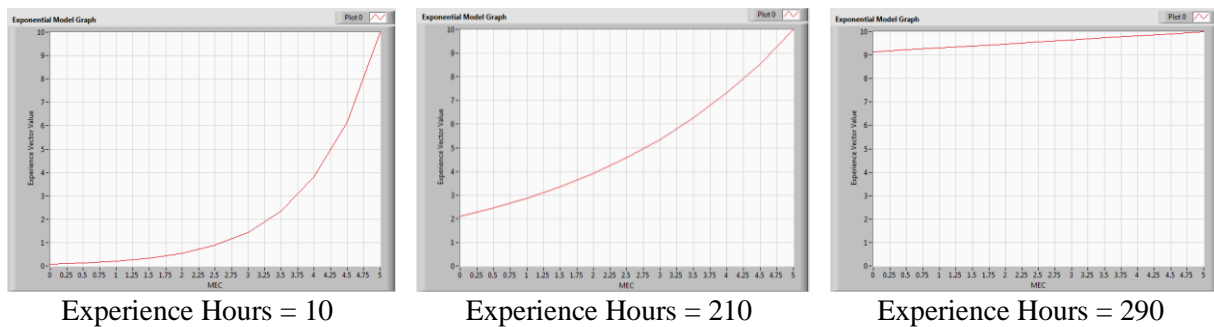


Figure 63 Identification of Relationship between Experience to MEC

There is clearly a causal link between the levels of K&S and the number of hours experience, it is expected that as experience levels increase, the minimum level of K&S set for RSE will increase; however, if the opposite occurs, according to expert judgement, it is an indication that acquisition of K&S using the training technology for the student pilot is somewhat defective. The calculation of E_y can be evolved (fine-tuned) to suit student pilot ability to attain MEC levels. The full behaviour to generate the RSEs can be found in 'Create RSE for Manipulation' Use Case in Volume II, Figure M 5.

The design and analysis arrays produced from the relationship analysis are saved in the ROSETTA 1 design and analysis directories for the respected student and an illustrative example of the design and analysis arrays are seen in Figure 64 & Figure 65. Each training technology is separated in the arrays by information regarding the number of parameters within the framework with the number of headers; with the final headers used to separate the different technologies. The 1st column signifies the specific K&S with the other columns identifying all the relevant MECs associated with the mission.

Result from ROSETTA 1 Design

DTH-24			
6.000000	0.000000		
14.000000	4.000000		
Phase of Mission	Assess and integrate information	Dynamic execution	Process and Analyse information
0.000000	0.000000	0.001504	0.000000
1.137652	0.353383	0.254135	0.353383
1.647773	0.705263	0.545455	0.723308
1.939271	1.066165	0.822556	1.084211
2.449393	1.427068	1.090909	1.436090
2.668016	1.806015	1.363636	1.806015
3.287449	2.184962	1.636364	2.148872
3.181818	2.545455	1.909091	2.545455
3.636364	2.909091	2.181818	2.909091
4.090909	3.272727	2.454545	3.272727
4.545455	3.636364	2.727273	3.636364
4.271255	4.000000	3.000000	3.971429
Time restrictions	Assess and integrate information	Dynamic execution	Process and Analyse information
0.000000	0.000000	0.022556	0.000000
0.890688	0.353383	0.263158	0.368421
Experience Hours	Assess and integrate information	Dynamic execution	Process and Analyse information
0.000535	0.000000	0.000000	0.000000
0.001430	0.363636	0.272727	0.363636
0.003824	0.727273	0.545455	0.727273
0.010226	1.090909	0.818182	1.090909
0.027345	1.454545	1.090909	1.454545
0.073125	1.818182	1.363636	1.818182
0.195548	2.181818	1.636364	2.181818
0.522926	2.545455	1.909091	2.545455
1.398385	2.909091	2.181818	2.909091
3.739499	3.272727	2.454545	3.272727
10.000000	3.636364	2.727273	3.636364
10.000000	4.000000	3.000000	4.000000
	Assess and integrate information	Dynamic execution	Process and Analyse information
DT-KS			
6.000000	0.000000		
14.000000	4.000000		
Phase of Mission	Assess and integrate information	Dynamic execution	Process and Analyse information
0.000000	0.000000	-0.007519	0.000000
1.465587	0.317293	0.299248	0.326316
1.793522	0.651128	0.569925	0.732331
2.194332	1.012030	0.795489	1.084211

Figure 64 Subset of ROSETTA 1 Design Database

Result from ROSETTA 1 Design

DTH-24			
6.000000	0.000000		
0.000000	4.000000		
Phase of Mission	Assess and integrate information	Dynamic execution	Process and Analyse information
0.320483	0.000000	0.001504	0.000000
0.393349	0.044173	0.033083	0.044173
0.465591	0.088346	0.064662	0.088346
0.537211	0.132519	0.096241	0.132519
0.608207	0.176692	0.127820	0.176692
0.678581	0.220865	0.159398	0.220865
0.748332	0.265038	0.190977	0.265038
0.817460	0.309211	0.222556	0.309211
0.885965	0.353383	0.254135	0.353383
0.957009	0.397368	0.290550	0.399624
1.027370	0.441353	0.326965	0.445865
	Assess and integrate information	Dynamic execution	Process and Analyse information
DT-KS			
6.000000	0.000000		
0.000000	4.000000		
Phase of Mission	Assess and integrate information	Dynamic execution	Process and Analyse information
0.531374	0.000000	-0.007519	0.000000
0.596712	0.039662	0.030827	0.040789
0.661580	0.079323	0.069173	0.081579
0.725981	0.118985	0.107519	0.122368
0.789912	0.158647	0.145865	0.163158
0.853375	0.198308	0.184211	0.203947
10.000000	3.818182	2.863637	3.818182
10.000000	3.863637	2.897727	3.863637
10.000000	3.909091	2.931818	3.909091
10.000000	3.954546	2.965909	3.954546
10.000000	4.000000	3.000000	4.000000

Figure 65 Subset of ROSETTA 1 Analysis Database

7.6.3 ROSETTA 1 Evaluation Process

The design data developed from the previous stage is required to be visualised within a framework that is constructed to capture the relationships between the non-functional worker oriented requirements (K&S) and the functional oriented requirements (MECs).

The requirements are denoted as members of a set of dependent variables $\{R_1, R_2, \dots, R_n\}$, and the design metrics are denoted as members of independent variables $\{m_1, m_2, \dots, m_p\}$. In the body of the ROSETTA matrix, the SMEs are identifying the strength of the relationship between R_i and given m_k ; this represents a relational transformation between requirements space and the metric space. For this $\{i, k\}$ slot, the transformation is represented by a partial derivative, $\delta R_i / \delta m_k$. The value of the solution against all requirements is evaluated against an overall weighted function (importance factor). However, the claim that the partial derivative represents the sensitivity between the two parameters is dependent on the assumption that the requirements are completely independent of each other and that the metrics are completely independent of each other; moreover in majority of practical applications this assumption is false and thus these partial derivatives are not sufficient for capturing the full sensitivities between parameters between two spaces. When this occurs the use of a roof and greenhouse is used for multivariate analysis to evaluate the mean difference between two or more dependent variables. Fortunately in this system problem, each K&S is seen as being independent to each other, likewise, each MEC are also independent although some can be consistent through the training pipeline and when one MEC finishes another begins. The resultant ROSETTA framework produced can be summarised in Figure 66, which describes the design and analysis framework which is used for evaluating trade-off solutions for each technology.

The requirements (R_i), which in this stage are the K&S parameters (Ks_n), are situated next to the greenhouse on the left; the design metrics (m_k), which are the MECs, are situated near the roof of the framework. The roof and greenhouse are used for analysing any dependencies between parameters and provides the facility for equations to be included in the sensitivity calculations. The roof can also function to provide a compensatory metric if there is a direct dependency between design parameters i.e. to account for control adjustments for CO₂ emission levels between engine and aftertreatment systems of a diesel engine: the dependency is associated with the evidence that the control structure of both can reduce the level of emissions and the trade study is which system to adjust to optimise reduction of

emissions. In the instance of training characteristics, it is reasoned that all parameters are independent of each other and the roof and greenhouse have not been included in the evaluation, consequently, concentration is given to the functionality and analysis of data using the analysis 'Q' frame, which is described in the body of the framework.

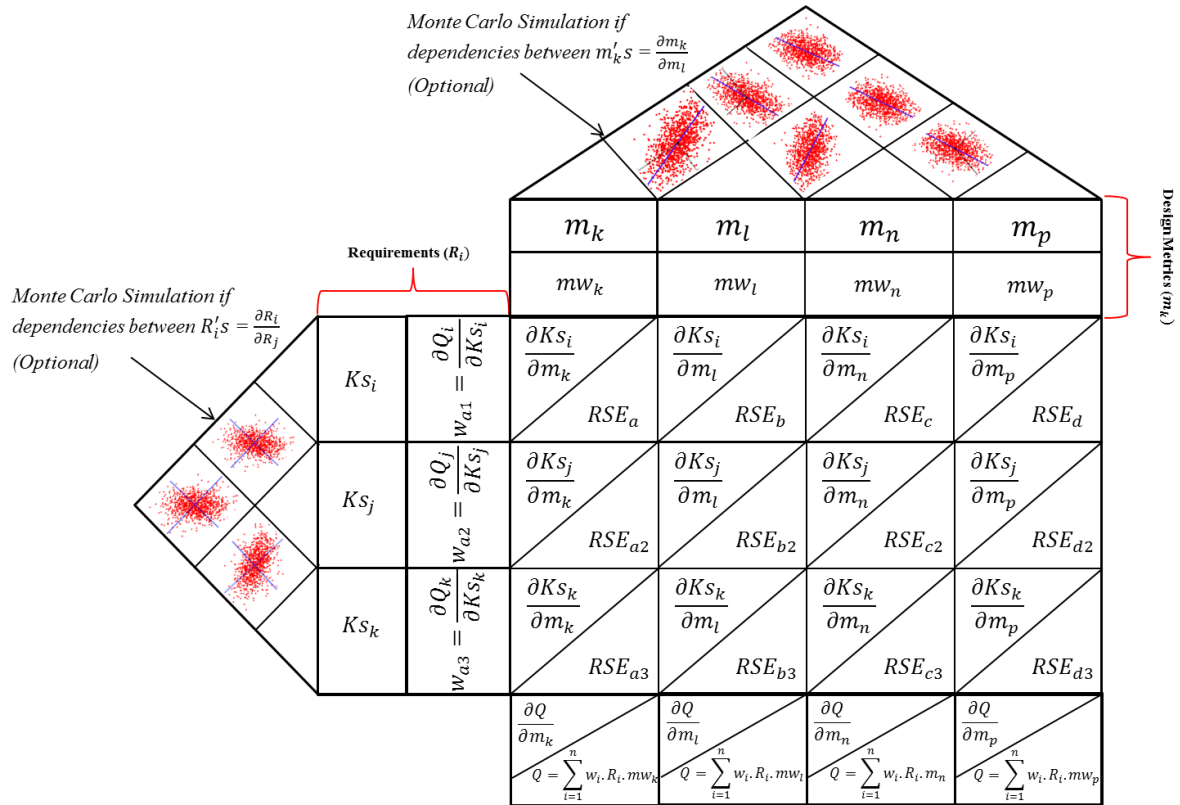


Figure 66 Summary of ROSETTA 1 Framework Analysis Method

The requirements and the design metrics have already has weights (w_i) associated with them from previous stages, however, uncertainty exists with the parameters and between them a relational structure can be developed, similar to the Q frame, where response surfaces can be created for further investigation. The RSEs, which describe the mathematical relationships between R_i and m_k , are developed and evolved within the method of Chapter 7.6.2; these are gathered from the database files saved within the ROSETTA directories (actual test/response data, if available, between a requirement and a design metric can be used within the slots of the 'Q' frame, which then become RSEs, computer resources permitting). The information gathered from the files is used the scale the framework and provide the graphical data within the slots of the framework.

The M&S environment provided by the framework produces the information required for the decision maker to visually and mathematically investigate the surrogate models in sensitivity analysis using a prediction profiler. Crosshairs in the each slot are used to select arguments

hidden in the RSE that determine the attribute value levels required for each parameter and the attribute values selected by the crosshairs in the prediction profile will determine magnitude of the slope of the partial derivatives, which gives an indication of the robustness of the training mission when it relates to the individual MECs. The process is to identify, using the decision makers experience and knowledge of both the student pilot and the technology, the optimal MEC/K&S level the technology is capable of realizing for ToT to the real world without substantial changes in the partial differential values or training attribute levels either side of the crosshair. The architecture used for ROSETTA 1 analysis and the second stage technology elimination is illustrated in Figure 67, with the behaviour described by the ‘Perform High Level ROSETTA 1 Analysis’ Use Case in Volume II, Figure M 6.

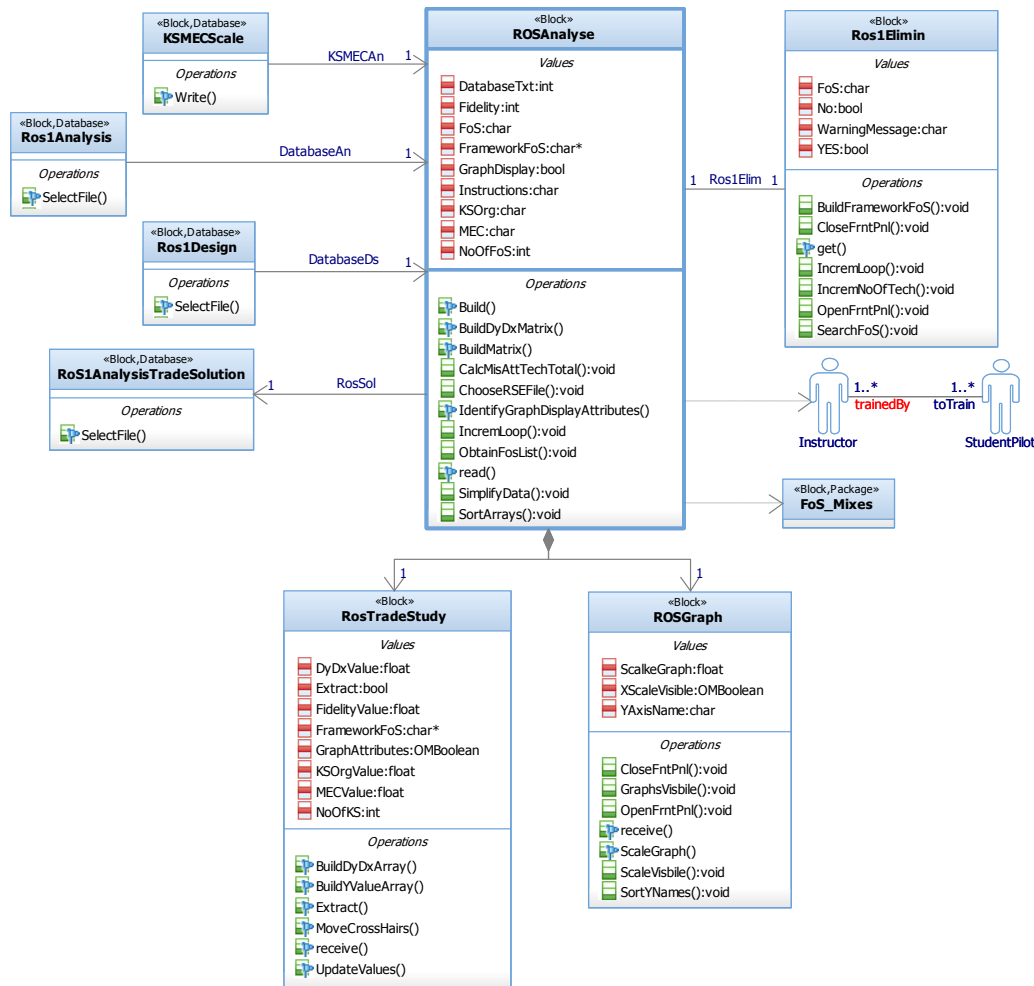


Figure 67 ROSETTA 1 Analysis Architecture

In the sensitivity analysis any movement in crosshairs will instantly update the sensitivity values and the predicted responses. Therefore, the sensitivities between a metric and requirement varies as a consequence of other assumed values, which implies that the values

cannot be represented by a single value, but are partial derivatives that are functions of the other metrics and is strongly dependent on assumptions made about the given problem and how much additional information is provided by SMEs.

An illustrative example of the ‘Q’ (analysis) frame is examined in Figure 68 concerning the DTH-24 training technology. The procedure requires the decision maker to converge on two critical parameter levels of a particular K&S and MEC and evaluate the effect on other multi-attribute values within the framework; in this case ‘Multi-Tasks’ has been considered the main requirement parameter along with ‘Assess and Integrate’ MEC, which is used to direct the trade study analysis. Aspects of MEC importance to the success of the mission is also taken into consideration for the start of the analysis. For the purpose of describing the functionality of the ROSETTA framework in the methodology, the example RSEs are shown in red and the magnitude of the partial derivatives is illustrated by the linear black line (re: main body of Figure 66), whose behaviour is described by the ‘Perform ROSETTA 1 Analysis’ Use Case in Volume II, Figure M 7.

The decision maker is examining for a robustness indication that can be visually realized by the change in slope of the partial derivative and ideally the value of the partial derivative needs to be minimal between the two critical parameters. In the example shown in Figure 68, the estimated level of ‘Multi-task’ skill required the complete the mission is ~ 2.5 (Ks_i attribute), as this level gives a clear indication that there is little change in MEC value when checking for the rate of change of the partial derivative at either side of this set-point, i.e. $(\Delta \min. (\frac{\partial R_i}{\partial M_k}))$ and not necessarily a measure of being optimised. This robustness measure is used to ensure when training to a set level of K&S, there is a degree of tolerance to action the K&S in the evaluated mission and still achieve a minimum acceptable level of both K&S and MEC when compared to the set levels.

The method of reading the framework is to identify the level of K&S required by moving the red crosshair to the desired position, i.e. the current expected training level required; and to read as follows: ‘For this level of K&S, it is expected that the student pilot will attain this level of MEC at that level of K&S using this type of training technology’. As the student pilot progresses through the training pipeline the level of K&S is expected to increase in relation to the level of MEC required for successful completion of the mission, hence, the RSEs that describe the relationship between requirement and design metrics is projected to

move upwards in correlation to the levels of K&S acceptable for the experience level of the student pilot.

This proposed method provides two options for the decision maker, the first is to permit all sensitivities to be dependent on the critical parameters, i.e. the other values of parameters, and location of the crosshairs in all the slots of the framework are dependent on the requirement value and thus control the location of set-points of all other slots; or depending on how intrinsically linked or coupled both requirements and design parameters are, each design parameter can incur an independent trade study within the same framework, i.e. if there are no dependencies between any parameter then the design can incur independent values in the trade study. It is the responsibility of the decision maker and SMEs to decide whether designs incurs a trade-off with other training attributes; it may be possible to need different levels of K&S to perform different MECs and a result each column (design metric) may not be fully dependent on the levels of K&S set for other MECs. The decision on the level of dependency within the framework can affect the trade study outcome and is seen as one of the most important choices to be made for the functionality and results of the trade study.

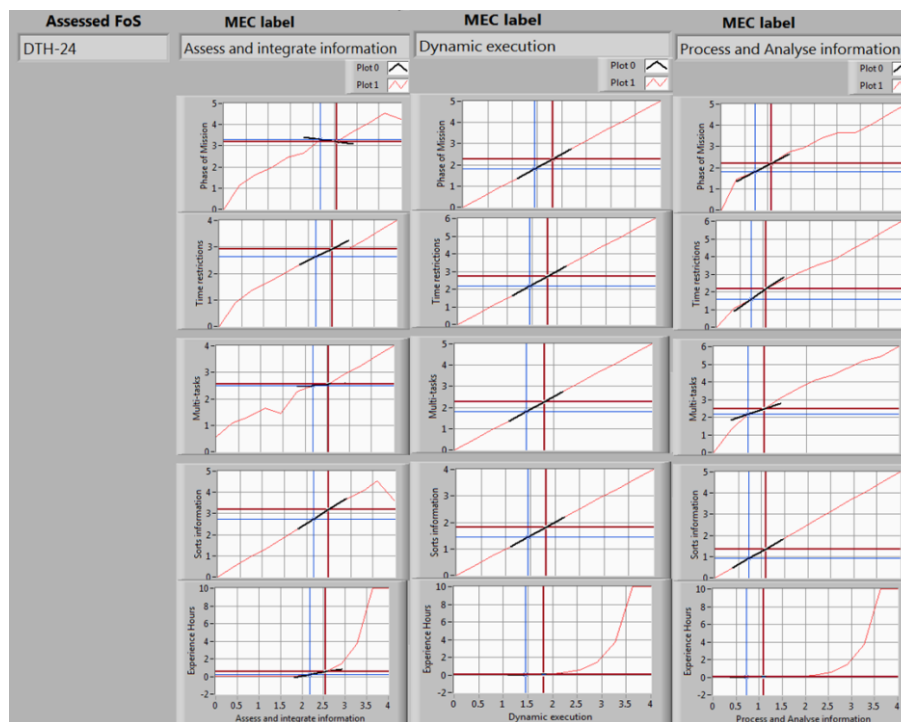


Figure 68 Example of Analysis Frame for One Technology

NOTE: The relationships between requirements and the design parameters are for illustration purposes only. These relationships need to be developed by SMEs with years of experience within the flight training domain and ideally gathered and transformed to surrogate models from current qualitative subjective assessments on performing relevant MECs with technology systems.

The visibility of the frame is dependent on the capability of the tool and the capacity of the computer RAM and speed of processor. (Current version of ROSETTA uses a commercial tool to produce the workflow process including the ROSETTA frameworks and thus uses the scrolling function to view other columns of RSEs for analysis of the partial derivatives and levels of training attributes at the set level.) The current architecture is limited to a maximum of thirty rows of RSE data due to software tool constraints (although scripting with automatic graph generation this limit can be extended), with the number of design metrics limited to the number required to perform the trade study and should be decided on using abstraction techniques. For complex designs, is advised to keep the ROSETTA frameworks as simple and usable as possible. In the current design of the framework tool, the number of slots (graphs) shown is directly dependent on the number of rows (requirements) and columns (design metrics) identified in previous stages of assessment, the program will automatically scale the UI to suit the number of relationships required to be visible. Once the trade-study for a particular training technology is complete the analysis transfers to any other training technology that remains and a new framework is produced with new sensitivities to be analysed.

The values of the requirements levels and design metrics are saved along with the partial derivative values within databases. The total quality function is found with respect to each metric, where the weights are obtained by differentiating Q with respect to each element:

$$\frac{\partial Q}{\partial m_k} = \sum_{i=1}^n \sum_{k=1}^p \left(\frac{\partial Q}{\partial K S_i} = W_{ai} \right) \cdot \frac{\partial R_i}{\partial m_k} \cdot \left(\frac{\partial m_k}{\partial m_l} \vee mw \right)$$

The equation gives a natural way to determine the sensitivities of metrics to the full set of requirements, regardless of the information content. However, the equation is oblivious to the source of the data contained within the framework, hence; qualitative estimates relying on SMEs to fill the gaps can be used for the direction of the solutions and justifications of the trades using the RSE stored within the framework. Along with the partial derivatives, it is important to use the RSEs and the attribute values to check the decision method used for concentrating on one training parameter for the basis of the trade solution, thus, the total quality function is also used to check solutions and comes in the general form similar to the standard QFD approach:

$$Q = \left(\sum_{i=1}^n (w_i \cdot R_i) \right) * mw$$

The behaviour of the evaluation using the framework is described by the ‘Perform Trade Study’ Use Case in Volume II, Figure M 8. The results of the evaluation are summarized to simplify the elimination process to enable the decision maker to visually scrutinise the analysis from the framework and a subset can be seen in Table 13. The table displays the results of all the above calculations and separates the technologies along the columns in the table. The first results shown are the total quality function values, followed by the impact of the values of the metric rankings described by using the partial derivatives equation for $\delta Q/\delta m_k$.

Table 13 ROSETTA 1 Analysis Results for Elimination of Technology

DTH-24					DT-KS			
	Assess and integrate in	Dynamic execution	Process and Analyse inf			Assess and inte	Dynamic execution	Process
Q	683.768819	512.826614	683.768819		Q	253.635351	338.180468	253.6353
	Assess and integrate in	Dynamic execution	Process and Analyse inf			Assess and inte	Dynamic execution	Process
dQ/dMk results	-10.291305	-7.718478	-9.056050		dQ/dMk results	71.861958	207.448587	75.91709

The method used for elimination of technology is similar to that used for ROSETTA 0 elimination, however concentration is needed on the values displayed. Ideally the decision maker is looking for the technology with the highest value of Q, hence the highest training value, with the lowest value of $\delta Q/\delta m_k$ to ensure robustness in training. All the MECs are still available for analysis as such six discrete values need to be assessed and the technology which incurs a low value in Q and a high value change (either in +/-) in $\delta Q/\delta m_k$ should be under scrutiny in relation to suitability for training with respect to training objectives. The technologies that are evaluated as being suitable will be tagged to the end of the saved database file that has all the partial derivative values along with the attribute level values of the requirements to be used for further stages of analysis; an example of the elimination process is seen in Figure 69.

The screenshot displays the ROSETTA 1 Elimination Function interface. It features a table titled 'Final Output for Consideration' with columns for DTH-24 and DT-KS, showing various metrics like Q and dQ/dMk results. To the right, there is a 'training choice' section with a text input field labeled 'Training Tool(s) to Further Assess'. Below this, a 'Warning Message' section asks 'Do you wish to eliminate the VRTE tool designated DTH-24 From future Analysis' with 'N/A' entered. At the bottom, there are 'YES' and 'NO' buttons.

Figure 69 ROSETTA 1 Elimination Function

The training choice of blending mix indicates that this elimination stage is not to completely decide on one training technology suitable for the mission scenario, but to give an indication

that more than one solution can be accepted for further elimination later. The warning message is visible and predominantly for future use to inform the decision maker of any issues with regards to current training technology being displayed e.g. issues with availability, faults, configuration changes, etc. The information presented to the decision maker removes clutter from the analysis phase to provide simple metricised values with which to make final decision with. The behaviour of the elimination function is described by the 'Perform Technology Elimination ROS1' Use Case in Volume II, Figure M 10; with full explanation of the behaviour and operation of ROSETTA 1 Assessment stage, please refer to Volume II, Appendix M.

7.6.4 ROSETTA 1 Summary

It might be possible to stream student pilots into different ability groups in relation to the mathematical functions, which describe the relationships that are directly dependent on the suitability of the technology to the student pilot and the specifics of the planned mission scenario. The RSE shape can then be dynamically modified to suit student pilot progress to the training technology/system as they execute the mission scenarios. In time, the type of response surfaces that describe the relationship in training effectiveness (i.e. the training technology/system(s) effectiveness at training a pilot for a given K&S attainment level) given a specific MEC set to a level that can be sustained by a training technology (moreover the type of VRTE configuration), the choice of elimination of unsuitable solutions can be more optimised given pilot's experience and learning ability. Knowledge gained from the response surfaces and the K&S level achieved from performance feedback mechanisms within a given time constraints should assist decision makers in the suitability of VRTE or technology configuration settings (e.g. layout, HOTAS, environment, no. of displays, etc.) for a given training mission.

The feedback from performance indicators and subjective evaluations from the mission can be used to clarify whether the student has achieved the estimated levels of K&S with the chosen blending mix. As training progresses the amount of tolerance allowable for the student pilot to achieve the minimum acceptable level of MEC with the training technology, which becomes more difficult as they progress through the training pipeline, reduces and a situation of large rates of change in partial derivative values is expected to occur between either sides of a specified set-point. As a result, it is believed the amount of experience the student pilot gains should ensure that progress swings to the upper right hand side of the

mathematical relationship; swings to the left hand side of the relationship can be a sign of negative ToT and be an indication that intervention is needed to discuss what aspects of are causing negative relationships to be prevalent within the ROSETTA frameworks.

Issues with the generation of the ROSETTA framework, especially in designs that incur a high degree of dependencies between requirements and design metrics involve the real time and efficient operation of the data within the frameworks. The commercial tool used for the creation of the ROSETTA framework is governed by rules dictated by the capability of the technology (computer) and the memory and processor management techniques embedded in the tool itself. Larger frameworks greater than 6-by-6 parameters, with two plots shown in each slot, uses considerable computer resources and will slow the natural flow of information updates from one column or row to others when it relates to the regeneration of a new sensitivity relationship to match the new trade study position of the crosshairs. Abstraction techniques using hierarchical representations of the system can alleviate this issue to a degree, but as more knowledge is gained about the specific components of the system that trade studies need to be performed on, the ROSETTA frameworks requirements will become more constraint when it relates to performance efficiency. Consequently, commercial tools that are designed for multiple purposes do not have sufficient capability to create large ROSETTA frameworks, thus, if the information within the frameworks is useful, for practical purposes, a specific tool needs to be developed with the capability of producing large graphical data in slots of the analysis frame, as per Figure 68, without incurring substantial use of processor or memory workload. The tool also requires allowing integration with other commercially available tools with the ability to handle textual and spreadsheet data easily and providing plug-ins to other commercial software for added functionality concerning planning, resource allocation, budgeting software, etc.

7.7 Pilot Awareness Rating Scale Model (PARS)

Further to the discussions in Chapter 4.3, the Situational Awareness Global Assessment Technique (Xvelin et. al., 2009) was designed around real-time, human-in-the-loop simulation of a military cockpit, where the simulation is stopped at random times and the pilot asked questions to determine their SA at that point in time and comparative analysis is completed with baseline answers with the difference providing an objective measure of SA. However, with random interruptions to the simulation it is difficult for the student pilot to be immersed in the simulation and thus lead to inaccurate results with respect to obtaining

information regarding ToT. Nonetheless, the method of indicating objects of interest at strategic points in time is useful for scenario planning and to permit a more accurate briefing session given to the pilot especially for accurate aviation through waypoints.

Self-rating systems gain a subjective assessment of participant SA (Naderpour et. al., 2015). The techniques are administered pre-mission execution to assess the student pilot's understanding of the objectives of the mission and what constraints in SA are preordained to be based on the chosen blending mix used; and post-trail for comparison of pilots actual experience of SA during the execution of the scenario; both of which are accomplished via a simple SA related rating scale. The advantage of this type of technique are the ease of application (efficient and low cost) and their non-intrusive nature, however, there is a danger of mistaking SA with performance and issues with poor recall from participants (Fernandez & Braarud, 2015). The PARS assessment is based on the strengths of the Situation Awareness Rating Technique (SART) (Salas & Dietz, 2011), which is generally administered post-mission; and the Crew Awareness Rating Scale (CARS) (McGuinness & Foy, 2000), which is for assessment of command and control attributes in SA and workload (Drury et. al., 2006). PARS uses measures relating to: familiarity of situation, focussing of attention, information quantity, information quality, instability of the situation, concentration of attention, complexity of the situation, variability of the situation, arousal, and spare mental capacity. PARS also integrates questions based on Endsley's three level model of SA consisting of three statements designed to elicit ratings based on ease of identification, understanding, projection (i.e. levels 1, 2, and 3 SA) (Endsley & Connors., 2008). The fourth statement investigates the participants understanding of relevant task related goals in the scenario. The workload subscale consists of four statements, designed to assess effort in relation to the projection of future states of SA related to elements in the situation.

The self-rating system for SA is attractive to use for a varied number of reasons: quick to administer, require little training, simplistic in nature; and can be structured to match certain key identifiers within mission scenario tasks. Accordingly, SA ratings may be correlated with performance outcomes that might result in recognition of periods of time when student pilot's possessed different levels of SA and hence could lead to identification of certain tasks at key times when performance outcomes was affected.

7.7.1 Student Understanding of Scenario, State of Mind and Success

The elimination of blending mixes using the frameworks for mixed method analysis bequeaths systems that can adequately train the required levels of K&S to achieve the required goals. A pre-mission brief should be used to give the student pilot the knowledge specifically bounded by mission goals and objectives, but included is information regarding the features, control, functionality and the environment of the remaining systems to give the student pilot knowledge of certain constraints with respect to usability, ergonomics and SA of the blending mixes. The student pilots use the knowledge of the tasks with information regarding the limitations of the technology to complete a self-assessment PARS form, described in Volume II, Assessment Form 5, to give the decision maker an indication of the pilot's mental context and processing of the activities involving in actioning the task activities using the technology. Under consideration is information relating to the SA dimensions to allow the decision maker to collect important information about how affective the student pilot will be in executing the mission with one or more disparate technologies. Dimensions under consideration include:

- The context and processing of perception of received information
- The measurement of understanding or comprehension
- The context and processing of possible future development and actions
- The identification of objects or operations that affects mission success
- The context and processing of task knowledge for mission success

The information gathered from these dimensions can form the purpose of keeping a positive relationship between instructor and student pilot. If the student pilot is having difficulty understanding the complexities of recognition of key actions and objects within the pre-mission brief, it is unlikely that performance will match the perception of current ability levels.

7.7.2 PARS Implementation Design

The PARS architecture is illustrated in Figure 70, which uses the information content of the PARS assessment form for the basis of identification of attributes. After the mission brief the student pilot(s) are required to complete the online assessment form that uses a Likert scale to metricise their agreement or disagreement with a particular statement. Once completed, the database is interrogated by the 'ReadDatabase' block and the question numbers and answers are converted to integer values. The 'PilotAware' block categorises the questions and answers to separate the understanding of SA and the understanding of mission related tasks

given consideration of the remaining blending mixes given in the pre-flight brief. The student database file is updated with the total percentage for PARS, percentage understanding of mission tasks and percentage understanding of SA. The behaviour of the PARS assessment is described by the ‘Obtain Student Pilot Understanding of Mission’ use Case in Volume II, Figure N 1.

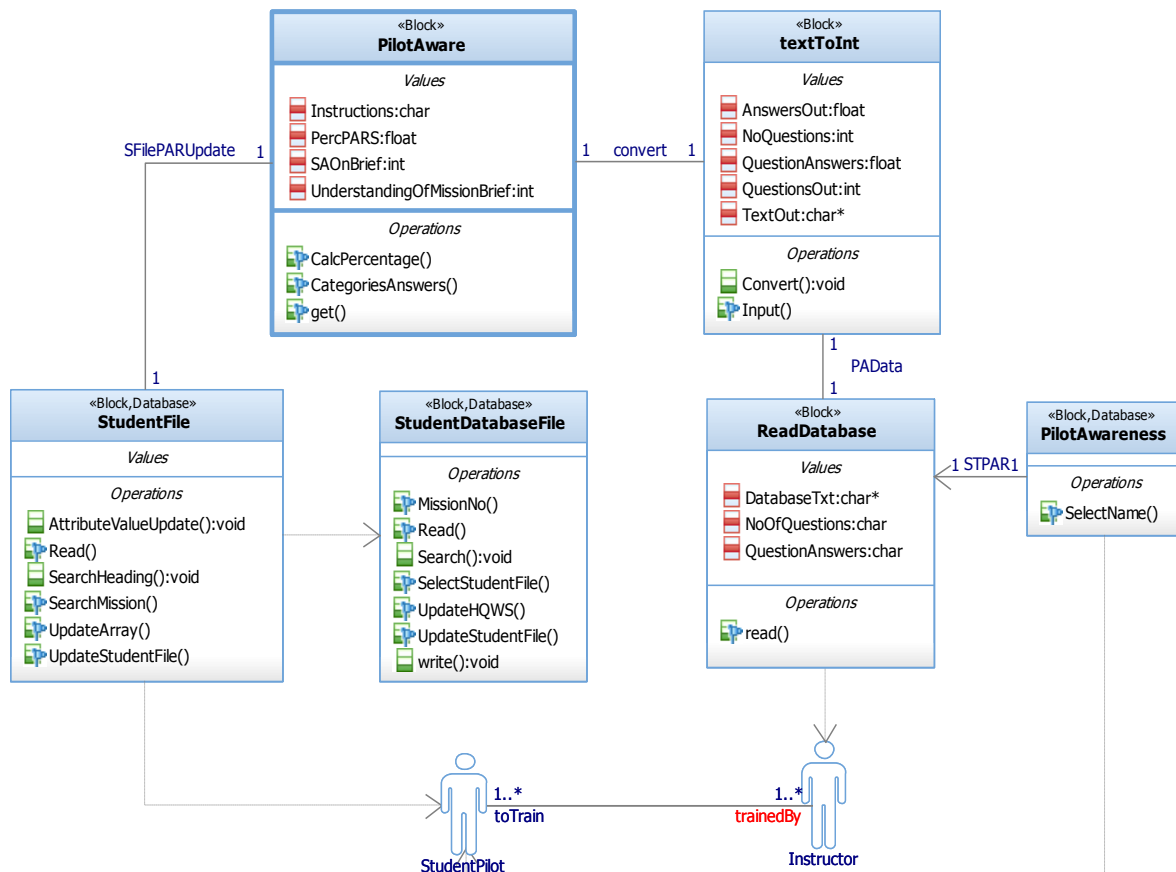


Figure 70 PARS Design Architecture

7.7.3 PARS Evaluation Process

Evaluation of the PARS questionnaire involves a 7-point rating scale ranging from 1 (difficult) to 7 (easy) depending on agreement to the statements. The student pilot answers are used to gain an indication as to the ‘comfort’ and ‘confident’ levels. The total percentage PARS is calculated first followed by SA and mission understanding, as seen in Figure 71. If the total PARS is low or likewise if the grade is high it could indicate that the student pilot has not fully understood the tasks in the mission scenario and consequently the performance outcome could be lower than expectations without confirmation of understanding. The understanding of mission scenario brief concentrates on the situations the pilot will come across in the flight and the feedback should again cover ‘middle ground’ scoring. A high

score should indicate either that the planned mission is below the pilot's ability to cope with the planned tasks or is too overconfident of completing the tasks; a low score signifies that there is possibly a lack of confidence in their ability to perform tasks in the mission using the blending mixes remaining. The SA assessment grade concentrates on various aspects of awareness when it relates to identification and understanding of using the technology for key objects that requires a change of state of the pilot and hence the aircraft for successful completion of the mission tasks. A low score could indicate that the pilot believes it will be extremely difficult to identify key objects within the simulation and the configuration of the technology makes operation of controls difficult; a high score could indicate their belief that all objects and actuators are very easily identified and manipulated; and concentration can be given to decision making tasks with the information feedback from the technology, over control actions: an indication of overconfidence or lack of knowledge with using the technology configuration and the possibility that the goals of the training mission has been misunderstood. It is perceivable, however, that as the student pilot becomes more skilled (via plasticity) and muscle memory takes precedence, high grades in the evaluation will become the norm (Hohmann & Orlick, 2014).

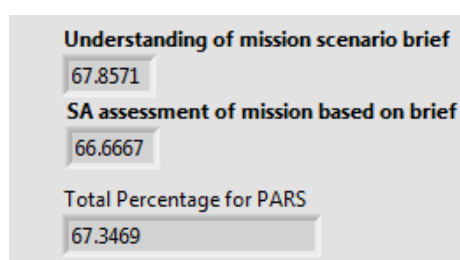


Figure 71 Calculation of PAR specific Attributes

Once information has been retrieved and the grades assessed, the workflow tool will save the new grades within the main student database file on the row of the relevant planned mission scenario for future reference.

7.7.4 PARS Summary

The PARS assessment is an efficient and necessary approach for identification of pilot's understanding and difficulty ratings based on the mission brief. The results also represent how effective the instructor's presentation and explanation skills are to pilot gaining the knowledge required for them to complete the mission scenario. As the training continues through the pipeline, the answers given should correlate to previous if the mission brief is consistent with others. A deviation in answers can give an indication that the explanation

given was not adequate for the student pilot to make an unbiased decision about the mission scenario specifics. If the grade is consistently low but performance outcomes are high, the student pilot could be under assessing his belief in his own ability and be an indication of low confidence levels and anxiety. However, an overconfident student pilot is often the most difficult to train and to transfer new knowledge to (Leveson et. al., 2009).

The findings from the assessment can provide an important distinction in the study of knowledge acquisition especially in relation to the suitability and mapping of the student pilot to blending mixes. This distinction will manifest itself in the consistent performance data of future missions and subjective observer assessments of the student pilot whilst executing the mission scenarios. It is the behaviour of the student pilot that ultimately demonstrates that their understanding is concurrent with their PARS assessment; often individuals have difficulty avoiding behaviour related issues when understanding of objectives and goals is defective and are uneasy with technology suitability. Subjective opinion of decision makers will determine how the assessment affects the student pilot progression to more difficult mission scenarios.

7.8 ROSETTA Level 2 Model

The ability to establish quantitative relationships between fidelity characteristics and training effectiveness should enhance the value of training programmes that utilize blending mixes for training. The relationship to fidelity takes a further abstraction, from that discussed in Chapters 5.2 & 7.4, by considering not only training objectives but also organisational objectives and assists the decision maker in selecting which training technology to use for each respective student to gain the necessary training levels to achieve organisational objectives, conversely, the decision is influenced by and highly dependent on blending mix numbers and availability.

Due to the trade-offs allowable for specification of technology it is vitally important to understand the fidelity dimensions in relation to training effectiveness (Graham, 2004);(ICF, 2013). To this extent, the definition of the fidelity dimensions has to be understood. Physical fidelity can be classed as the degree to which the physical simulation resembles the operational environment. Functional fidelity can be classed as the degree which the VRTE represents or relates to the actual cognitive nature of the task when performed in the real world. Psychology fidelity is the degree to how effect the VRTE is in producing the sensory and cognitive processes within the student pilot as experienced in the real world.

Furthermore, some researchers and practitioners have decomposed these dimensions into more detailed dimensions to permit further details of the architecture with respect to the fidelity dimension, for example see Table 2.

The ability to determine the legitimate amount of all fidelity characteristics and requirements necessary to achieve organisational objectives, which correlate to training mission objects, and then be capable to choose the suitable blending mix for the desired level of training would reduce the chances of conducting training with inadequate levels or excessive levels of fidelity. Addressing the issues concerning blending mix choice for individuals, requirements and training characteristics with capabilities are needed to be sought to achieve the largest benefits in training and determine the most appropriate evaluation technique for the planned training mission.

Table 14 describes the general factors to be considered for assessing the suitability of the blend to use for execution of mission scenarios. Of interest to the training organisation is saving lifecycle cost of training thus attributes relate to both time and cost. These attributes are highly dependent on the training decisions made and objectives set for each student. The variables of interest relate to the technologies capability to efficiently train the necessary levels of K&S for ToT to the real world and become parameters within the ROSETTA framework. The technology is also assessed for its ability to permit the student pilot to exhibit trained behaviour, replicate the behaviour of the real world aircraft and provide emulated feedback to allow the pilot to become immersed in the simulation to satisfy training objectives.

Table 14 Checklist of Factors to be Considered for Fidelity Trade-off Studies

General	Training methods
Time	Training Objectives
Cost	Variables of interest
Systems to be included	Behaviour of interest
External factors of interest	Procedures and skills to be considered
	Psychological & Physiological factors to be stimulated

Fidelity is generally a metric free subjective value, which can vary dramatically from one SME to the next. Giving more attention to a quantitative method of defining fidelity, determined from workshops with training experts, could establish a specific value of fidelity that incorporates the facility of configuring the fidelity settings to match the technology constraints using RSEs to describe predictive relationships between parameters of the

framework. This ensures the ability to obtain fundamental differences between comparisons of one blending mix to another based on predetermined fidelity values related to training levels and organisational objectives. The shape of the RSEs and crosshair set point in the analysis can be mapped to the performance and validity for the human perceptual system. Over time, the type of blending mix, measures evaluated, RSE shape/curve, and fidelity characteristics can be obtained that is perceived would produce a common set of values created from subjective and observer evaluations that can be stored in training databases available for future mission scenarios and students.

7.8.1 Organisational Objectives to Fidelity Characteristics of Training Tool

The structure and strategy of using blending mixes depends moreover on the organisational objectives and the decision making activity should concentrate on the top layer of management interests. The organisation analysis commences with the service they provide, which leads to the identification of how technology affects routine operations and the methods by which they are accomplished. The introduction of VE within the training programme has experienced a dramatic change in the methods used and the relationship between the GBS and the organisational objectives have yet to be fully understood. The common approach to adapt the organisation strategy is to use subjective characteristics, measures and opinions to drive the use of technology to satisfy objectives primarily based on assumptions that vary between SMEs. The ROSETTA methodology offers a means to formalize subjective opinions that can be directly associated with rules that govern the organisation. Distinguishing between training levels (job levels) and organisational objectives using mathematical relationships to the technology characteristics can overcome the problems of weak or inconsistent correlations that current flight training programmes, which integrate blended mixes, are based.

7.8.2 ROSETTA 2 Implementation Design

The behaviour of ROSETTA 2 Design can be found in ‘Obtain Sensitivities and Correlations between FoS’ Use Case in Volume II, Figure O 2. The design of the framework consists of the decision maker deciding on the level of abstraction to be used for analysis. The organisation requirements and fidelity characteristics for the technology is stored in a database illustrated in Volume II, Table IX; each of the identified parameters within the table includes the boundary limitations of the scale to be used to create the surrogate models, which describe the relationship between the requirements and design metrics. Figure 72, describes the output of the database interrogation that is used for the purpose of scaling and

plotting the linear relationships within the ROSETTA framework design sub-stage using a high-level of abstraction for real time analysis. The behaviour for retrieving database information is described by the ‘Select Level of Detail for Analysis’ Use Case in Volume II, Figure O 2.

Do you wish to perform a detailed design or a high level design for the analysis phase.

☐ Detailed

☒ High-Level

OK

organisational elements		
Practice emergency	0	10
Ease of Distributed	0	10
Availability	10	100
Affect on Attention	1	10
Affect of Situational	1	20
Affect on Salience	1	20
Direct Operating Cost	1	100

Training Device Fidelity		
Visual	10	100
Physical Fidelity	20	100
Functional	10	100
Task Fidelity	10	100
Psychological	0	100
Verbal Fidelity	0	100

(a)Choice of detail

(b)Framework Parameters & ROI values

Figure 72 Demonstration of High-Level Abstraction Scales from Fidelity Database File

The framework requires information regarding the remaining technology references to be used for the final stage of elimination: the RoS1AnalysisTradeSolution database is interrogated for a list. The design reuses the ‘DesignLoop’ and ‘DesignSlot’ blocks from ROSETTA 1 design sub-stage and repeats the behaviour of ROSETTA 1 when it comes to the development or evolution of the RSE shape between each requirement and each design metric (fidelity characteristic). The design sub-stage outputs data to two databases, one that describes the original RSE shape used to design the mathematical function; the other is the analysis RSE shape that is the mathematical regression with additional sample point added within the RSE for more detailed analysis. The architecture of ROSETTA 2 Design can be seen in Figure 73.

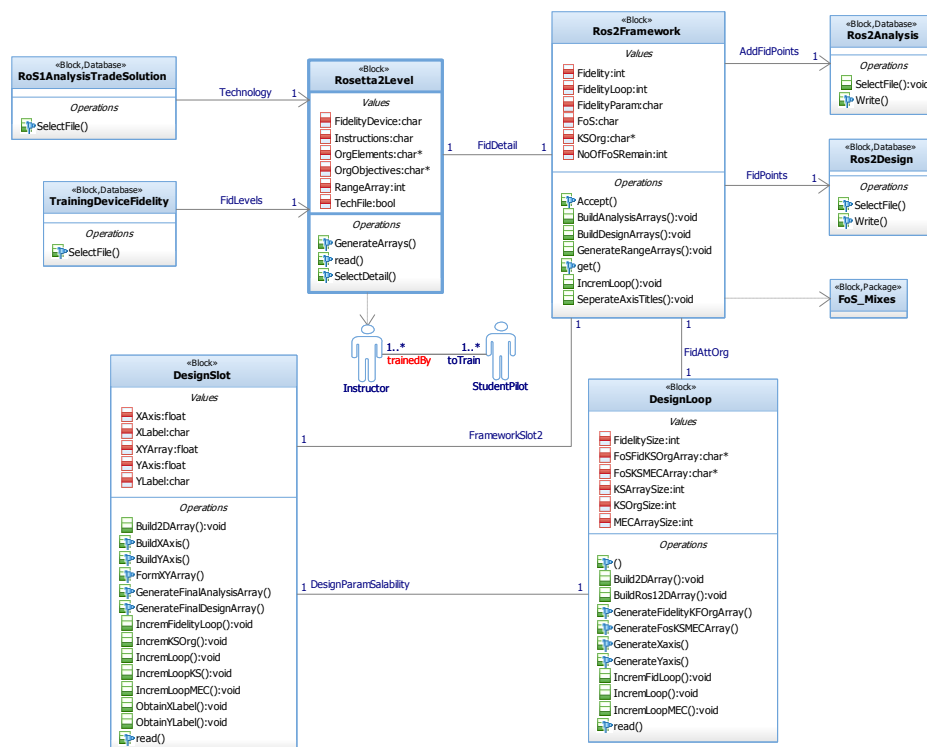


Figure 73 ROSETTA 2 Design Architecture

The K&S identified from ROSETA 0 takes on a different dimension within this stage of analysis; concentration is directed at the relationship between K&S to the organisational objectives and therefore has to relate to the training goals of the organisation. The knowledge dimension and associated RSE shape is assessed in relation to the effect of the fidelity dimension to the body of information applied directly to the performance expectations relating to the planned mission scenario of a student pilot executing the required tasks. The stages of experience on the knowledge dimension range from 1(weak) to 10(outstanding) and can be associated with semantic description as described in Table 15.

Table 15 Stages of Expertise for the Knowledge Dimension

Knowledge Level Needed for Mission Scenario	Grade	Range
Weak	1%-24%	1.0-2.4
Limited	25%-39%	2.5-3.9
Adequate or Satisfactory	40%-49%	4.0-4.9
Acceptable or Competent	50%-59%	5.0-5.9
Good or Commendable	60%-69%	6.0-6.9
Excellent	70%-79%	7.0-7.9
Outstanding	80%-100%	8.0-10

The skill dimension and associated RSE shape is assessed in relation to the effect of the fidelity dimension to the performance expectations to the planned mission scenario of the student pilot in cognitive functioning and physical movement in executing tasks in the training technology. The stages of experience of the skill dimension have a range from 1(novice) to 10(expert) and is associated with semantic descriptions described in Table 16. For further details on the explanation of all parameters that can be used within the framework please refer to Volume II, Appendix O for full description and guidelines.

Table 16 Stages of Expertise for the Skills Dimension

Skill Level Needed for Mission Scenario	Grade	Range
Novice	1%-25%	1.0-2.5
Capable	26%-39%	2.6-3.9
Skilled	40%-49%	4.0-4.9
Advanced	50%-59%	5.0-5.9
Proficient	60%-69%	6.0-6.9
Expert	70%-100%	7.0-10

The technology fidelity characteristics relationship to the objective requirements is assessed to support capability based trade-off decisions for selecting optimal flight training mixes. The specifics of the blending mix are assessed with the ability to trade-off decisions and obtain relationships between each objective requirement and a specific technology fidelity characteristic. Levels of performance of each technology can be used for trade studies using

The calculation of the total quality function follows the same method as that used in the calculation for ROSETTA 1. The structure of ROSETTA 2 Framework is seen in Figure 75. The requirements (R_i) include the new dimension(s) of K&S together with various organisational objectives relating to the efficient administration of training objectives; the requirements are recognised as not experiencing any dependencies between them; full descriptions are given in ‘Organisational Design Variables’ in Volume II, Appendix O.

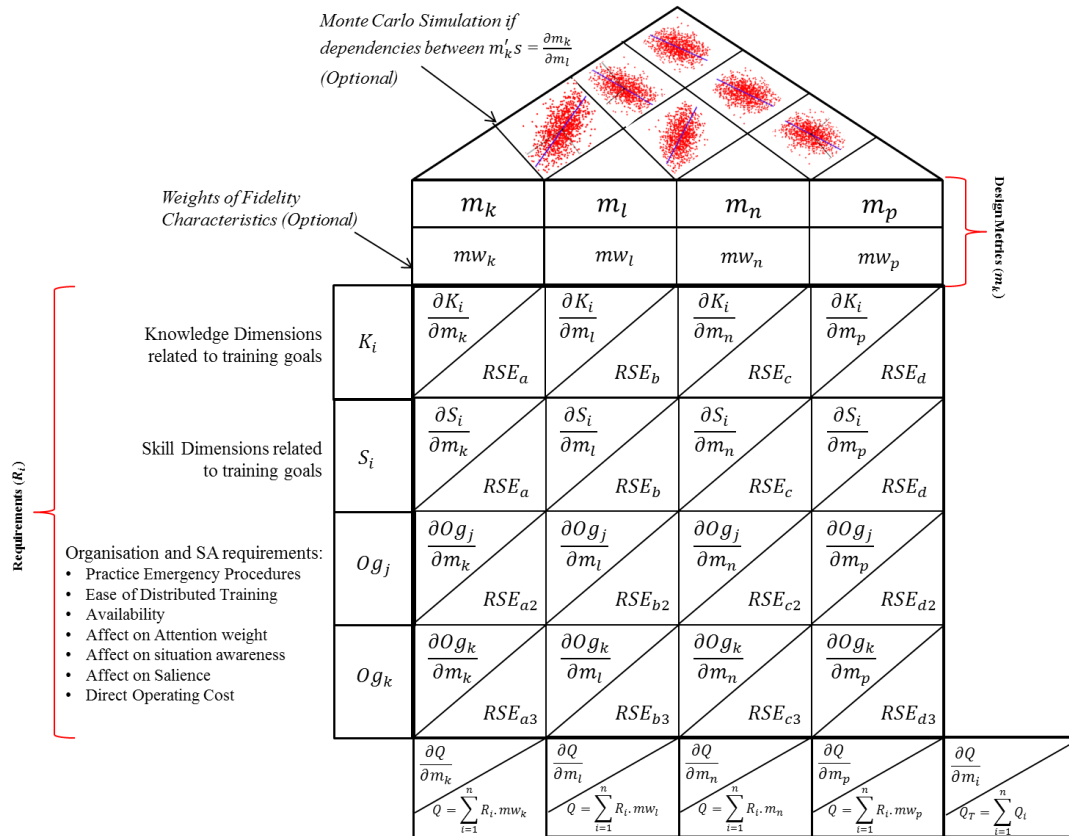


Figure 75 Summary of ROSETTA 1 Framework Analysis Method

The design metrics (m_k) are the fidelity characteristics of the technology and the decision of placing weights to fidelity is a choice dependent on the specifics of the mission scenario. The design metrics can have dependencies between them; the level of detail needed for the analysis will determine if the Monte-Carlo simulations are necessary, however, in real time situations it is deemed that this is excessive for an indication of technology suitability to training, full descriptions of fidelity characteristics can be found in ‘Technology Viewpoint’ in Volume II, Appendix O. Time in a student pilot’s day, technology fidelity and availability along with pilots experience can be considered as attribute values that can change to explore different trade-offs. The information content stored within the surrogate models of the

framework should consider availability in terms of both pilot time and aircraft availability and can be limited to repetitions in certain blending mixes.

The graphical representation of the framework is displayed to the decision maker who uses the crosshairs to perform trade-studies on each remaining technology independently. The total quality function is found with respect to each requirement as follows:

$$\frac{\partial Q}{\partial m_k} = \sum_{i=1}^n \sum_{k=1}^p \frac{\partial R_i}{\partial m_k} \cdot \left(\frac{\partial m_k}{\partial m_l} \vee mw \right)$$

The equation gives the sensitivities of metrics to the full set of requirements, regardless of information content. The source data contained within the RSEs is also used to calculate the quality function, similar to the QFD approach as follows:

$$Q = \sum_{i=1}^n R_i \cdot mw$$

The results of both equations are used to generate the aggregate total quality function significant to total suitability and robustness metrics of a technology using the equations below:

$$\frac{\partial Q}{\partial m_i} = \sum_{i=1}^p \frac{\partial Q_p}{\partial m_k}, \text{ and } Q_T = \sum_{i=1}^p Q_i$$

Each result is summarized in a simple table similar to the one described in Table 13 of ROSETTA 1 where the decision maker concentrates on the maximum value of Q_T and the minimum value of $\frac{\partial Q}{\partial m_i}$ with which to base decisions on. The decision maker uses the identical procedure to perform the trade study; after accepting the positions of the crosshairs within the framework the data is simplified for elimination choice, as per ROSETTA 1; the behaviour is described by the ‘Attain Fidelity Value of Systems’ Use Case in Volume II, Figure O 6. However, the elimination stage should display a message to the decision maker indicating that only one system can be accepted for the mission scenario, as per Figure 76. Once the ‘No’ function is selected the technology reference number is saved within the respective student database file as the chosen blending mix with which to execute the mission; the behaviour of the elimination function is described by ‘Procure Effective Blending Training Mix’ Use Case in Volume II, Figure O 7.

The screenshot shows a software interface for a training tool. It features a title bar labeled 'training choice' and a subtitle 'Training Tool to Execute Mission Scenario'. The main content area displays a warning message: 'Do you wish to eliminate the VRTE tool designated DTH-24 From future Analysis One System of the FoS ONLY to be chosen for execution of the Mission Scenario'. Below the message are two buttons, 'YES' and 'NO', for user response.

Figure 76 Elimination Warning Messages for Decision Maker

For a full explanation of the behaviour and operation of ROSETTA 2 assessment stage, please refer to Volume II, Appendix O.

7.8.4 ROSETTA 2 Summary

The ability to establish quantitative relationships between training effectiveness and fidelity should enhance the value of training with different training technology. The predictive models provided by the RSEs give the training organisers and designers of VRTE and aircraft technologies the ability to compare the identical type of assessment concerning the impact of various levels of fidelity on training effectiveness prior to using or investing in any new training technology; this can give an early indication of suitability to training.

The ability to choose the level of abstraction with which to perform the design and analysis at this stage provides added flexibility within the evaluation process. The more detailed analysis of fidelity characteristics can be performed to isolate any potential issues with the use of the technology for a particular training level, in addition, can be used as the basis to assess future training technology. The high-level analysis permit real time assessment of the choice of blending mix with little technical knowledge on the system, signifying that training the decision maker for this level of assessment should be cost effective and efficient; the more detailed assessment can be reserved for SMEs with a high level of technical and pilot training knowledge.

The record of blending mixes stored with the student file allows for assessment of performance levels using such technology to be evaluated separately and thus can assist in identifying technology that student pilots are having difficulty with. The inclusion of salience and attention weight gives strength to the ability of the technology to display relevant objects during the mission scenario that the student pilot has to recognize and use SoPs to change state of the aircraft. The direct operating cost is a key parameter under consideration to the level of K&S needs; the trade-study is directed at organisational objectives and it makes sense to choose technology that provides the correct levels of training for less cost but with sufficient capability for the other dimensions under consideration.

7.9 Task Load Model

Task load is defined as the mental workload inflicted by the technology to be controlled by the pilot (Bennekorn & Tuinen, 2010), although the definition of task load varies among SMEs. Workload increases by the time criticality of the tasks required for the pilot to perform. Nevertheless, measuring task load during execution does not give a direct prediction of how the student pilot will handle each task in the mission, as each task may have different numbers of activities involved and different time constraints with which to perform them. The types of tasks and activities involved to perform the mission task are seen as important determinants of workload experiences of pilots. Task load is influenced by a number of factors:

- Energy rate demand – whether or not mission or task constraints can be met; this is more apparent to the pilot when pressure to meet the constraints continues into the final part of the task.
- Time Available – Actions which need to be performed within time constraints (directly related to distance available in the task), which becomes more important as the time available reduces.
- Final velocity and intercept angle – constraints required by the mission requirements with respect to waypoint intercept.

Predictions of task load are heavily influenced by operational factors, such as small deviations between the actual aircraft thrust, modelled aircraft thrust and pilot behaviour (Gillet et. al., 2010). There is a strong bias towards developing pilot models to be used within M&S to predict per timed sample the task load a pilot will have to manage during execution of a task. However, deviations from SoP by pilots can completely overshadow the M&S predictions due to the inherent human factor issues for pilot actions, as discussed in Chapter 4.4.

7.9.1 Task Load Implementation Design

Before the decision maker makes a conscious decision on the tolerance set-point of accuracy for activities in the tasks, a general indication task load is needed with which to base measurements. Attributes under consideration include:

- $i = V$ (visual), A (auditory), C (Cognitive), or M (Motor)
- $s =$ Flight Procedures (approach, Landing, Scan Patterns, Checklists, and others)

- n = Number of accomplished activities or tasks
- $TL_{i,s,n}$ = task load of each task activity
- $\Delta t_{i,s,n}$ = duration of each task activity (sec)
- t_{total} = flight time (sec)

The calculation of task load, therefore, requires knowledge of the duration of tasks and the workload dimensions associated with each task with which to base calculation on. The baseline scenario, described in Volume II Appendix H, gives the decision maker the ability to upload the baseline mission and use the simulation values to calculate task load. The architecture for the task load stage can be seen in Figure 77 with the behaviour described by the ‘Assess Task Load for Each Task within the Mission’ Use Case in Volume II, Figure P 1.

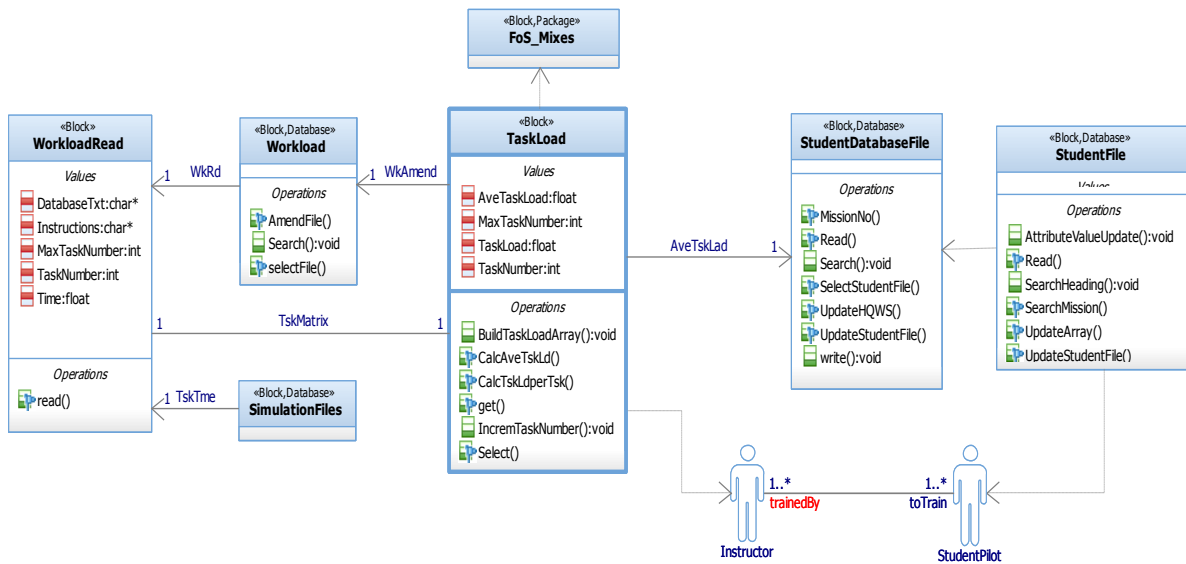


Figure 77 Task Load architecture

The ‘TaskLoad’ block requires information from the workload stage of Chapter 7.5, which gives the estimated workload values for ‘i’ in each dimension. The simulation database for the planned mission is interrogated to obtain both $\Delta t_{i,s,n}$ and t_{total} for the mission scenario. The results of task load calculation are then saved in the ‘TaskNoandDetails’ handling qualities database file for the mission within the Student_Pilots directory; and the total task load for the whole mission is saved within the main student database file.

7.9.2 Task Load Evaluation

The task load is calculated by first considering the total time available for the task (or activity) along with the total time for the mission to calculate τ :

$$\tau = \frac{\text{time available for task } (\Delta t_{i,s,n})}{\text{Total Flight time for Mission } (t_{total})}$$

This calculation gives a fraction of time in the mission with which to perform each task. This calculation is used within the workload values to calculate task load as follows:

$$TL_{i,s,n} = \frac{\% \text{Workload per task} * (1 - \tau) - \% \text{average Workload per task}}{\tau}$$

Task load values ≤ 45 gives a clear indication of adequate time available for the pilot to perform the mental tasks or activities needed without over exertion, however, task load values ≥ 60 gives an indication that the number of tasks or activities to perform for the successful completion of the mission tasks has considerable time limitations and this increases the amount of mental workload processes that has to be maintained during the task. The planned mission tasks should be directed to present to the student pilot task load level ranges between 46-59 to ensure mental process overload conditions do not occur in early stages of training. However, as the student pilot progresses through the training pipeline and the mission tasks increase in difficulty it is expected that the pilot has adequate experience and proficiency to manage tasks or activities with greater task load values.

For the subset of tasks described throughout this chapter, the workflow tool outputs the evaluation result and ‘tags’ it to the ‘TaskNoandDetails’ file stored within the student directory, as seen in Table 17. This small database file created from the workload assessment stage is used to store useful information regarding the instructor’s opinion of how the student will cope with the tasks. In this proof for functional context example, the task load value for task 1 scores higher even though the task description expresses minimal control required for the aircraft; however, upon further investigation the timescale needed to complete the tasks and activities produce a task load greater than the bank turn. This information is important to consider when predicting performance success of each student; from the results more mental workload is required for monitoring avionic systems and the environment to ensure course (or target) is maintained in the first task due to time constraints than in the second even though the number of activities and operations in the second manoeuvre is greater.

Table 17 Example of Output from Task Load Assessment

TaskNoandDetails Student Directory File			
Mission Scenar	Oscar123		
Task Number	Task Description	Pilot Specific Rating	Task Load
1.000000	Maintain height and heading	52.5741	53.1699
2.000000	begin bank turn at 30Deg	60.9074	46.135

The total average task load for all tasks within the mission is saved within the main student database file for quick reference.

7.9.3 Task Load Summary

Although workload has been assessed in a previous stage the effects of workload pressure being instigated on the student pilot with respect to time was not considered. The task load assessment fills the gap in workload measures by using the knowledge gained from the workload assessment and adding a time dimension to workload considerations. The outputs give an indication that simple tasks with strict time constraints can affect the mental processes of the student pilot, which were previously ignored. The task load value can affect how the decision maker predicts performance and the maximum tolerance acceptable in flight manoeuvres or targeting operation/task accuracies and precision using the chosen blending mix. For example: if a task within the mission scenario requires substantial control adjustments in a minimum amount of time and the chosen technology characteristics and configuration is not best suited to distinguishing between operations (buttons/switches), this can incur additional workload for student pilot to ensure the correct action has been executed and thus affect timely completion and accuracy of manoeuvres and objectives within mission tasks.

7.10 Prediction of Task Performance with Chosen Training Tool

“Prediction is very difficult, especially if it’s about the future.” – Anonymous

Most human performance measurement data informs of the degree to which users accept and believe they perform with equipment and technology; the measurement gives an indication of good design practices to integrate HITL. A measurement of task performance, however, suggests how effective and efficient the particular human machine interface (HMI) to suit a particular user has been. The primary utility of task performance data is to influence system

design throughout the development process; if the scenario is known and data from previous executed mission scenarios is available, decision makers can concentrate upon design characteristics that influence performance of certain tasks.

The evaluation of success of the choice of blending mix, and generic training SoPs, is revealed in the performance data that could indicate design characteristics that degrade performance, as such could denote areas to which redesign or technology configuration changes can improve human task performance and thus total system performance, as briefly discussed in Chapter 4.4.2. The knowledge gained from the performance database provides a useful basis for decision makers in determining performance constraints and training success. Inputs to M&S and mission simulation exercises can be used to predict future performance results and decide on future selection of training tasks using disparate blending mixes.

The measurement tool for performance based measures should exhibit a reliable effect regardless of sample size, the subjective assessments and the mathematical formalisms used to predict performance has to concentrate on individual success rates rather than group.

7.10.1 Performance Prediction Implementation Design

Accuracy of activities or tasks with the FTS is generally based on observer led assessments in qualitative terms such as ‘good’, ‘expected’, ‘unsatisfactory, and so on. However, for prediction of accuracy quantitative measurements are needed to give confidence quantities that the aircraft (virtual or real) can be controlled within certain margins of error from the ideal. The question of ‘We are confident that using the technology interfaces the aircraft can be maintained within ___degree of ideal using the chosen blending mix” is required to be asked (re: Figure 12). The predictive measurements of interest with answering this question are accuracy and precision of the tasks or activities. To assist the decision maker in realizing a valid prediction of flight precision, data is gathered regarding the understanding of SA from the PARS assessments along with PSR and task load to calculate the default standard deviation needed to generate Gaussian curves for the task activities of the mission: time accuracy, altitude accuracy, heading/position accuracy, and others if necessary and when required. The predicted accuracy refers to how close performance is believed to be from the predicted mean, whereas the precision refers to how consistent performance is to the mean, i.e. able to maintain control of aircraft to planned mission requirements, as seen in Figure 78.

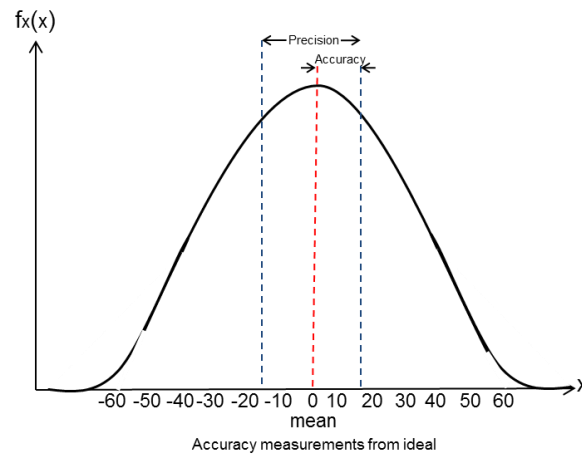


Figure 78 Graph Illustrating difference Between Accuracy and Precision

Each task of the mission is evaluated with the decision maker having an opportunity to vary the precision levels to incorporate his/her own views on the ability of the pilot using the chosen blending mix to maintain mission requirements. The generated curves are then saved within the distribution files database for further analysis with actual performance data from the executed mission scenario, the architecture used for this assessment stage is described in Figure 79; with the behaviour described by the ‘Estimate MoP of Students’ use Case in Volume II, Figure Q 1.

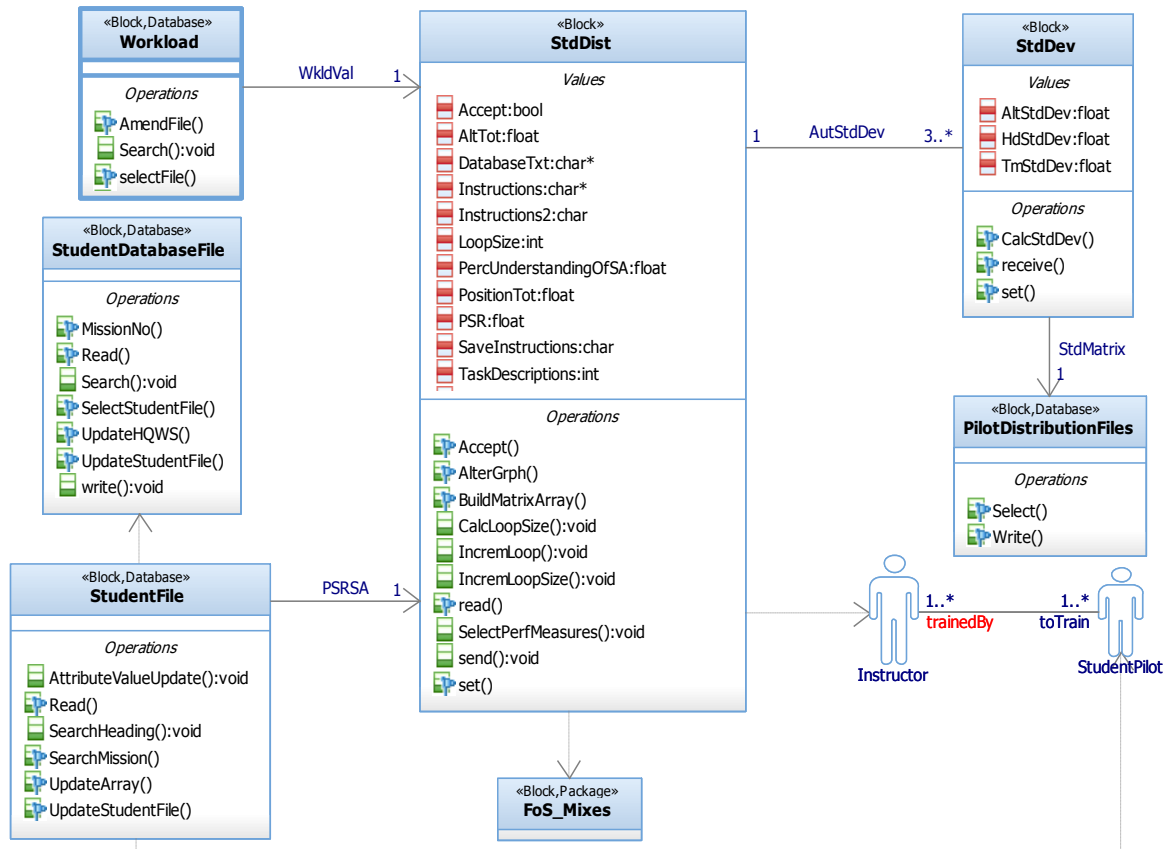


Figure 79 Prediction of Accuracy and Precision Architecture

7.10.2 Performance Evaluation Process

The default standard deviations are gathered from SME workshops as being an adequate set-point for precision flying to achieve task goals, the data gathered from the main student database file and the workload files are then used to amend the standard deviation to more accurately predict tolerance settings gathered from previous assessments. These set points can be amended to suit experience gained with the chosen blending mix, thus, the new StdDevOut value can be used to amend the default standard deviation settings for all accuracy dimensions being predicted. The results for the subset of Oscar123 mission scenario example used for this proof of functional concept are illustrated in Figure 80, which describes the Gaussian distributions for accuracy and precision for the first task. In this example, for time accuracy it is perceived that the student pilot is able to maintain SA enough to allow them to fly the aircraft to waypoints within ± 1.5 seconds, at an altitude of ± 17 feet and maintaining heading between ± 7 degrees from planned – these default values can be evolved as the training programme progresses. Furthermore, the decision maker has a number of functions that can scale the precision settings of each accuracy dimension that will amend the shape of the Gaussian curve and respective array values to suit the subjective opinion of the abilities of the student pilot using the chosen blending mix, hence, adjust the precision and accuracy set-points to suit judgement.

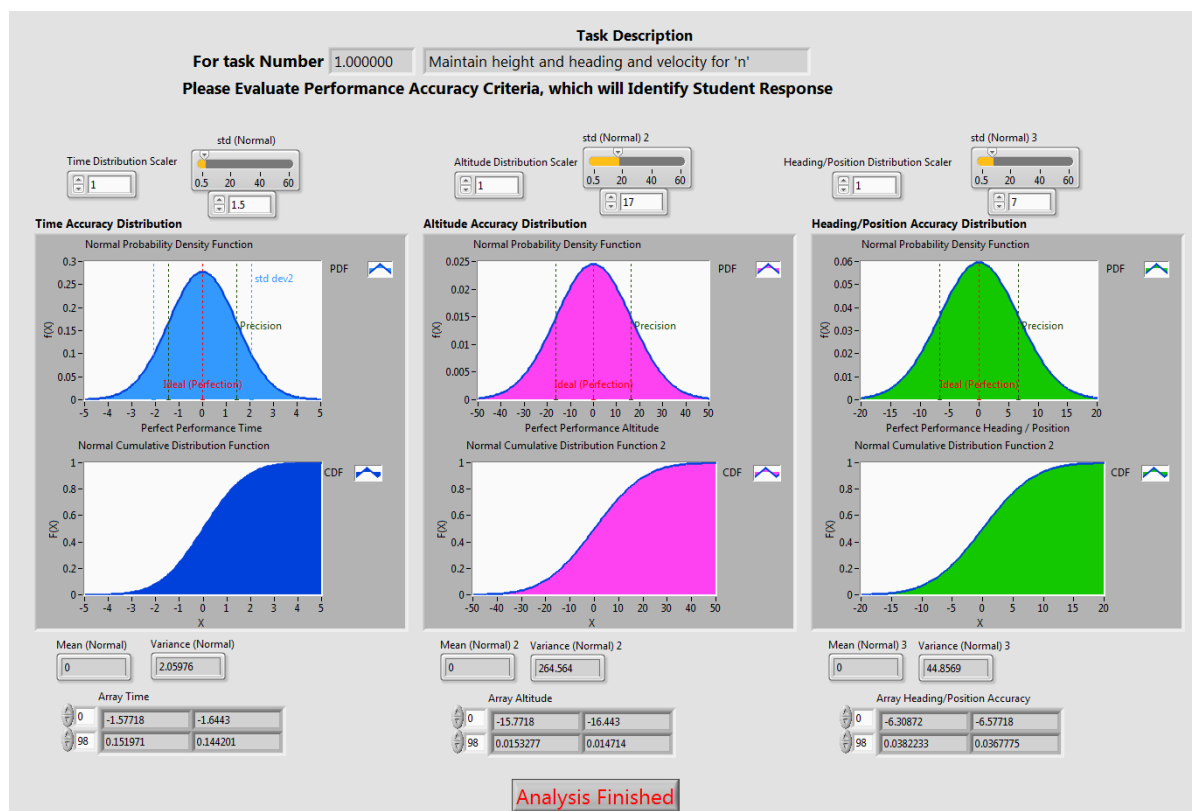


Figure 80 Accuracy Predictions for Subset of Oscar123

Once the decision maker is happy with the accuracy settings of the task, the ‘Analysis Finished’ function is selected, which will transfer to the distributions for the next task within the mission and store the array values along with precision and accuracy values within a 2-D array. The same process of analysis of precision settings occurs until all the tasks within the mission have been assessed, as seen in Table 18. The data stored within the table is to enable comparison of distributions from one mission to the next to check the direction of improvement i.e. the precision tolerances should decrease as experience is improved and with it the acquisition of greater levels of K&S.

Table 18 Example of Results from the Sub-set of Accuracy Predictions

Additional Assessment Results												
Task Number	1.000000											
std Deviation	1.935185	21.932099	9.030864									
timex	5.000000	4.932886	4.865772	4.798658	4.731544	4.664430	4.597315	4.530201	4.463087	4.395973	4.328859	4.2617
timey	0.007321	0.008003	0.008737	0.009528	0.010377	0.011288	0.012265	0.013311	0.014428	0.015620	0.016890	0.0182
Altitudex	50.000000	49.328859	48.657718	47.986577	47.315436	46.644295	45.973154	45.302013	44.630872	43.959732	43.288591	42.617
Altitudey	0.001353	0.001450	0.001552	0.001661	0.001775	0.001895	0.002022	0.002155	0.002294	0.002440	0.002594	0.0027
Headingx	20.000000	19.731544	19.463087	19.194631	18.926174	18.657718	18.389262	18.120805	17.852349	17.583893	17.315436	17.046
Headingy	0.003803	0.004060	0.004331	0.004615	0.004914	0.005228	0.005557	0.005901	0.006261	0.006637	0.007029	0.0074
Task Number	2.000000											

The row with heading Std Deviation in Table 18 indicate the predicted accuracy tolerances for the task for the respective student pilot using the chosen blending mix, which will be used to score the pilot performance accuracy to the set mission scenario using the baseline scenario developed in Chapter 7.1. As such, each planned task should be scored independently of others, consequently, the performance accuracy from one task will not directly affect the score for the second, although it is expected that any deviation in course and time will be corrected by during execution of the mission. The predicted outcome can be used for preliminary attribute values of CC for each task output and therefore lead to a predicted EC_{out} grade for each student (re: Pipeline model of Figure 40).

7.10.3 Performance Summary

The prediction of performance is a difficult task to accomplish and is open to considerable bias. However, the proposed workflow tool described in this methodology presents information to the decision maker in an easy to understand visual manner, which will store the information contained within the distribution curves within a database that can then be compared to actual performance data from the chosen blending mix (training technology/system). This ensures that prediction moves from being an undocumented subjective assessment to something that can be used to document progression that incorporate the judgement of the decision maker on the performance outcomes using the chosen blending

mix. This type of documentation can assist in removing some of the biases that might exist by forcing the decision maker to justify their predictions based on previous performance levels and their own understanding on the blending mix suitability to mission tasks and the respective student pilot.

This type of distribution used for the accuracy dimension can also be used in workshops to debate the usefulness of the training technology in relation to ergonomic design, visual acuity and operator comfort. If a training technology shows consistent increase in tolerance levels, based on performance feedback scores from predicted tolerances, then a discussion needs to occur on whether: continued training on the technology is beneficial; if tolerance levels have to be increased when using this technology configuration; or whether concentration needs to be focussed on other aspects of performance outcomes relating to acquisition of K&S rather than accuracy and precision in student pilot operations and subsequent behaviour within the execution of the mission.

7.11 Pre-Flight SA Assessments

In this method, the student pilot is asked to complete a self-rating SA questionnaire before the decision maker completes the last subjective assessment to predict task and mission success. The self-rating technique is used due to its non-intrusive nature and to compare thoughts and feelings with post-trial assessments; the technique requires very little training to complete as it is designed in an easy to understand language. The Pre-flight SA assessment should be administered via an online form whereby the pilots rate their own SA based on their current state of mind. Information regarding the previous night sleep, how their day is proceeding and their opinions about the upcoming mission are integral to estimating the success before the student pilot is evaluated; the assessment form can be seen in Volume II, Appendix R, Assessment Form 6. These SA ratings are used to correlate performance along with success of mission task goals with the decision makers predicted responses.

7.11.1 Pre-Flight Implementation Design and Evaluation

The data for each student pilot is stored within the PreFlightSA database that used a Likert scale to transform feeling into metric values. The 'ReadDatabase' block will request the decision maker to search the database file for the name of a student pilot. The 'textToint' block converts the textual data to integer values and the function of the 'PreFlightSA' block will use the answers to calculate the average percentage pre-flight SA score and save it within

the respective student pilot database file. The architecture can be seen in Figure 81; with the behaviour described by the ‘Obtain Situation Awareness’ Use Case in Volume II, Figure R 1..

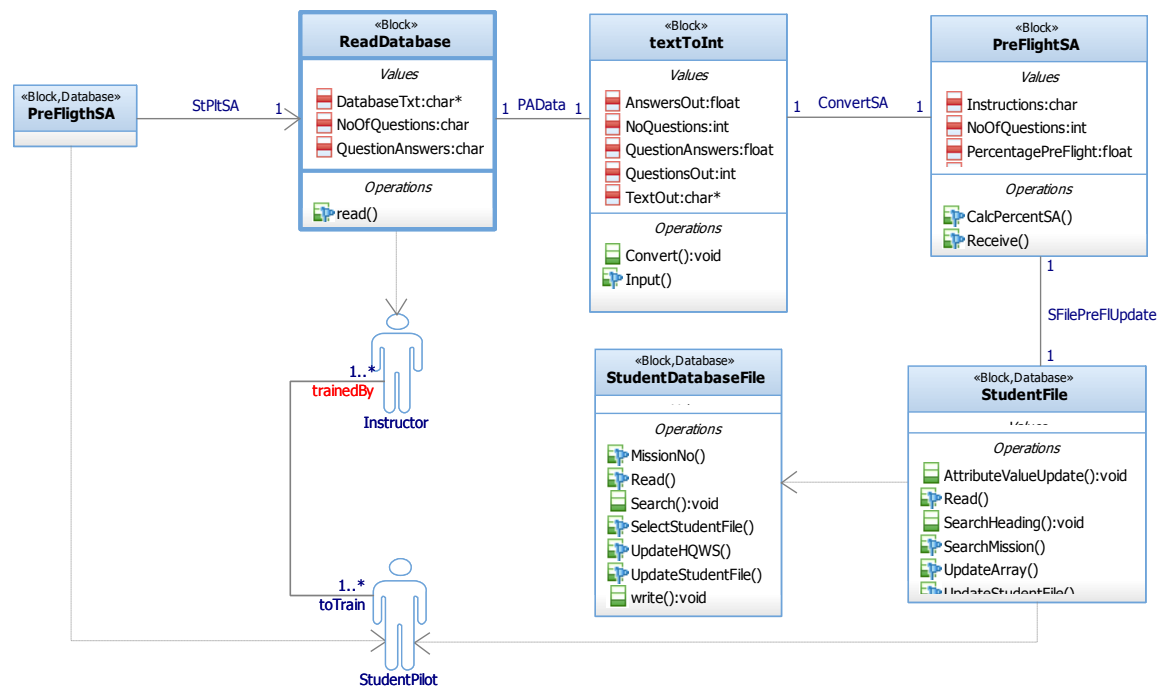


Figure 81 Pre-Flight Assessment Architecture

Low percentage scores should prompt further investigation of the pilot's state of mind before execution of the mission. For instance, if the pilot is feeling ill, it might be best to postpone the evaluation of the mission for another day; however, if the pilot is not motivated to fly the mission the decision makers subjective assessment on the success of the goals will inadvertently be lowered that may affect the next planned mission scenario and slow progression of the pilot; all of which are undesirable and contradictory to organisational objectives.

7.11.2 Pre-Flight SA Summary

The method of assessment, although simple, provides adequate information with which to base the success of mission goals on for each pilot. The questionnaire takes a matter of minutes to complete and less than a minute to evaluate. The completed questionnaire answers are available within the database should the assessment score be low to allow for further investigation and subsequent decision of the suitability of the pilot on training aspects of the mission at that point in time. However, the assessment can be optional as most SA assessments are completed post-flight (Pritchett, 2012); (Pool et. al., 2012).

7.12 Goal Modelling Techniques for FTS

Every mission scenario is developed with a goal and with its objectives, i.e. precision of flight, to achieve the goal. A goal is defined by Lamsweerde (2009) as ‘*a prescriptive statement of intent that system should satisfy*’, i.e. an achievement or accomplishment for the effort given, whereas objectives can be seen as time-related to achieve certain tasks and are measurable (see Chapter 7.10). A goal can also be defined as a statement of intent and the identification of strategic goals is through keyword analysis of the mission statement and the specification technology available (Pokoradi, 2011); using binary reasoning to clarify the satisfaction of goals, measured with a degree of satisfaction elicited from SMEs opinion, with the chosen blending mix. There is of course a matter of uncertainty due to unfamiliarity with technology and configuration of controls and indicators. Questions such as: how likely is the goal to be satisfied and how likely an obstacle is going to occur in the current configuration setting?, have to be asked in the development of quantitative approach, which will provide a probability of satisfaction (O’Hagan et. al., 2006). In a situation where technology is involved, in the absence of engineering and software knowledge (as SME are experienced in instructing and flying aircraft), eliciting quantities of interest i.e. satisfying training goals, from experts has to be based on the experiences of those making decisions. The probabilities allocated to the sub-goals when summed should equal to one.

Goals along with relationships between them provide a logical basis for arguing on the MoE related to mission goals as per the holistic view of learning illustrated in Figure 11 of Chapter 4.2. Performance goals are often set by the decision maker to control the workload based on their subjective opinion or feedback performance results. However, it is extremely hard to assess whether a goal has been met unless the goal is systematically decomposed into specific requirements with specific margins. Assessment has a strong dependence on SME opinion, about probabilities and weights, with specific K&S gained from training and experience (Kerkering, 2012); for instance, for task goals the goal to be assessed can be: achieve success in maintaining aircraft course and time accuracy within predefined task tolerances. Consequently, the objective is to achieve the greatest amount of accuracy in flight to match task tolerance levels; and goal modelling is used to estimate what the decision maker believes can be achieved in relation to task goals, as per Figure 15.

The abilities of the student pilot have to function in the background of K&S, which will be used in most activities executed in the mission scenario. Thus, using ‘X’ for K&S, ‘G’ for

goals, and ‘A’ for actions, the process to improve performance of the student pilot can be described as:

Belief (based on ability)+ Applicable ‘X’ + G => HOW TO ACHIEVE PERFORMANCE

Achieving performance concerns aspects of gaining confidence and experience to continue achieving higher levels of performance by acquiring greater levels of K&S using the blending mix when compared to previous missions. This predicate well-formed-formula brings clarity to quantities and quality levels. This can be used during planning of training missions for goal formulation, performance analysis and evaluation of instruction. For the proposed workflow process, prediction of goal satisfaction in relation to the understanding of the pre-mission brief is described as a functional relationship with key objects to be recognized within the mission as follows:

$$f : O * +G \rightarrow A$$

Where O^* represents certain objects of interest that should be recognised as prediction of a change in state, G represents the goals of the task in relation to the object and A is a set of possible training actions based on training rules (SoPs) using the chosen blending mix. This function represents a mapping from object and goals directly to actions. Generally, probabilistic simulation techniques moreover Monte-Carlo analysis (Belvardi et. al., 2012), is used to analyse goal satisfaction level and uncertainty in expert opinion. Thus, mechanisms are introduced to measuring sensitivities of an overall goal to different factors to necessitate the extent of any improvement necessary for satisfying an overall goal.

It is natural to be able to validate learning objectives by using modelling techniques in association with performance feedback provided by the training technology to be able to quantitatively reason about the satisfaction of current K&S levels in achieving the overall missions goals and further identify the characteristics and behaviour of the technology and the pilot that must be improved to increase the goal satisfaction. An approach known as goal modelling (Kelly & Weaver, 2004) is suited for prediction of goal success of tasks and for the FTS consists of three main components enabling quantitative assessment: the goal models themselves, SME elicitation (O’Hagan et. al., 2006), and probabilistic simulation. Goal modelling will establish for the decision maker, in a formalized way, the level of confidence that a student exerts and functions using a particular technology and configuration within pre-defined limits set by SMEs.

7.12.1 Why Model to Estimate Success of Mission Objectives

The mission scenario is decomposed into structured goals (Lamsweerde, 2009) consisting of attributes that can be measured using the technology i.e. airspeed / time accuracy, altitude, course heading, etc. Each goal is further categorised and associated with relevant K&S being evaluated and verified between each waypoint or mission related task. Expert elicitation techniques (Meyer & Booker, 2001);(Forester et. al., 2004) are applied based on evidence from previous subjective assessments, instructor opinion and/or previous performance evaluations, for the mitigation of potential biases; however, it has to be noted that dependability has a strong reliance on expert judgement (Littlewood & Wright, 2007); in time, previous performance results will offset this dependency and reduce any potential negative bias from the instructor. One of the key areas that can be enhanced by using goal modelling techniques is to identify significant areas of uncertainty, e.g. a new student with swings in personality traits that can lead to significant ambiguity in expert judgements, knowledge deficiency of blending mix training effectiveness, define an objective fitness criteria to reduce uncertainty; and create a more robust evaluation based on subjective assessments, daily pilot traits, and previous performance results.

The first task when constructing the goal model is to elicit the mission scenario task details and descriptions; then, identify the qualitative evidence which will enable quantifying the probability of each of the goals being satisfied through decomposition and precedence order of each goal (Holden & Dickerson, 2013). High level goals describe the general strategic objectives of the mission generally related to performance and K&S levels, lower-level goals are more specific in nature to how the mission is going to be completed, i.e. what technology is the pilot going to use to complete mission objectives; this incurs the specification of what the strategic goals are that the technology must satisfy along with identifying any potential obstacles that could affect success of the goals. The obstacles that can prevent the student pilot from achieving the set goals include the lack of understanding of the specific operations required by the tasks in the mission, the unsuitability of the technology for the student pilot that increases their workload unexpectedly and the lack of SA caused by deficiencies in K&S. SME opinion must be elicited on the chosen blending mix for use of accomplishing the strategic goals and then a decision must be made about whether goals need to be amended to suit the technology or the pilot (/ptissues of Figure 82). If the level of satisfaction gathered from the performance feedback or observer evaluations is low, sensitivity assessment can be completed to identify the input quantities with the most significant impact of overall goal

satisfaction, and thus will assist in creating an action plan to make improvements such as: collecting more evidence, additional provisions in technology, or identification of suitable pilots to achieve the set goals and ejection of pilots who consistently fail to meet the 'grade'. Alternatively, if the pilot performs the tasks as predicted it can be considered that SME judgement is accurate to the capabilities of the technology / system and abilities of the student pilot interacting with it, the process can be described in Figure 82.

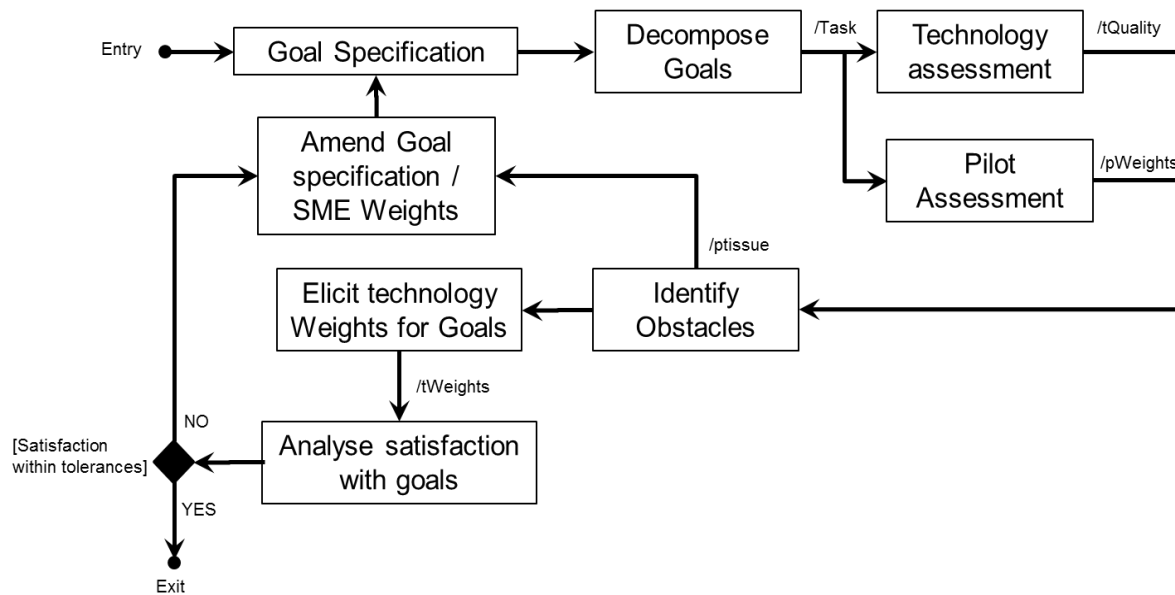


Figure 82 Goal Modelling State Behaviour

7.12.2 Goal Model Implementation Design

The probability of satisfaction of the goal (G_i) given supporting evidence items ($E_1...E_n$) is obtained from subjective opinion and questionnaires from previous completed assessments relating to the student pilot with concentration given to the objectives and actions of the goal and querying: based on $E_1,...,E_n$, how likely is G_i to be satisfied. The goals can be made at the task or activity level relating to the goal of the task the pilot is attempting to satisfy. These probabilities will be in the form of a triangular distribution, which are highly intuitive, account for uncertainty and promote ease of understanding. To define the distribution, the decision maker has to decide on three parameters: the maximum performance probability of satisfaction of the goal (b), the minimum performance probability of satisfaction of the goal (a), and the most likely performance probability of the goal (m), with concentration given to the maximum likelihood value first, as seen in Figure 83.

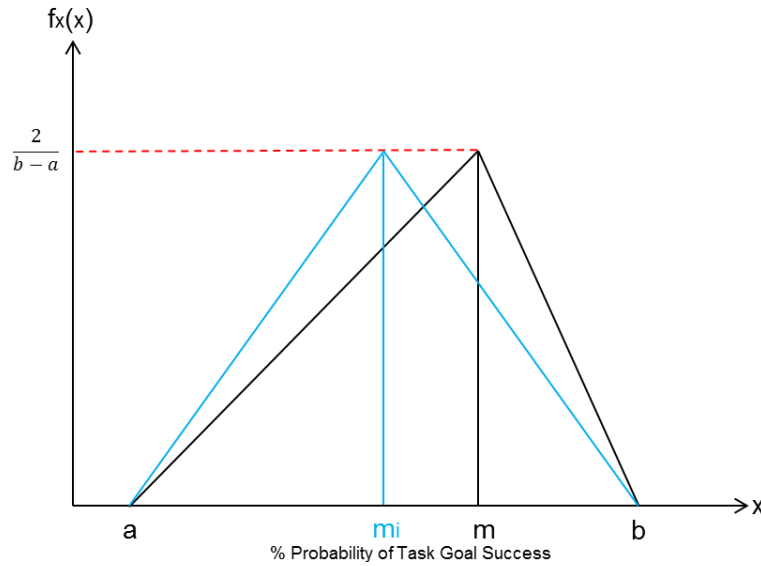


Figure 83 Goal Modelling Assessment Technique

The triangular distributions are calculated using a density function formula as follows:

$$f_x(x) = \begin{cases} \frac{2(x-a)}{(b-a).(m-a)} & \text{if } a \leq x < m \\ \frac{2(b-x)}{(b-a).(b-m)} & \text{if } m \leq x \leq b \end{cases}$$

The shape of the distribution is the important quantity to assess as missions are evaluated. The 'blue' triangular distribution informs the decision maker that between the two limits of satisfaction there is an even chance that the student pilot will succeed with the chosen blending mix, indicated by m_i . The task for the instructor is to ensure that when predicting goal satisfaction, 'm' is consistently transitioning to the right of m_i or the previous location of 'm' for a similar task to obtain confidence that training is transferring the correct levels of K&S with the chosen blending mix for the pilot to cope with tasks. Any deviation to the left of m_i or the preceding set-point of 'm' gives a clear indication that the view is the pilot may struggle to achieve task goals at this set difficulty level of the task and that possible negative ToT is being experienced with the chosen blending mix. To assist in easy analysis a cumulative triangular distribution can also generated to concur with decision maker's assessment with crosshairs giving identification of improvement or deterioration in decision makers views on progress, as follows:

$$f_x(x) = \begin{cases} 0 & \text{if } x < a \\ \frac{(x-a)^2}{(b-a).(m-a)} & \text{if } a \leq x < m \\ 1 - \frac{(b-x)^2}{(b-a).(b-m)} & \text{if } m \leq x \leq b \\ 1 & \text{if } x \geq b \end{cases}$$

Once the satisfaction probability values are obtained for all the tasks within the mission scenario, an evaluation on overall mission goals with the chosen blending mix can be performed using point value algorithm based on the elicited probabilities (Robert & Casella, 2005). This method is based on repeat random sampling of a set input probability variables decided by the decision maker to obtain an estimate for the degree of satisfaction of a strategic goal either per task or per mission depending of the level of abstraction used for the assessment i.e. if task goal assessment is used then the Monte Carlo analysis is to find overall mission satisfaction for a student pilot; the algorithm can be seen in Figure 84.

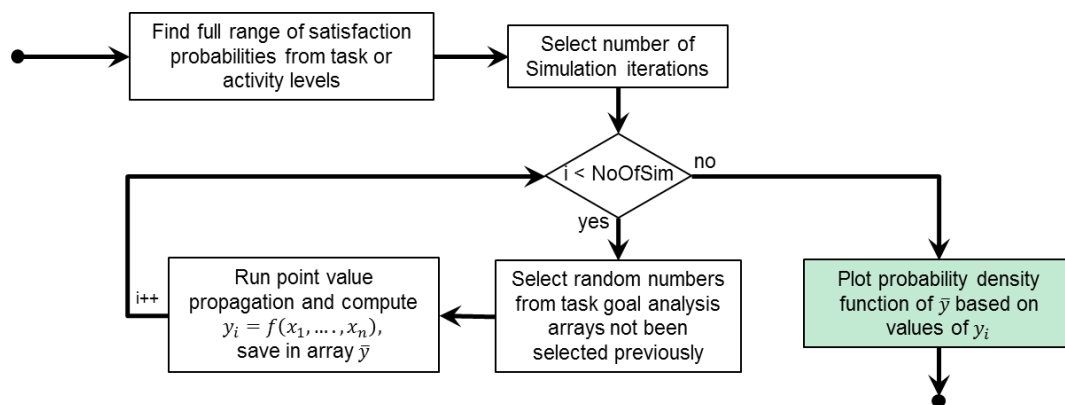


Figure 84 Goal Propagation Algorithm

In each iteration of the algorithm the input variable to the algorithm is generated from the probability distribution for each of the variables. The point value propagation algorithm, described below, is then used to record the value of y_i in an array. After a number of iterations specified by 'NoOfSim' an approximate probability density curve is generated with values falling within the value ranges of minimum and maximum of the limits of the input variables.

Point Value Propagation Algorithm

- **While** (Child Goal(s) Probability attribute is not done) – All tasks or activity distribution attribute values not sampled
 - **For** every goal satisfaction distribution Not DONE
 - Apply rule $P(G) = \prod_{i=1}^n P(G_i)$;
 - Mark Goal variable as 'DONE';
 - **End For**;
- **End While**;
- **Output** value for child goal satisfaction distribution.

The architecture used for goal modelling assessment is illustrated in Figure 85; with the behaviour described by the 'Estimate Mission Success' Use Case in Volume II, Figure S 1. The numbers of tasks are gathered from the 'TaskNoandDetails' database file for the respective mission and the 'GoalModel' block uses the data to present to the decision maker

details of the tasks. Once the assessment is complete the point value propagation mean value is saved into the main student database file for the respective student pilot.

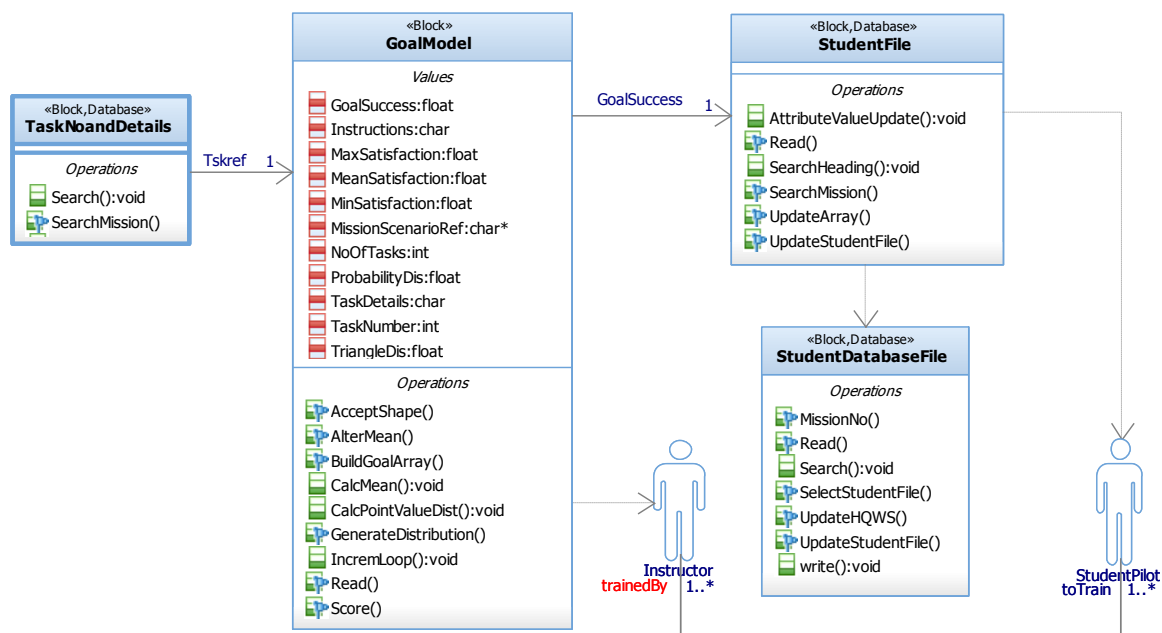


Figure 85 Goal Modelling Architecture

7.12.3 Goal Modelling Evaluation Process

In this proposed method, the assessment begins with retrieval of the task numbers and details which are displayed on screen for the decision maker to make subjective assessments on the success of the tasks based on current information about the student pilot. At this point, three windows appear in succession asking for input data regarding the maximum, minimum, mean goal successes for each task, as per Figure 86. The data is based on the goals of the task and should be related to the precision estimates given in the prediction of task performance assessment stage.

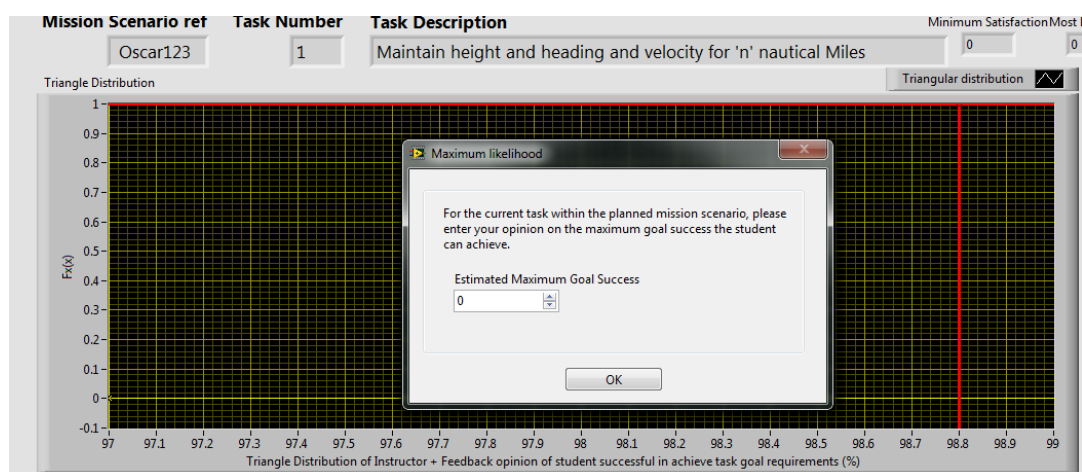


Figure 86 Inputting Data into Goal Modelling Assessment.

Once the values have been entered, the workflow tool produces a triangular distribution based on the inputted data to give a visual representation of the decision maker's objective assessment of the success in achieve the task goals, as seen in Figure 87. Hence, all the data needed to allow the decision maker to visualise their judgments on the success of the task is displayed.

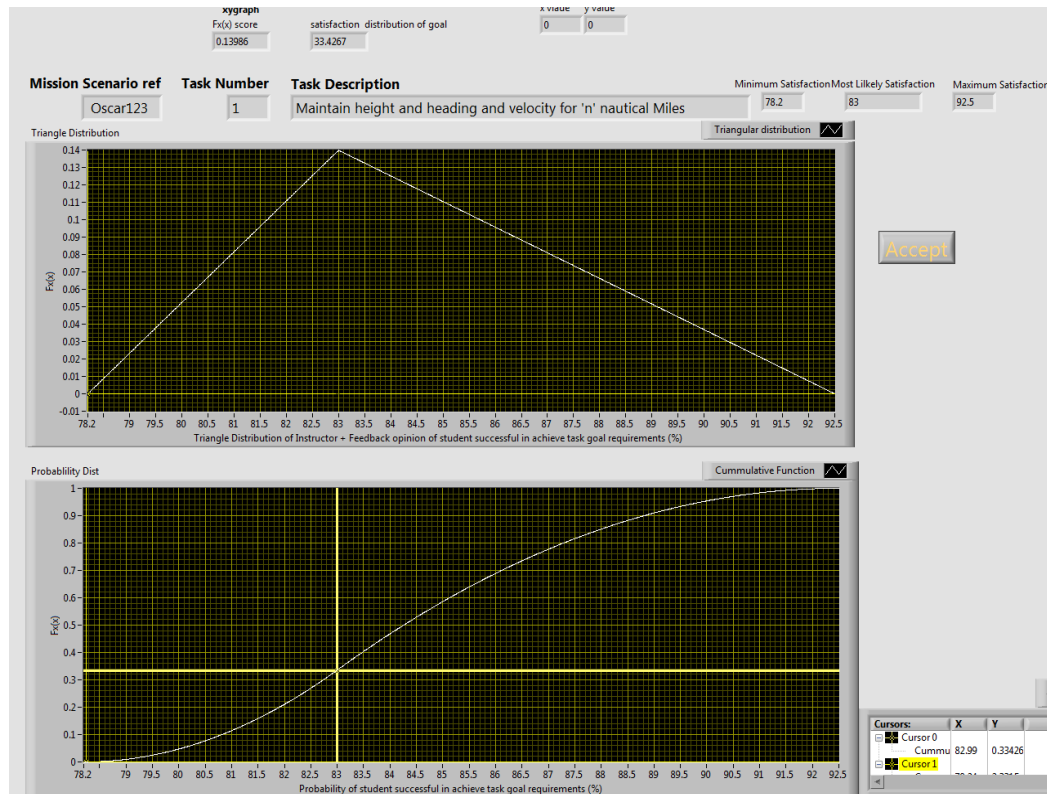


Figure 87 Example of Assessment on Student Pilot Success for Task 1

At the top of the screen, the $fx(x)$ score that is associated to the 'yellow' cursor is permitted to be adjusted by the decision maker if the mean value of the distribution needs to be amended. To the right, is the satisfaction distribution of goal, which gives a percentage metric value of predicted success using the chosen blending mix for the task: below 50% causes a concern about the necessary K&S levels to achieve the goals of the task. Below these metrics are details about the mission reference, task number and details followed by the inputted data by the decision maker. The top graph is the triangular distribution and below it is the cumulative triangular distribution that is used to give the satisfaction value of the assessment. Once the success of the task has been evaluated, the accept function is activated which transfers the assessment to the next task within the mission and the identical assessment process resumes, as seen in Figure 88.

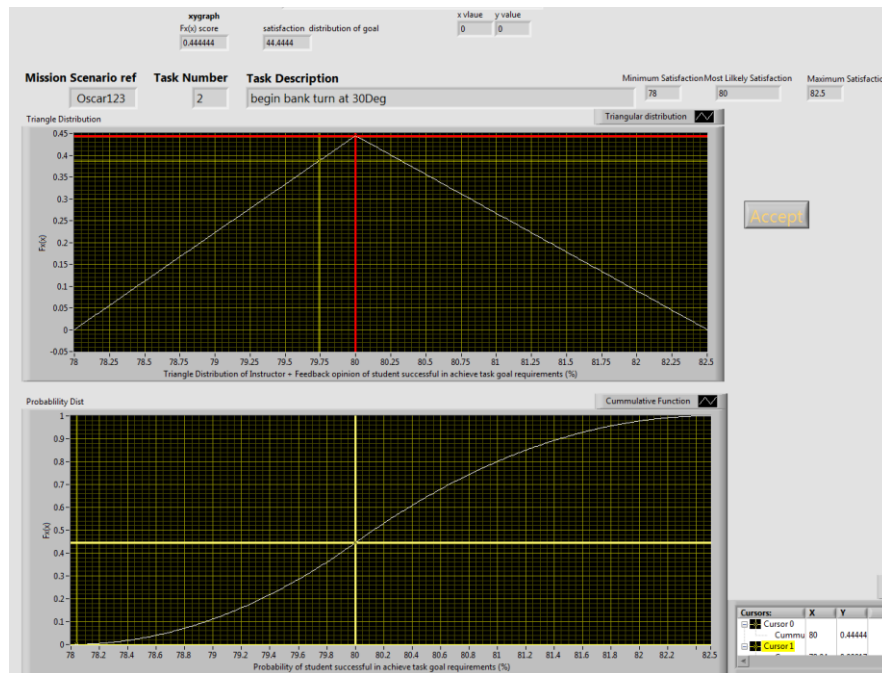


Figure 88 Example of Assessment on Student Pilot Success for Task 2

The results of full mission goal analysis are stored within a 2-D array; the mean and % satisfaction metrics are located in the first column followed by the data used to generate the triangular distribution, as seen in Table 19. The array is further saved within the Goal_Modelling directory under the relevant mission scenario reference for the respective student pilot (see Figure 39 for details about the database architecture). The data can be used to compare triangular distributions for similar tasks in different mission scenarios to indicate whether the decision maker opinion of ability is improving or whether any possible interventions are necessary.

Table 19 Goal Assessment Results

Goal Modelling Results											
Oscar123	Mean & % Satisfaction	x - Probabilit									
Task 1	83.000000	78.200000	78.210000	78.220000	78.230000	78.240000	78.250000	78.260000	78.270000	78.280000	78.290000
	33.426719	0.000000	0.000291	0.000583	0.000874	0.001166	0.001457	0.001748	0.002040	0.002331	0.002622
Task 2	80.000000	78.000000	78.010000	78.020000	78.030000	78.040000	78.050000	78.060000	78.070000	78.080000	78.090000
	44.444444	0.000000	0.002222	0.004444	0.006667	0.008889	0.011111	0.013333	0.015556	0.017778	0.020000

The results of the point value propagation algorithm is generally used to predict mission success, but can also be used to grade student achievement by sectioning the triangular distribution and grading the section lines, as per Figure 89, which is highly dependent on the shape of the distribution.

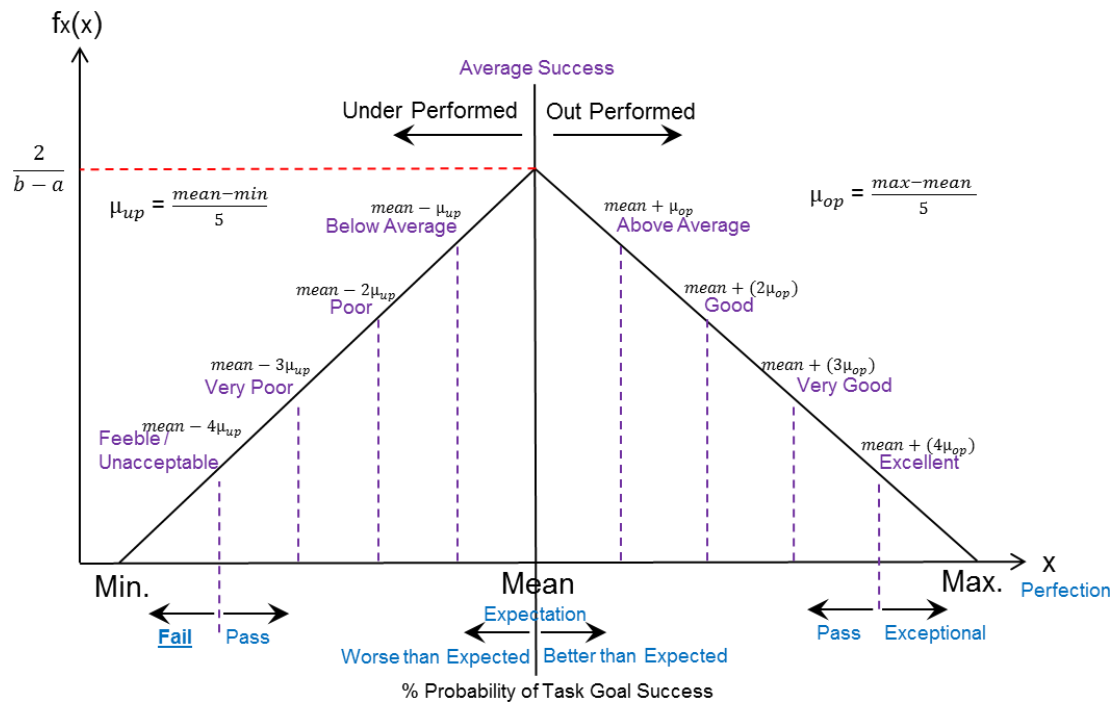


Figure 89 Potential Goal Modelling Grading System

If the mean (expectation) is towards the maximum value and the minimum value is low compared to the maximum, there is greater scope for the student pilot to underperform in task goals but still be within expected ranges (greater scope of bounds between mean and minimum acceptable achievement levels). The method, however, is to propagate the mean value on each goal assessment for different missions towards the maximum as an indication of improvement in pilots ability to understand and with it perform task goals using the chosen blending mix, which can be used as an early indication of ToT effectiveness. Note: it is perceived that the distribution will be far from a perfect triangular distribution, but the result should still determine a peak goal satisfaction for the mission or task (depending on the level of abstraction used for the analysis).

The estimations of task or mission success can be clarified with feedback data from the technology and further analysed with subjective and objective feedback from both the instructor and the student pilots. This feedback data is used to display to the decision maker the performance outcomes of previous missions as they are assessing the goal success for the next mission and therefore can be used to check their own bias viewpoint with respect to set goals.

7.12.4 Goal Modelling Summary

This assessment stage describes a visual approach to quantitatively assess the goal success and how to systematically elicit expert opinions. A key advantage to goal modelling is the

intuitive manner with which information is presented to the decision maker to conduct analysis that is easy to use for other non-technical personnel. The main novelty involves seamlessly combining expert elicitation and goal modelling in a coherent implementation for characterising the clarity of arguments through graphical representation to allow the decision maker to reason about task and mission success. The legitimacy of the quantitative outcome has to be complemented with previously completed assessments within the workflow process and pertain to the decision maker's qualitative judgements to reflect insights into views on progress and blending mix suitability in acquiring the relevant K&S for task and mission success.

The method of associating predicted goal success to performance precision estimates can assist in discovering any change in views on the student pilot and the chosen blending mix based on how the student pilot feels before executing the mission scenario for evaluation. The association also assists in examining the attributes of concern that determines whether the pilot has successfully achieved task goals whilst systematically failing (or being graded low) certain task performance objectives, during post mission analysis (see Chapter 7.13), i.e. it is possible for a pilot to achieve task goals (destroy target), but fail objectives (maintain accuracy of flight). The distinction between performance attributes and goal attributes can determine what aspects of pilot's abilities (K&S, decision making, understanding of requirements, etc.) and/or blending mix capability or configuration setting and/or fidelity, needs improving and thus assist in directing further mission scenarios to deficiencies.

7.13 Performance based Comparison

The primary purpose of performance measurement is to identify strengths and weaknesses in K&S necessary for readiness to focus future training on (Bennet et. al., 2002), although this type of measurement system can be highly dependent on the chosen blending mix capability. The training attributes per task profiles created in Chapter 7.4 (re: Table 7) can be used to trace progress and tailor exercises based on student pilot and technology weaknesses on specific training attribute areas. In a common measurement framework and workflow process observation- and simulation- based data can be integrated to provide assessments at the training attribute levels. The accomplishment of a task undertaken by a human operator is generally referred to as performance with emphasis given to the discussion in Chapter 4.4. These performance measures can be classified into a number of categories that include (Beaubien et. al., 2015):-

- Accuracy – the degree of correctness
- Time – duration of completion of the task
- Task battery – a number of different tasks performed in series or parallel to measure a range of abilities or effects

In this proposed method, each executed mission scenario is analysed against a baseline scenario of a flight executed by computer based simulation. The baseline scenario is used to base measurements of the accuracy of aircraft controls and how long it should take for the aircraft to fly to known locations or waypoints. The simulation represents the behaviour of an ideal pilot with a certain degree of validity and precision (see Chapter 7.1 and Appendix H); however, the simulation could never replace the human aspects and timings of flight executed by an experienced fully trained pilot. Data gathered from either is deemed suitable for baseline comparison studies to grade mission success using the chosen blending mix and time requirements of the mission. Special constraints and ratings for straight and level flight, entering and maintaining bank turns, and exiting bank turns are given to provide indication of situation awareness and accuracy in control; these constraints are directed to the accuracy and precision assessments completed in Chapter 7.10, where any change in state of the aircraft is assigned a separate task number. The tolerances in accuracy are perceived to be scored in the following manner:

- If the pilot controls the aircraft consistently within the set accuracy tolerances measured at each time sample point of assigned airspeed (described by time accuracy), assigned heading and assigned altitude, the pilot is rated a maximum of 4.

- If the pilot controls the aircraft mostly within the set accuracy tolerances, the pilot is rated a maximum of 3.
- If the pilot controls the aircraft mostly outside these set accuracy tolerances results in a rating of 2.
- Pilots who could not control the aircraft within these accuracy tolerances is awarded the lowest rating of 1. (These constraints consider the fidelity and sensitivity of the flight controls and flight instruments).

The total average percentage for each task is then used to compare with the predicted goal achievement to evaluate the subjective opinion of the decision maker with actual performance results in an effort to amend possible bias used to assess the success of the goals for future missions. The method for performance analysis can be summarised in Figure 90; the first two phases of the assessment are for quick visualisation of areas of interest that requires to be further investigated followed by an evaluation of the execution of the mission with the baseline scenario on a per task basis. The next two phases involves comparing the executed mission scenario to predicted performance estimates from previous stages, which is used to grade mission success but also to evolve future predictions based on performance outcomes. The pilot control adjustments are then investigated for spikes or constant adjustments and the data gathered can be used for discussions in the post-mission briefs.

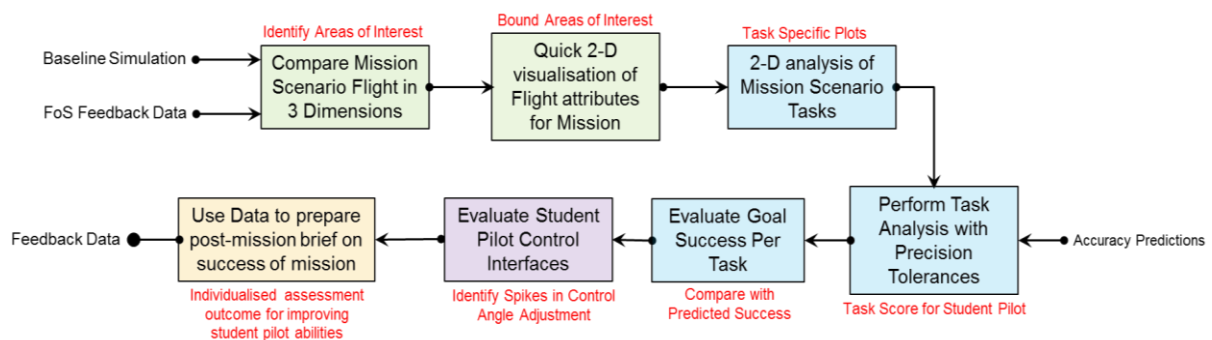


Figure 90 Simple Block Diagram Describing Method of Assessment for Performance

7.13.1 Performance Based Comparison Implementation design

There are a number of different views used to analyse the student pilot flight statistics including 3-D and 2-D plots. Data from the baseline scenario flight, either from computer simulation or expert pilot using a system in the FoS for the identical mission scenario, is retrieved along with the data from the evaluated mission scenario for the respected student pilot, which is imported by the 'SimulationData' block, as seen in the architecture described in Figure 91.

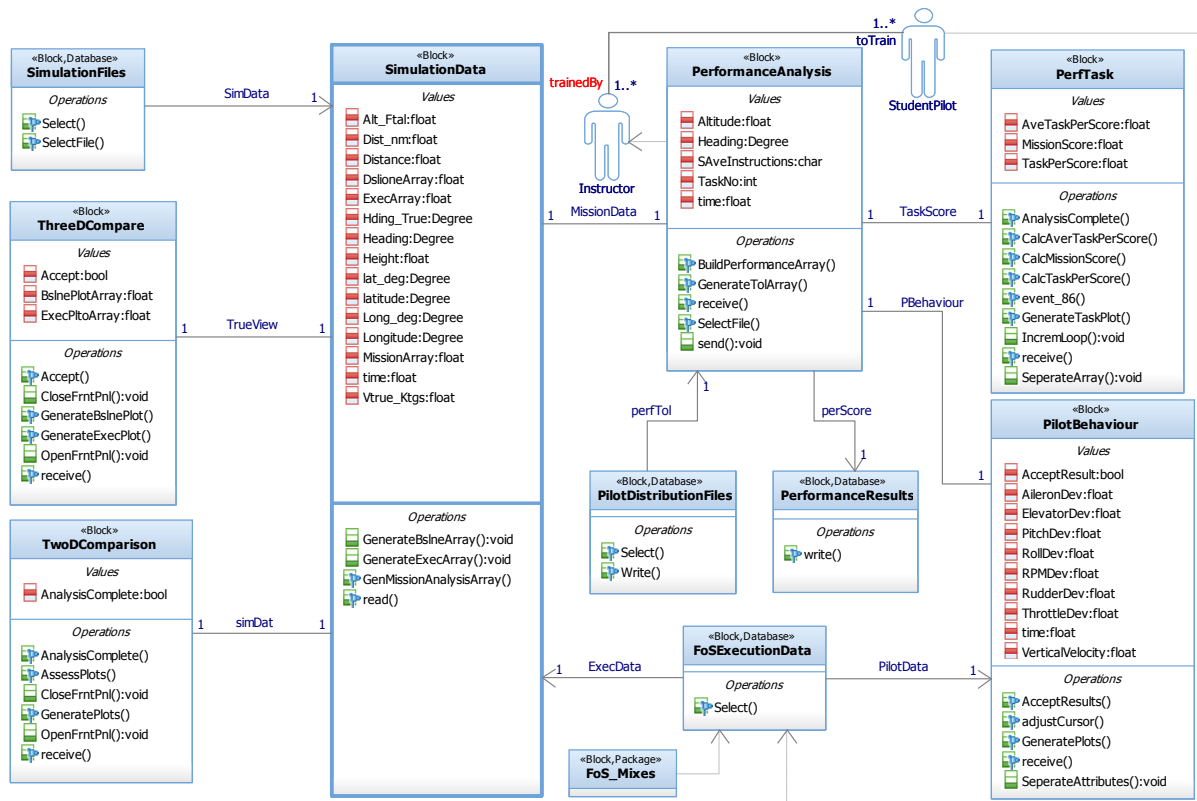


Figure 91 Performance Based Comparison Design Architecture

The 'ThreeDCompare' block provide the functionality to display the flight in real time to allow for discussions on any time instance that caused a deviation in flight from the planned; there are a number of views available from separate graphs to both the simulation and executed data on one 3D graph. The 2-D graphs are generated by the 'TwoDComparison' block that displays aircraft flight attributes with respect to each other or time for a more detailed analysis of potential causes of concern identified within the 3D comparison. The 'PerformanceAnalysis' block retrieves data from the PilotDistributionFiles database for accuracy and precision data to permit grading on the accuracy of flight. The block also provides the functionality to investigate flight attributes on a task-by-task bases, which is mainly used to provide visual evidence of any issues to discuss with the respective student pilot during post-mission briefs and obtain further feedback on the views of technology suitability. The architecture is further advanced to allow analysis of the changes in state of the aircraft controls for investigation into the nerves of the pilot and to their ability to control the aircraft smoothly using the chosen blending mix; this also can give an indication of possible 'startle' effect that causes a temporary loss in concentration showing spikes in controls due to overcompensation events. The performance results from the analysis are saved within the PerformanceResults database for the respective student pilot. The behaviour

of the architecture is described by ‘Assess Students Performance’ Use Case in Volume II, Figure U 1.

7.13.2 Performance Base Comparison Evaluation Process

The data (Longitude, Latitude, Altitude) from the baseline simulation and the executed mission scenario from the system of the FoS is used to perform a real time analysis of both flights using either 1 3-D graph or separate 3-D graphs to identify any deviations of time defects with the executed flight, as seen in Figure 92.

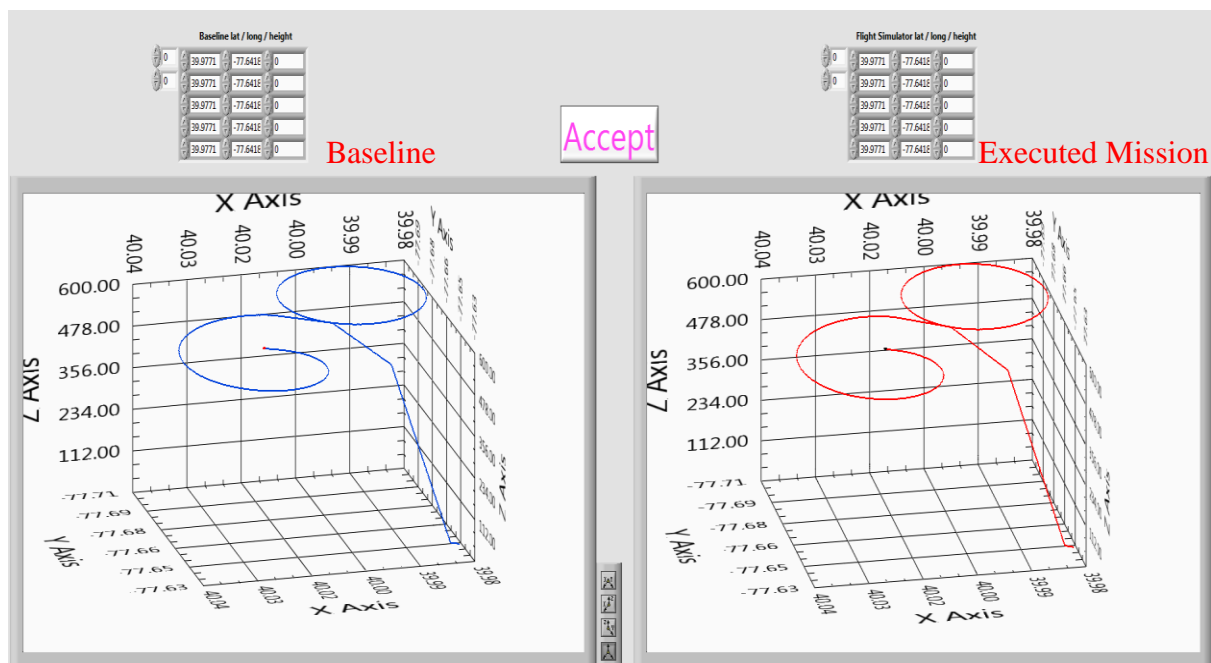


Figure 92 3-D Comparison of Flight with Baseline.

Any issue with timing or course deviations can be efficiently identified and the decision maker can mark the time stamp for further analysis. This analysis stage is for quick identification of areas of concern that require further analysis in more detail using 2-D graphs of the mission. The UI permits the decision maker to rotate the graphs or zoom in on the graphs for a more details comparison. The behaviour is described by the ‘Compare Mission Execution in 3 Dimensions’ Use Case in Volume II, Figure U 3. From the identification of any control issues identified on the 3-D graph analysis, the flight attributes are plotted against each other for a more detailed analysis of the full flight. The graphs plot: latitude vs time, longitude vs time, latitude vs longitude, distance vs time, heading vs time, height vs time, latitude vs height, longitude vs height, and distance vs height; as seen in Figure 93.

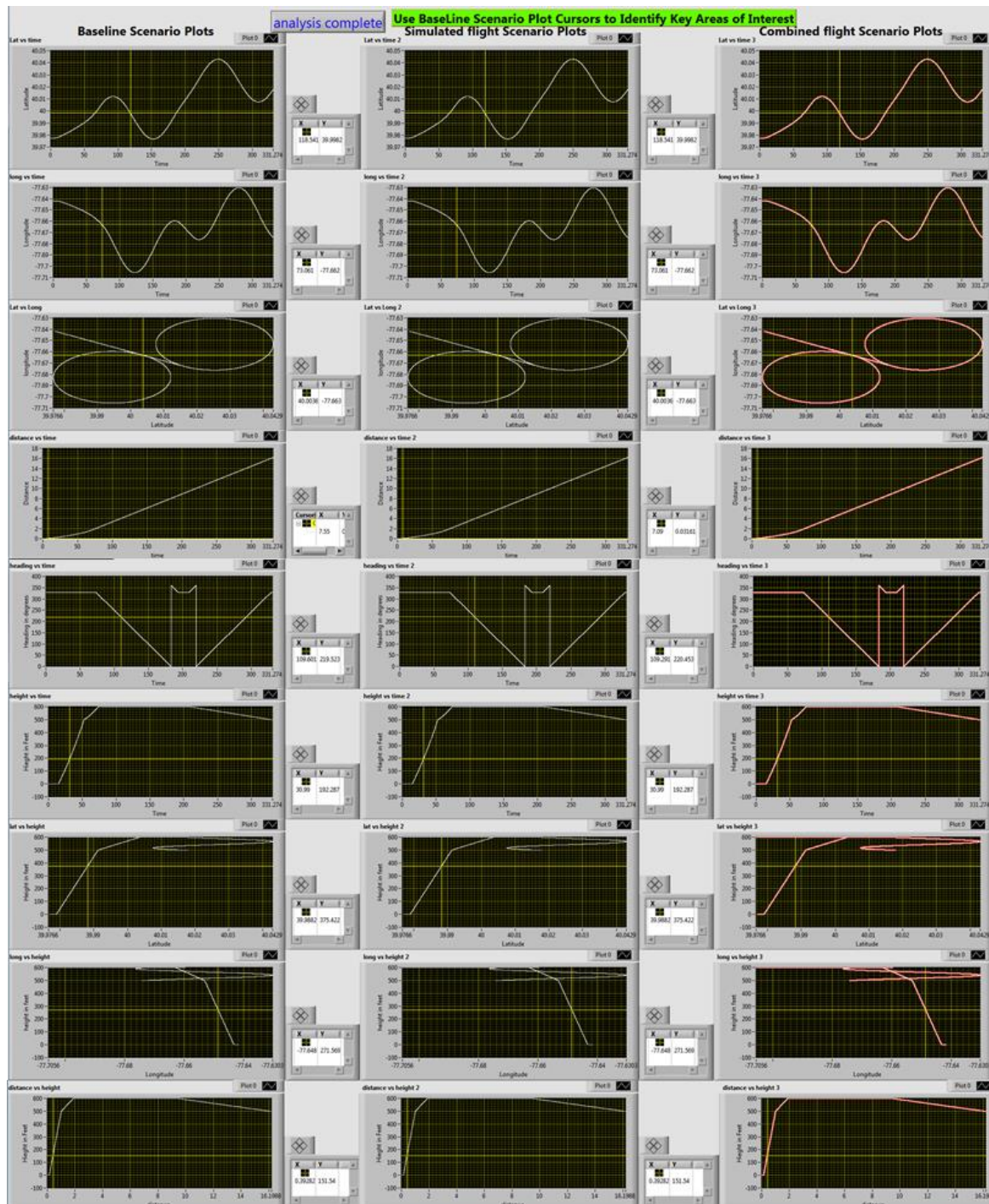


Figure 93 Analysis Using 2-D plots of Flight Attributes

The graphs produce a 9*3 frame with the first column describing the baseline scenario plots, the second the executed scenario plots and the third are combined plots within the same graph. For a more experienced pilot the combined plot may not distinguish flight deviations easily in a visual manner, thus separate plots are displayed to easily identify differences in shape of the relationship. The decision maker has a number of options within the UI:

- The baseline scenario graph controls the position of the crosshairs of the corresponding graphs to the right of it, i.e. a change in position of the cursor will automatically change the position of the 'x' axis coordinates of the crosshairs on the other two plots to the right. Along with the change the attribute values are displayed to give a clear identification of the deviation from the baseline.
- There are options to allow zoom into the graphs to allow for more detailed display on areas of concern identified from the 3-D analysis.
- The decision maker can make notes on attribute values, i.e. time, altitude, etc. with which to base future discussions on.

Once the analysis is complete, the decision maker selects the 'Analysis Complete' virtual button to transfer to monitor the task behaviour. The behaviour can be described by the 'Compare the Flight in Detail' use Case in Volume II, Figure U 4.

The precision tolerance values from the predicted performance assessment stage are retrieved and used to grade on the success of task objectives. The identical 2-D arrangement, as per Figure 93, is used for analysing the results of the executed mission on a per task basis, as seen in Figure 94. The executed mission data is separated into tasks using the baseline scenario time stamp to search through the data for specific stages of the flight. Once the details in the graphs are assessed, the decision maker uses the 'Analysis Complete' function to permit the workflow process tool to transfer to the second task details for the decision maker to analyse. This type of detailed analysis focusses the decision maker on identifying key areas and time within the tasks when potential issues arose and can lead to key information traces through the mission scenario to identify any objects that the student pilot should have identified or any obstacles caused by the technology that affect issues with the change the state of the aircraft.

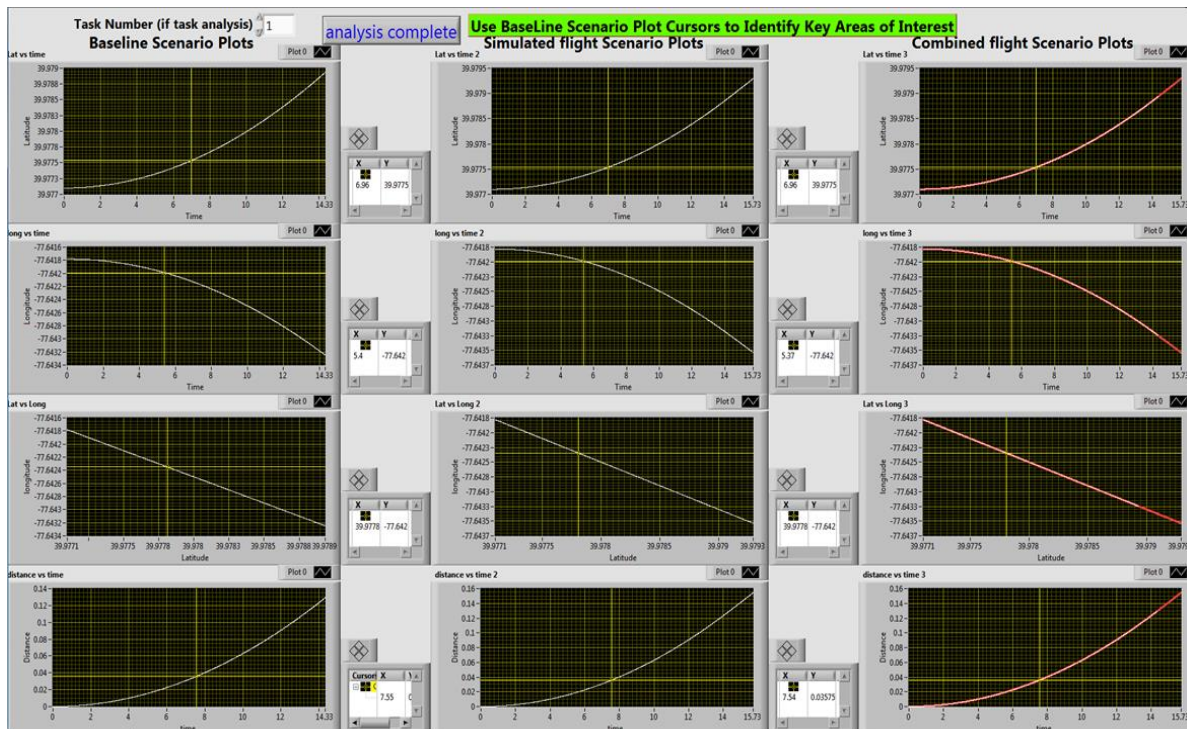


Figure 94 Subset of Performance Per Task Analysis

The tool will use the information stored within the plot arrays to calculate the grade achieved at each sample point within the task. The average score per task will then be calculated followed by the total score for the whole mission scenario. The score/grades concentrate on the performance of the chosen blending mix using the performance outcome from the student pilots executing the mission scenario on the chosen blending mix. All the scores are saved within performance distribution database for future reference. The behaviour is described by the 'Perform Performance Task Analysis' Use Case in Volume II, Figure U 5.

To add substance to the task analysis, data concerning the pilot controls can be retrieved from the FoS database, which pertains data involving the angle adjustments of the pilot controls. The values of the control attributes are plotted against the real time of the mission scenario to gain insight into stages of flight that the pilot had spikes in control adjustments, a simple illustrative example of a subset of task data can be seen in Figure 95. All control attributes are link by the mission time attribute, therefore the crosshair control for all the graphs is directly dependent on the cursor position of the elev vs time graph plot (master with all other graph crosshairs are slaves with regards to 'x' axis cursor position). The decision maker can adjust the position of the crosshairs to identify exactly what control position all interfaces are in at one point in time. This can give an indication of periods of time that the student pilot lost SA and quickly corrected the rotation of the aircraft or increased/decreased throttle to maintain course. Cross referencing the time stamps from the aircraft attributes to control

attributes can identify any causal links of any course deviation from planned and will assist in post-mission briefing with the student pilot and can, to some degree, confine areas of the flight where the incorrect SoP was actioned or confusion caused indecision or ‘startle’.

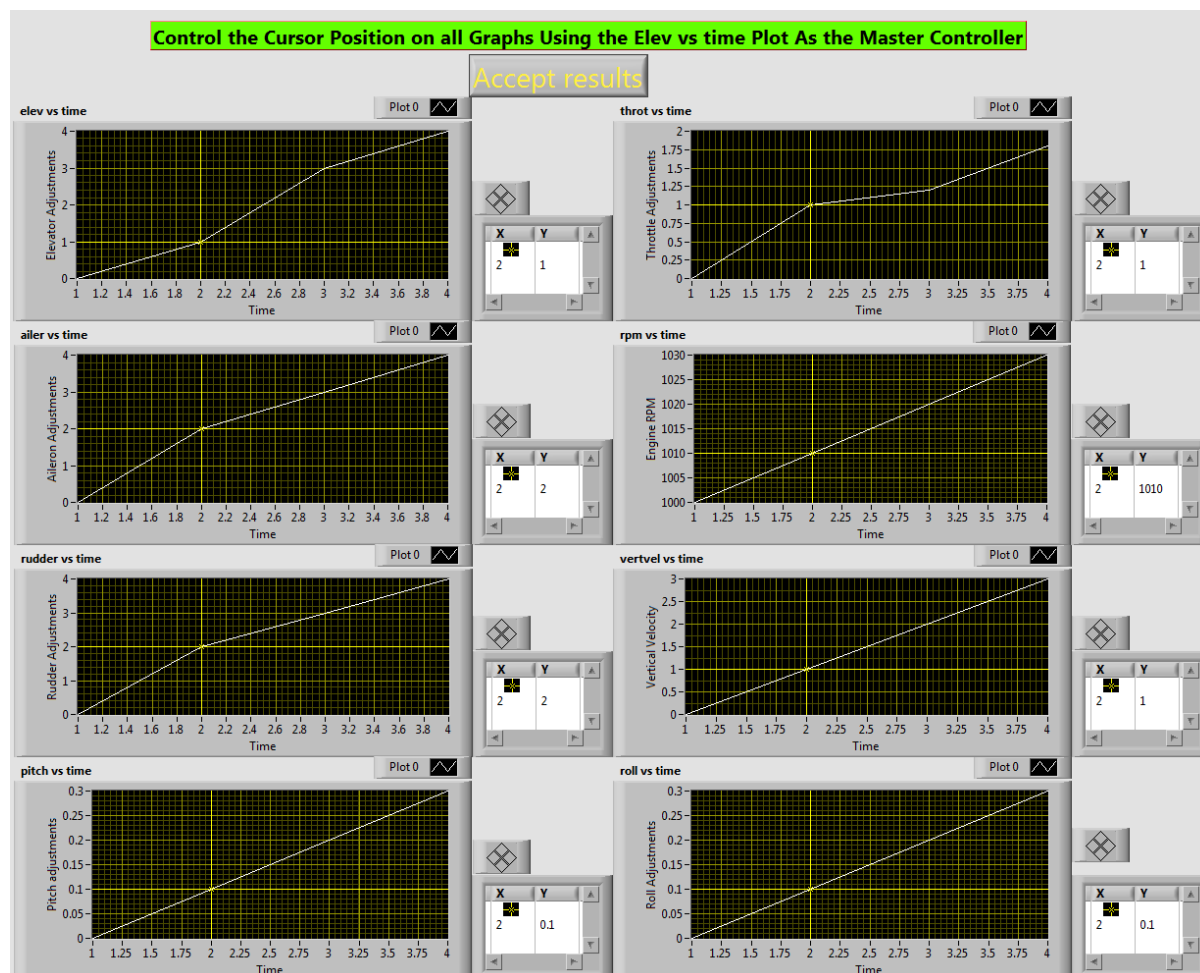


Figure 95 Example of GUI Arrangement for Pilot Control Adjustment Assessment (subset of task)

The data contained within the graphs can be stored within the performance distribution database, if required. The behaviour can be described in the ‘Assess Pilot Operation Behaviour’ Use Case in Volume II, Figure U 6.

7.13.4 Performance Base Comparison Summary

The initial phases of analysis concentrate on the full spectrum of the executed mission scenario with which to identify the ROI that are causing concern. Although, these phases can be ‘skipped’ it is important to narrow the detailed analysis to areas that require time at the expense of other phases of flight that are routine and not causing any training concern with respect to training attributes. Once identified, at the task specific graph assessment phase, certain mission tasks can be accepted immediately whilst time is better used at comparing the

scenario tasks and associating the deviations of flight with identified training attribute relationships acknowledged in the mission planning stage and discussed in Chapters 7.3-7.4. The proposed workflow tool described by this methodology can easily be modified to save certain time constraints within the mission task from this phase to be highlighted in the graphs when assessing control interface responses to assist in identification of any causal links between deviations of flight and loss of control. The dependency of the crosshair controls in the graphs can assist the decision maker in analysing the data contained within them by providing visual cues without the need to re-adjust the cursors (or eyes) and is seen as an integral function for ease of analysis. With performance issues identifiable on a per task basis, concentration can be given to the investigation of the relevant training attributes per task created from ROSETTA 0 assessment (Chapter 7.4) and assist in evolving the training programme to concentrate on the reasons why those training attributes in that stage of the mission scenario are deemed to be weak. The decision then is whether the problem is caused by the suitability of the blending mix, either in relation to its capability to perform correct functionality to the planned K&S levels or to the student pilot, or deficiencies in the student pilots' K&S levels; comparative study with student pilots who used the identical technology flying the same mission scenario could direct the investigation to the appropriate area.

The ability for the legitimacy of the comparison of the executed mission scenario to the baseline scenario model lies in the capability of the baseline simulation to synchronise to the training technology sample rate. The synchronicity needs to provide sound predictions within certain set bounds with which to grade on both performance and goal success; these bounds are centred on the phases of flight. If the differences between model simulation and training technology output data are unacceptably out of synch, confirmation of the performance grade is deemed unverifiable and consequently the outcome has to rely on the instructor judgement of performance and goal success. The outcome of the assessment proposed in this method concentrates on the performance of the chosen blending mix using the performance outcome of the student pilots executing the mission scenario using the technology to determine whether the capability of the training technology is suitable for practicing and exercising the levels of training attributes needed for the transfer of training (ToT). It is then for the decision maker to make a judgement of what the causes of any performance issues are and how the outcome affects the evolution of training attributes used to direct the training programme, however, how this decision is made is outside the scope of the research problem, as described by Figure 2.

Software Ontology Issues with Integration of Disparate Training Systems

Retrieving data from disparate blending mixes involves the use of a simple ontology checker to ensure the correct attributes are selected, i.e. parameter names and types. Consequently, if the proposed workflow tool was to be integrated within a FTS, the control feedback parameter names for all the technologies would need to be retrieved and added to this stage. An automatic decision can then be made based on the chosen blending mix, completed in ROSETTA 2, as to which attribute names are to be used to search through the database file. Figure 96, illustrates the issue with interrogating a database file produced by disparate training technology. The attribute names are specific to the training technology being used, in this example the FSTD uses specific names for file headers, but, for practical purposes using constants in this way would cause major issues when retrieving data. The header names would need to be stored within a database that is directly associated with their respective technology. The algorithm would then use variable names rather than constants that are updated with the information regarding the training technologies output file header semantics gathered from previous stages.

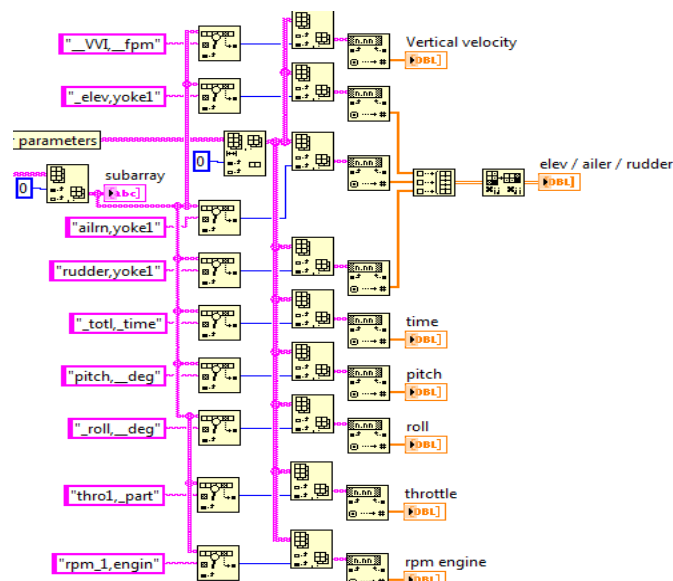


Figure 96 Illustration of Semantic Issues Relating to Database Search

7.14 After Execution Assessments

The association between motivation and performance feedback is linked to the inclination that giving feedback provides people motivation as people have a natural desire to know the effectiveness of their actions and behaviour (Alliger et. al., 2013). By formalising feedback it offers incentives for people to continue to develop K&S, and in the maverick or 'butch' nature of military aviation offers competition between pilots to achieve the best possible

target objective(s); within this target setting, the output can be directed towards the individual's efforts for achieving specific goals. This chapter concerns discussions within the methodology relating to the feedback mechanisms to evaluate the assessments used to choose the blending mix and the task difficulties involved in the mission scenario. The assessments discussed include observer-rating techniques, post-mission SA assessments by the student pilots and their objective views on the suitability of the blending mix for training the levels of K&S required by the mission to investigate the functional relationship between the technology and SA factors, as per Figure 13 in Chapter 4.3. The after execution assessment features a number of paper based assessments along with computer based questionnaires for completeness of the proposed workflow process. All assessments require the decision maker to enter the student pilots name at the start of the assessment stage within the workflow tool and each stage will evaluate the scores automatically by searching through multiple databases to retrieve response data to the questionnaires.

7.14.1 Observer-Rating Assessments

Observer-rating assessments are used to evaluate performance during execution of tasks. The decision maker (instructor) observes execution of the mission tasks, either in the background or using video images, with which to base decisions on how the student pilot is managing their SA and understands what the goals of the mission are using the chosen blending mix. Using observation measurement techniques, the mission scenarios contain tasks and actions which require certain cues to exhibit behaviour that is of prime importance to measure readiness. Acceptable responses to the tasks should be known in advance to enable a simple checklist to be completed by the decision maker at or immediately after the time of the observation (Nahlinder, 2009). During the execution the decision maker should make real-time notes based on what is observed at key points within the mission tasks that should trigger a change of state with the aircraft, i.e. with the prior knowledge of the training SoPs the pilot has, the decision maker will observe to monitor actions based on objects of interests within the flight or simulation and then make subjective assessments on the ability to cope with tasks using the chosen blending mix. The details noted during real-time execution are then summarised using a questionnaire that is related to the pre-mission assessment criteria's.

The questionnaire designed for the methodology covers areas of behaviour to be observed and assessed as described in Volume II, Appendix T Assessment Form 7. The design of the questionnaire considers aspects of how detailed the mission brief was, how comfortable the

student pilot appeared to be during execution of mission tasks, the workload of the pilot, responses to objects of interest and the subjective opinion on the success of the mission. The technique involves the observer or decision maker to rate the student pilot using a 10-point scale based on observable SA related behaviours. The 10 point rating scale ranges from 1(poor) to 10(excellent) and are specifically designed to provide a SME human perspective on performance aspects of task analysis that are used to evolve subjective assessments through the pipeline. The validity of the observer based assessment by itself is the extent to which observers can accurately rate the internal constructs of another individuals SA, however, the legitimacy of the assessment is strengthened by the multiple disparate assessments and feedback entities into the workflow process and can indicate certain aspect of behaviour regarding SA.

The architecture used for the design of the assessment is seen in Figure 97, with the behaviour described by the ‘Conduct Subjective Evaluation’ Use Case in Volume II, Figure T 1. The ‘ReadDatabase’ block will search for a student pilots name within the InstructorFlightEval database saved in the Flight_Analysis directory to retrieve the questionnaire answers. Of interest is the total average percent grade the observer has given the student, which is saved within the main student database file. A low score will initiate further investigation by the decision maker on the specific element of SA or workload that was observed as being particularly weak to enable more focus of the K&S for future planned mission scenarios for the respective student pilot. The behaviour of the architecture can be described in Volume II, Appendix T Figure T1.

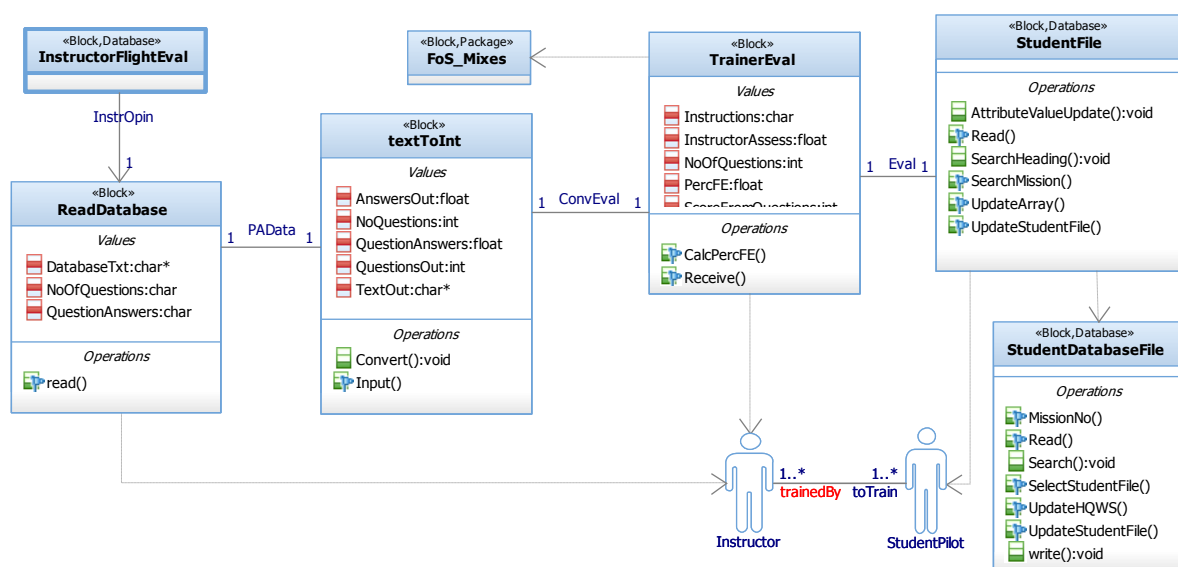


Figure 97 Instructor Observer Rating Assessment Architecture

7.14.2 Post-Flight Self-Assessments

The Post-flight self-assessment considers the student pilots subjective ratings relating to the perceived workload of the tasks in the mission scenario using the chosen blending mix decided in ROSETTA 2 assessment; the behaviour can be described by the 'Perform Post Mission Analysis' Use Case in Volume II, Figure V 1. It is expected with this type of assessment that a wide variability in results between student pilots will be received due to individuals preferring certain configuration arrangements. The workload is created by the task objectives, duration and fidelity characteristics of the training technology. The task demands imposed to each student pilot are unique and can refer to the behaviour of both the technology and the previous accomplishments of the pilot using the system. Student pilots are asked to rate their effort to measure their experience of workload and therefore it could be said that this type of subjective rating has a high 'face validity' value (Cain, 2007). The rating system used has to differentiate between experiences referring to the close correlation between workload subjective ratings and task difficulty (Salas et. al., 2011); (Stimpson et. al., 2012).

The experienced workload and psychological consequences reflect the performance outcome of the mission and obtaining information regarding 'comfort' levels is in the legitimate domain of subjective ratings, the questionnaire used to form the assessment can be seen in Volume II, Appendix V Assessment Form 8. The questionnaire attempts to satisfy formal requirements to quantify such experiences using rating scales that will result in modifiers from natural language descriptors to numerical values. The workload evaluation, therefore, is given with arbitrary scales labelled with numbers and verbal descriptions of the magnitudes represented by extreme values. The questionnaire includes the strengths of using rating scales in a NASA Task Load Index (Hart, 2006) to gather subjective results regarding: physical, mental and temporal demands along with information regarding misinterpretation or the lack of understanding of the objectives of the tasks, which cause a level of frustration during the execution that affect decision making activities (Moehlenbrunt et.al., 2011). The result of the comparison is used to create a weighted average of all scales. Included within the evaluation is feedback data involving decision making activities within mission tasks to clarify perceived learning aims of the mission, as per discussion on the readiness metamodel in Chapter 6.5, this will give a view of the understanding of the training goals. The rating offers a simple method of discerning specific aspects of workload that is reflected in the time taken to complete. The assessment is a hypothetical theory that represents the cost, in terms

of human characteristics, to achieve a level of performance and is based on the interaction between the requirements of the task, the environment which it is performed (chosen blending mix and environmental location), and the behaviours, K&S, and perceptions of the student pilot (Zeng et. al., 2011).

The evaluation is interested in obtaining the overall average percentage of the self-assessment and low results will initiate further investigation with comparisons of the decision maker's evaluation, as per Volume II, Assessment Form 7. If the student pilot's assessment considers the workload of a task using the technology to be excessive, their behaviour will be that of someone who behaves as overloaded even though the task demands are objectively low; this behaviour should also be noticed within the observer assessments of the executed mission. The student pilot would also experience period of ad hoc strategies appropriate for a high-workload situation (e.g. responding quickly, sharp control movements), experience physiological and psychological stress, which will manifest itself in the aircraft control performance feedback data assessments discussed in Chapter 7.13.2; and therefore adopt a lower criterion for performance. The designed architecture for the assessment is described in Figure 98 with the results of the assessment stored within the main student database file. The behaviour is described by the 'Obtain Pilot Self-Assessment of Mission' Use Case in Volume II, Figure V 2.

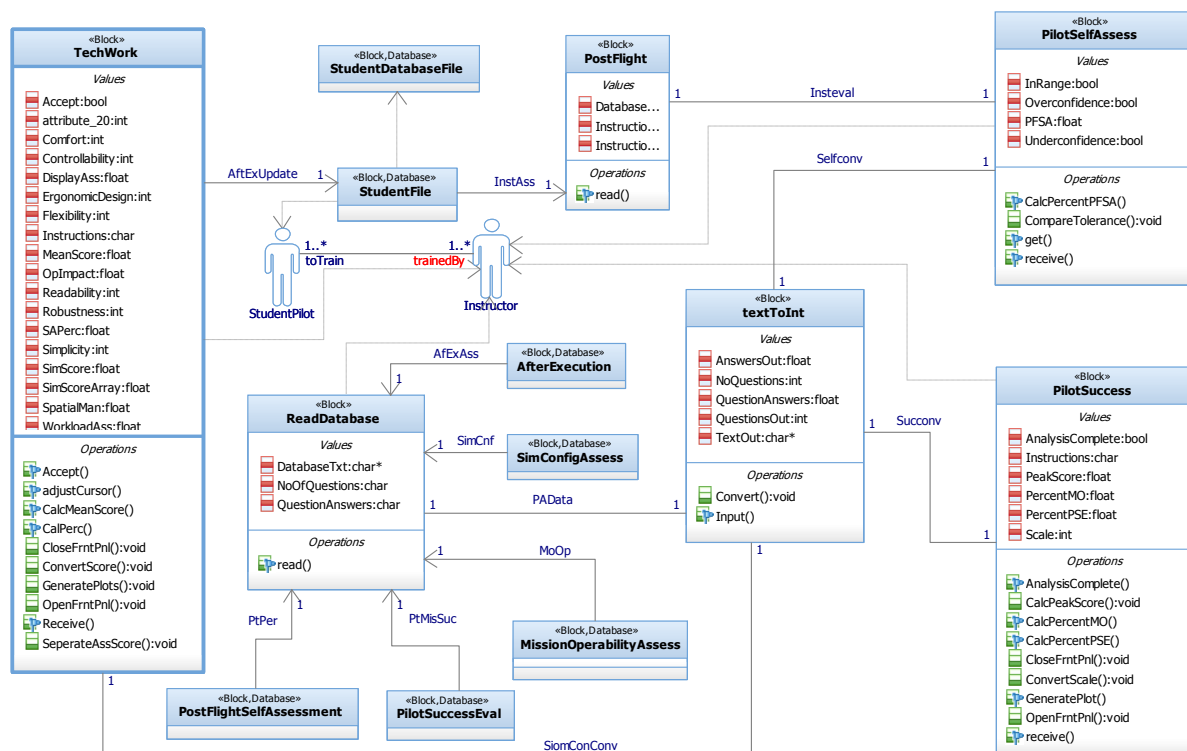


Figure 98 Post Flight After Execution Assessment Architecture

The results are automatically compared to the subjective assessment of the instructor during the observation of the mission scenario. If the results differ by ± 10 (can be amended as student pilot progresses through the pipeline), the workflow tool will display to the decision maker that there is an issue either with overconfidence of the pilot or under-confidence in their ability to perform tasks, as seen in Figure 99.

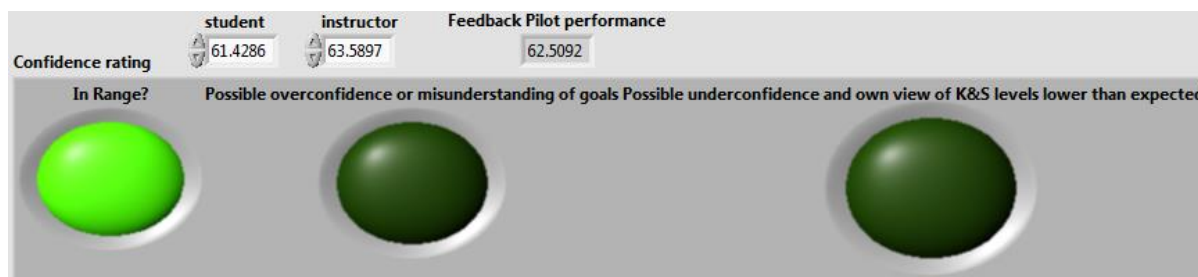


Figure 99 Subjective Confidence Rating Of Pilots Abilities

The results are easily identifiable and can form the basis of post-mission brief to examine why the scale of the difference between opinions is present. The feedback performance rating is merely the difference between the two attribute values, which is stored within the main student database file as an indication of the difference in opinions. The post-mission brief should reveal the explanation of the difference and assist in the identification of remedial actions to solve either the under-confidence or overconfidence of the student pilot, issues regarding the suitability of the chosen blending mix or the potential bias of the decision maker (instructor) when estimating performance and goal success of the student pilot based on subjective workload and SA scales completed pre-flight.

7.14.3 Pilot Flight Success Evaluation

The Pilot Success evaluation considers the student pilot's opinion on the level of immersion and realism of the technology for evaluating the training mission scenario and concerns the use of blending mixes for training. The blending mixes attempt to utilize an array of features to convey information through different means (e.g. sounds, imagery) and to create a realistic context. To prepare the student pilot for training, the pre-flight brief should be used to familiarise the pilot with the constraints of the technology to manage expectations and give confidence that the technology is suitable for acquiring relevant levels of K&S. The evaluation uses the student pilot responses to a questionnaire, described in Volume II, Appendix V Assessment Form 9 to obtain a subjective opinion on the information richness and realism with executing the mission scenario on the chosen blending mix. The level of richness concerns the technologies capability of supporting a high degree of interactivity

required for training in high performance environments. Of keen interest in the feedback is the opinion on the suitability of the technology for the successful completion of the tasks in the mission scenario.

The method of assessment concentrates the student pilot's opinion on set questions that adheres to a Likert type scale with language descriptions and metric values associated with each point on the scale. The scale is between -3(strongly disagree) to +3(Strongly Agree) with 0 being neutral. The total average rating is used to compare other subjective assessments along with actual performance outcomes; an average score less than neutral should initiate further investigations including comparing assessments from other student pilots in an attempt to find any correlation in results to identify the reasons for the low ratings. The average score is visualised to the decision maker using a triangular distribution as illustrated in Figure 100, with the total average percentage saved within the main student database file. In the illustrative example shown, the average suitability of the technology is between neutral and somewhat agree, with reference to the questions asked. The workflow tool should allow for the overlay of other student evaluations to appear on the graph for ease of correlation analysis. The behaviour of the architecture in Figure 98 is described by the 'Obtain Pilot Success Evaluation' Use Case in Volume II, Figure V 3.

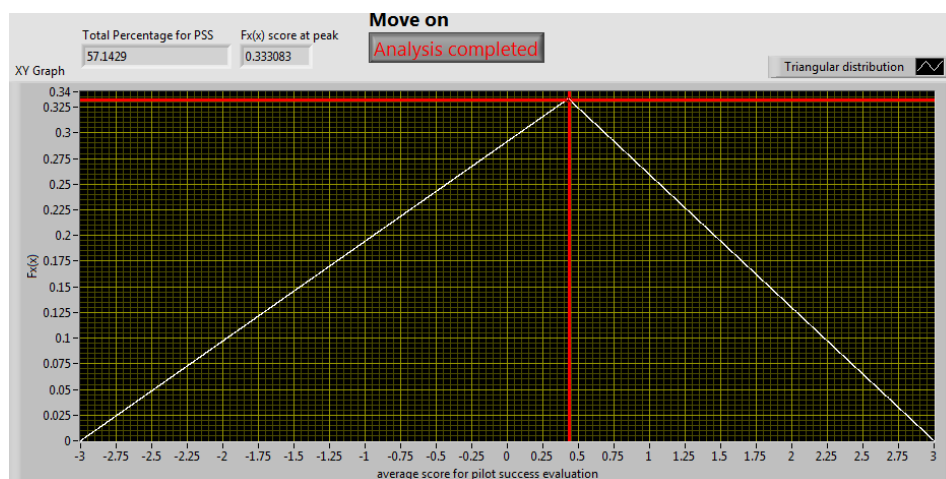


Figure 100 Total Average Rating of Post Flight Success Evaluation

7.14.4 Mission Operability Assessment

Mission operability concerns a holistic approach to assessing pilot workload and blending mix effectiveness using feedback from the student pilots to two very simple questions, described in Volume II, Appendix V Assessment Form 10. The assessment is based around rating the two dimensions using a choice of four sentences, which concern the tasks of the

mission scenario with respect to the use of the technology configuration. The questions are numbered 1(extreme/inadequate) – 4(low/capable) with the technology effectiveness giving a clear indication of adequacy for task execution. This questionnaire should to be used with others, described below, to provide feedback of the suitability of technology. The total percentage for each dimension is stored within the main student database file. This is an optional assessment that can be used independently for quick interpretation of suitability or used to strengthen or correlate the subjective assessments completed by the same student pilot. The behaviour of the architecture is described by the ‘Obtain Mission Operability Assessment’ Use Case in Volume II, Figure V 4.

7.14.5 Paper Based Assessments

The sequential judgment scale developed by (Pirella and Kappler, 1988) was used to rate the difficulty of vehicle handling tasks experienced by a driver, however, the rating scale measure can be applied to multiple domains and a much wider variety of tasks. As subjective difficulty rating is seen as one of the most relevant dimensions of workload, the interval scale properties permits the use of a parametric statistics on rating data. The rating is facilitated by the two level scale design that requires the student pilot to follow a flow and make judgements based on previous judgments made at the preceding level. The first judgment is made based on the description that concentrates on the effect of monitor numbers and size within the same block on the level centred on three basic categories: difficult, medium, or easy. The choice graphically switches subjects to the next level of specific instructions then onto a third scale based on numeric rating that is established on the constraints from the first level; and then to a smaller appropriate metric scale. The student pilots are permitted to cross over to the adjacent scales by virtue of the numeric rating system blending across the level two descriptors. Consequently, the ratings can be made with workload of two short scales but with the precision of a longer scale. The Spatial-Manipulation flow is illustrated in Volume II, Appendix V Assessment Form 11, where the rating is saved within the main student database file and uses the architecture described in Figure 98; with the behaviour described by the ‘Obtain After execution Assessments’ Use Case of Volume II, Figure V 5.

The modified Cooper-Harper scales used for this research retain the decision tree design of the original but the descriptors are changed to more suit the activities and domain of flight training (Miller, 2001). The Cooper-Harper scales are designed using the benefits of flowcharts to assist the student pilots in choosing workload ratings based on description

constraints that also identify whether improvements in the design or configuration of the training technology requires investigation. The Bedford-Harper scale (Wang et. al., 2013), utilises the identical decision tree of the Cooper-Harper scale but is centred on workload measurement by assessing spare capacity of pilots. With the scale designed as a simple flow chart with simple descriptors, the interpretation of workload is made with ease of reading the workload descriptions. These ratings should be given by the student pilot immediately after the completed mission scenario exercise whilst the experiences are still fresh in memory. Obviously, on evaluation of the assessments if more information about the specific tasks is needed a further interview with relevant student pilots is an option.

The Cooper–Harper rating scale is provides an unsophisticated approach in determining the workload of a task, where a rating of 1 indicates excellent or highly desirable technology characteristics and 10 indicates the technology characteristics in relation to the task requirements have major deficiencies. The student pilot is asked various questions within the flow chart that requires simple ‘yes/no’ answers. If the answer is ‘Yes’ then a further question is asked; however, if the answer is ‘No’ an indication to potential modifications required by the technology is given followed by various descriptions that the pilot has to decide which best describes their experience. When combined with additional system qualities, such as display arrangements, the Cooper-Harper scale has been found to be sensitive to psychomotor demands on an operator, especially for aircraft handling qualities (Hanson et. al., 2014). The scale can be easily amended to include additional cognitive functions such as, perception, problem solving, and monitoring, it also produces a quick feedback response due to the simplistic rating system and is easily integrated into a flight training system.

A number of disparate Cooper-Harper scale designs are presented by the methodology as a student pilot feedback mechanism on the configuration and suitability of the blending mix, given the training levels of K&S required whilst executing a mission scenario. There are five different types of Cooper-Harper scales described:

- Bedford Scale that uses a uni-dimensional scale that ranks whether the tasks could be complete with the technology (directly related to workload measures). For details on the Bedford scale, see Volume II, Appendix V Assessment Form 12.

- Modified Cooper-Harper scale that uses a uni-dimensional scale that ranks the technology's ability to provide sufficient capability to prevent degradation in pilot SA. For details on the SA scale, see Volume II, Appendix V Assessment Form 13.
- Modified Cooper-Harper scale that uses a uni-dimensional scale that ranks the display (size, configuration, refresh rate, etc.) capability to provide the student pilot with the visual feedback cues they require for decision making tasks. The results of the assessment can be compared to the spatial-manipulation assessment for concurrency. For details on the display assessment, see Volume II, Appendix V Assessment Form 14.
- Modified Cooper-Harper scale that uses a uni-dimensional scale that ranks the readability of the displays to provide visual acuity feedback to the student pilot for prevention of additional focus effort (symbolology, readability, visual workload, etc.). The result of the assessment should have concurrence with other display assessments. For details on the readability of display assessment, see Volume II, Appendix V Assessment Form 15.

To emphasize the student pilot feedback on the operational impact using the training technology, the Cooper-Harper scale is evolved to include a further six criteria dimensions that need to be assessed with identical rating constraints once the metric rating has been decided, the scale can be seen in Volume II, Appendix V Assessment Form 16. There are set limits on the desired level of grading criteria for the evaluation, the six criteria dimensions consider other operability factors in the assessment of operational impact, as shown in Volume II, Figure V6. The operational impact statements and grades are integrated with a flowchart applicable to the assessment criteria. A numerical result and a textual description of strengths and deficiencies are given for clarification. There are six major operability parameters: simplicity, flexibility, robustness, controllability, comfort and ergonomic design. The simplicity parameter refers to the technology capability to provide a simple, clear and consistent understanding on the technology functions and interfaces with respect to a reasonable number of tasks and processes that needs to be performed using the interfaces integrated into the technology. The operational tasks should be simple and efficient to perform on the technology allowing the pilot to concentrate on decisions to be made rather than the detailed operational sequences to be performed, which would increase the workload levels. The complex association of individual design characteristics with operational factors can be seen clearly in Figure 101.

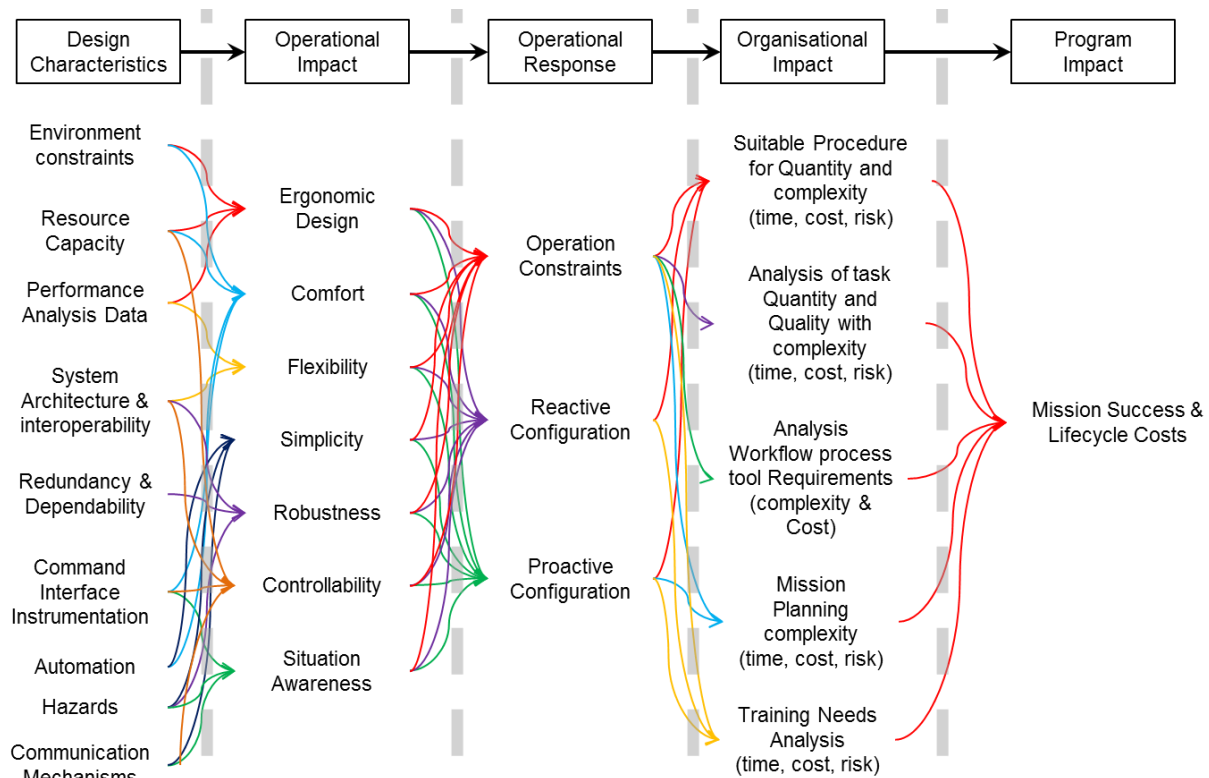


Figure 101 Design Characteristics and Operational Factors (amended from Crocker, 2010)

The flexibility parameter denotes the system's ability to cope with changes in tasks of the mission scenario or the ability of the technology to change configuration or to accommodate an update of subsystems to suit operational needs. Furthermore, high flexibility incurs the need for additional retraining and probable product development involving additional TNA assessments.

The robustness parameter signifies the technologies ability to cope with changing conditions for both nominal and off-nominal operations. When a component or subsystem develops a fault condition, the degree to which the whole system can continue to function is determined by the systems robustness. The tasks and goals in the mission scenario, especially in the event of a planned fault condition, bound the systems operational context and thus provide support for assessing robustness.

The controllability parameter represents the degree and difficulty with which the student pilot can direct systems performance during operation. Included in the assessment is the level of control the pilot must exercise and the ability for the pilot to become aware of control capabilities upon configuration change. The system should behave in a repeatable and predictable manner to each command.

The comfort parameter refers to the capability available to the pilot beyond which is required to execute the mission successfully and minimise operational constraints. The assessment should also consider a negative impact on performance with the technology and its environment to the concentration level of the student pilot for completion of tasks in the mission scenario. Environmental factors and pilot seating can increase workload and confusion and the injection of fault conditions during the mission scenario, which could cause premature abortion of the mission.

The Ergonomic design parameter refers to the capability of technology to provide controls and visual feedback mechanisms within human reach and sight limitations. The parameter has a dependency on the comfort parameter, but the assessment considers the layout of controls and the amount of movement required for the pilot to action tasks within the mission scenario.

The data collected from the after executed assessments are stored with a pilot subjective assessment database file within the Flight_Analysis directory with an illustrative example of the outputs of the assessments shown in Table 20.

Table 20 Database for Paper Based Assessment

Name	Flight Mission Reference	1. Spatial Manipulation	2. Workload Assessment	3. Situation Awareness	4. Display Assessment	5. Readability of Displays	6. Mission Scenario OP. Impact	6.1. Simplicity	6.2. Flexibility	6.3. Robustness	6.4. Controllability	6.5. Comfort	6.6. Ergonomic Design
TrevorH	Oscar123	9	4	5	6	4	4	5	3	2	2	6	7

The decision maker will not necessarily see this database, as the workflow tool should automatically interrogate the file to display to the decision maker the percentage score of all the paper based assessments, as seen in Figure 102. On the UI, are the current student name and the mission scenario reference number along with the percentage score for each dimension assessed on the paper based system. The percentage scores are stored within the main student database file and low scores should initiate further investigation of the paper assessments with possible interviews with the student pilots. Additionally, group assessment scores should be investigated for correlation analysis as technology that consistently scores low may not be suitable for evaluation of the mission scenario nor practicing for the relevant K&S needed for ToT. This along with the feedback data can form the basis for acquisition decisions and evolution of training technology used in the FTS along with elimination of blending mixes deemed unsuitable for training K&S for mission readiness.

Spatial Manipulation % 64.29	Name to Search (SurnameINITIAL) trevorH	Operational impact Simplicity% 50.00
Workload Assessment % 40.00	Mission Scenario Reference Number Oscar123	Operational impact Flexibility% 30.00
Situation Awareness % 50.00		Operational impact Robustness% 20.00
Display Assessment% 60.00		Operational impact Controllability% 20.00
Readability of Display % 40.00		Operational impact Comfort% 60.00
Operational impact % 40.00		Operational impact Ergonomic Design% 70.00

Figure 102 Output Example from the Paper Based Assessments

7.14.6 Simulator Configuration Assessments

The fidelity value of training needs to be sufficient to permit student pilot assessment with assurance that observed behaviour will transfer to the real world aircraft. Student pilots and SMEs must be able to differentiate between the technical and human factor criteria of the blending mix and its use for training K&S. For the purposes of the methodology and to gather more detailed information about the technology used to execute the mission scenario the simulator configuration assessment, see Volume II, Appendix V Assessment Form 17, investigates the student pilot's subjective views of the layout and configuration of the technology when it relates to the successful completion of the tasks in the mission scenario. Dimensions under investigation include:

- The position, type and stability of cockpit chair to permit the pilot to sit comfortably within easy reach of all relevant controls.
- The configuration and resolution of the displays including graphical representations of objects on the displays suitable for prompt decision making activities.
- The layout, type and emulation capability of the controls used for ToT capability.
- The temperature, humidity, lighting and noise of the surrounding environment.
- The software that drives the simulation - is fit for purpose.

The questionnaire is designed using a Likert type scale that ranges from -3(unacceptable) to +3(excellent) for each question and a number of questions are required to be answered for each dimension. It is envisaged the student pilot will complete the questionnaire online and the workflow tool interrogates the database file were the decision maker is prompted for the student's name to gather feedback data. The data for each dimension is used to generate triangular distributions for ease of visualising the results, as seen in Figure 103.

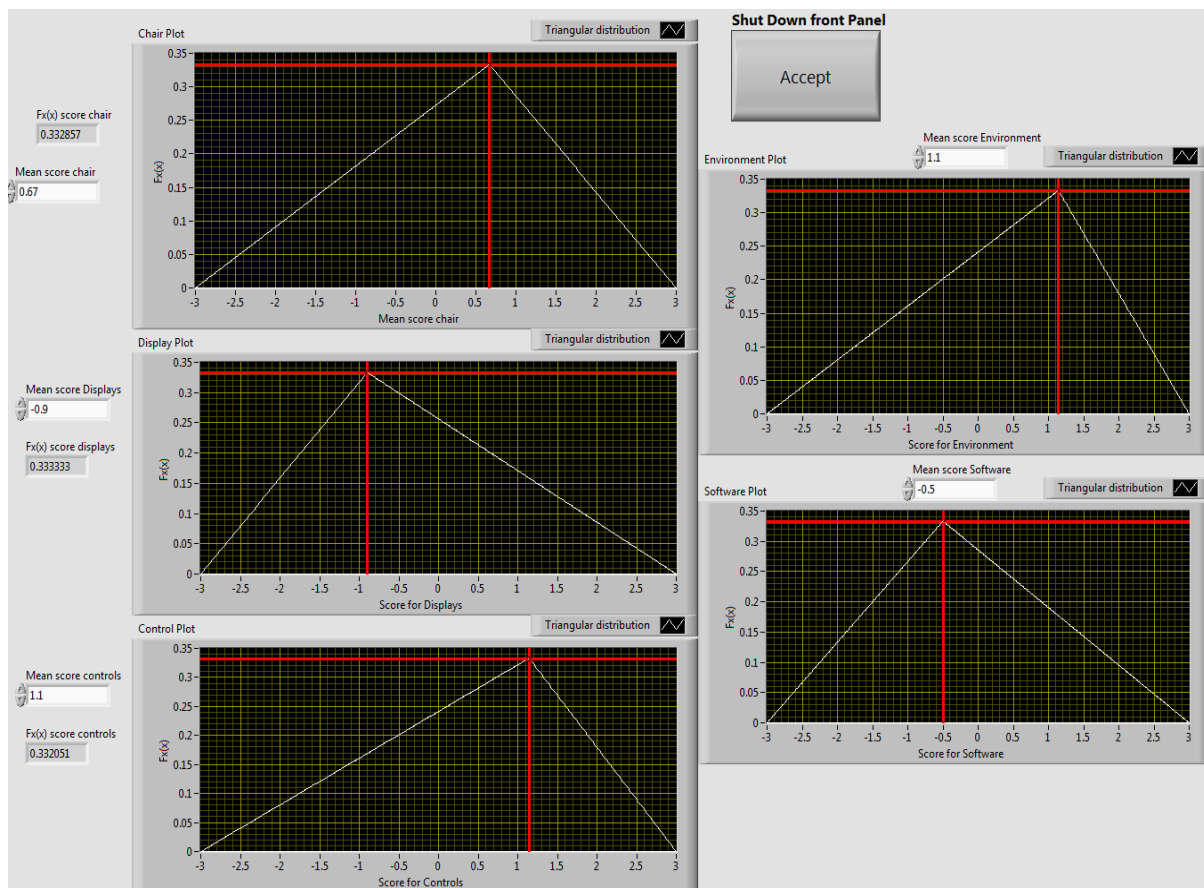


Figure 103 Triangular Distribution Results from Simulator Configuration Assessment

The mean value of the average score for each dimension is used to generate the plots to give an indication to the decision maker which particular aspect of the technology caused some discomfort and possible distraction for the student pilot in performing the mission scenario. In the example illustrated in Figure 103, the two critical dimensions that require further investigation are the capability or configuration of the displays and the simulator software, i.e. peak to the left of the neutral point '0'. (These results can also be compared to other student pilots by the means of plotting all feedback data for each dimension onto the graphs for quick identification of correlation of opinions). The data is then transformed to an average percentage score that is displayed at the side of the graph plots with the functionality to shut-down the graph display, as seen in Figure 104; the percentage score is then saved into the main student database file.

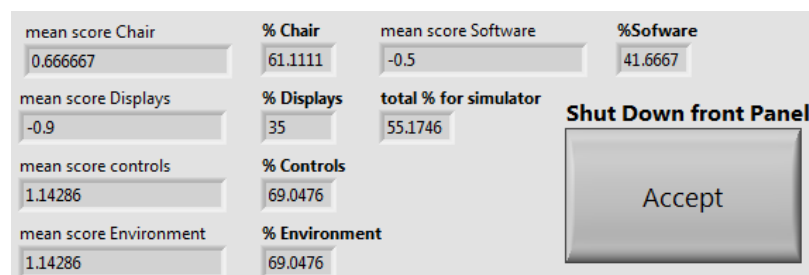


Figure 104 Final UI Example for Simulator Configuration Assessment

7.14.7 After Execution Assessment Summary

SA and workload assessments are used to determine the effect, positive or negative, resulting from proposed designs of scenarios and technological involvements with the choice of blending mix used for training. The issue with SA assessments is despite the increase use of a number of human factor techniques, there is little evidence that any of the methods actually work and are valid (Stanton & Young, 2003);(Zarghooni, 2009); hence, there is a need to assess their reliability and legitimacy. Reliability in SA refers to the degree to which the measure will generate the same data under the same scenario conditions in an ongoing training programme; furthermore the measure of SA for individuals progressing through the training pipeline should improve for the same scenario conditions as an indication of ToT with the acquisition of K&S. Legitimacy of the SA measure refers to how the measure is actually measuring SA and not some other parameter.

Workload feedback, which should be consistent to the pre-flight workload assessments completed by the decision maker, gives a strong indication of the negative effects of the mission scenario performed using the technology on pilot's decision making activities. The perceived workload can be seen as aspects of K&S that are weak in individuals, which they themselves did not initially recognise before executing the mission scenario. However, in many cases individuals who believe they were unsuccessful in achieving goals will inherently attribute the cause to third parties i.e. the training technology. It is crucial that concurrence analysis between the groups of student pilots is completed to obtain a more accurate representation on the influence of the technology to mission scenario task outcomes. The outcome of the comparison analysis between pre- and post-mission subjective assessments can be used to debate the use of technology configurations for training K&S as the training programme evolves overtime.

The number of post-mission assessments can reduce once the feedback has allowed the relationships stored within the ROSETTA frameworks to evolve to a degree that confidence is achieved on the shape of the RSEs that describe the relationship between training parameters. The confidence weights placed on the DSS for analysis will take a period of time to optimise to ensure concordance with organisation objectives. The more detailed analysis using the simulator configuration assessment can then be used when a new system is added into the FoS to determine suitability of training various levels of K&S on the technology.

7.15 Workflow Process Tool Summary

Each assessment stage generates data that is then saved into relevant directories of the database architecture with a summary of the results saved within the main student database file situated in the Student_Pilots directory. In the methodology, each student pilot is assigned all directories associated with each assessment stage, as per Figure 39. The contents of the main student database file is used for quick identifications of any issues, which can be further investigated using data stored in the slave directories containing all data for each assessment stage. The content of the main student database file is described by the header information shown in Table 21.

Table 21 Header Content of the Main Student Database File

Header Names					
Student Name	Male/Female	Age Rating	Pre-Study Total	Flight Hours	Flight + Simulated
ILS Score	Student Rating	% Personal Allocation	Mission Scenario RefNo.		
Mental Effort	Physical Difficulty	Time Available	Understanding Horiz. Position.	Time Criticality	Usefulness of Information
Pilot Specific Rating	% Workload	%Workload	% PARS	% Understanding of Mission	% Understanding of SA
ROSETTA 2 Tech. Choice	Task Load for Mission	% Pre-Flight	Mission Mean Goal Success	Ave. % Instructor flight evaluation	Pilot Performance Grade
Feedback Pilot Performance	Pilot Flight Success	% Mission Operability	Ave. Spatial manipulation	% Pilot Workload Scale	% Pilot SA Scale
% Display Configuration Assessment	% Readability of Display Rating	% Operational Impact			
% Simplicity	% Flexibility	% Robustness	% Controllability	% Comfort	% Ergonomic Design
% Chair	% Monitors / Display	% Controls / Interfaces	% Environment	% Software	% SimConfig Total

The database file is in the form of an excel spreadsheet with each new mission data from the assessment stages stored on a separate row to the previous one. The decision makers can also use the spreadsheet file as a quick reference or summary of previous assessments and performance results if required, although the proposed tool can permit the facility to display results on a UI. The UI may be beneficial to collect the data for each mission as throughout the student pilot lifecycle this file can become large.

The slave directories are intended to store large data within text files since the files are not intended to be viewed independently to the workflow tool. This is to allow the conversion of the textual data to visualisation data for ease of analysis. This notwithstanding, saving the

data as a text based system using header information within the files is designed to allow for human readable data without the need for a tool to visualise the information. There is an argument with respect to the baseline simulation architecture to save the data in binary format which would allow the software simulation to reduce memory resources and decrease the generated file size, though, for flights which are scheduled to last less than 10minutes the flight data is comfortably stored within arrays of the simulation that are then converted and saved as a text file. For flights greater than 10minutes, binary output files would be recommended, consequently, the file would not be able to be human readable outside the workflow process tool or other tools which permits transformation of binary to text data.

Throughout Chapter 7, the theme has been to present data to the decision maker in a way to make assessments and evaluations simple to identify aspects of the data that would not be obvious using static spreadsheet data. Data visualisation is inherently easier to interpret, saves time and produces clarity to factors that require careful analysis within the data set. Each assessment stage has the property to generate volume data that is used in arrays, which are transformed to textual spreadsheet data for ease of storing, however, the templates used to save the data provides a structured format that is used to direct the decision maker to relevant regions of interest, e.g. the first column and row being used for header reference and line spaces to allow the data to be visualised and searched through in an easy way. However, occasionally displaying metric data in tabular form does allow quick identification of the impact of the evaluations. Using abstraction techniques to present metric data in the assessments, i.e. using the ROSETTA stages for elimination of technology function to simplify the result as an example, permits the holistic view of the results of the evaluation to be displayed; when placed beside each other it does accommodate the viewer to investigate why any differences from the expected result, based on subjective opinion, are present. The actual generated data from the assessment can then be investigated in a graphical format to check understanding of the evaluation criteria used to generate the results, thus, can be used for early evolution of the relationships between framework parameters.

For any software driven management process, it is always prudent to develop an alternative approach to plan missions and administer pre-mission assessments on the student pilot and the technologies, which can then be inputted into the software databases at a later stage. Hence, to account for the general unreliability of network services that locks out users from the system it may not be possible to plan the missions or assess workload, SA or assign

training attributes to mission tasks using the software tool. Therefore, the methodology also offers a paper based record that can easily interface with the workflow process tool once computer or network services are resumed; an example of the paper-based planning system is seen in Volume II, Figure H 29. Each task is assigned a separate planning sheet and covers aspects of planning, such as: relevant MECs and K&S, flight plan details, a list of blending mixes and workload dimensions. Included on the planning sheet is additional data concerning cues in the simulation or flight which are of interest, cockpit operations required to be performed by the pilot and the number of trained SoPs available for the student pilot to achieve the task objectives. There are additional areas of technology specific characteristics that can be assessed in relation to human factor issues. The information contained in this paper based system is not overly excessive, but will take longer to complete than the proposed software based tool, which directs the decision maker to complete the information required in a sequenced format; as such will be easier to achieve evaluation results and reduces the issues of which information to enter first, where the information to review is, which can cause disorientation in many document based information management systems.

SUMMARY AND CONCLUSION

This chapter summarises the research project described in this thesis and discusses the development of the methodology for the decision support system integrating the ROSETTA framework to identify causal relationships between system elements of the flight training system (FTS). Included in the discussions is a synopsis of the measures used for human, technological and organisational factors involved within the methodology to work towards solving the problem of choosing the correct blending mix, along with the current limitations of the design. The chapter concludes with brief suggestions of how to further define evaluation criteria that can be used to track progression using the available systems from the FoS relating to awareness and finishes with some final remarks of the research area.

8.1 Summary of Research Project

The thesis commenced with a discussion of the about why research is needed in flight training systems with details about the complex socio-technical interaction between blending mixes. The SE methodology led to discussions on the complexity involved with the use of disparate training technology/systems to achieve a level of readiness. The natural path from this discussion steered to further investigation into the technology characteristics involving measurements of fidelity and performance indicators using the training systems. This allowed an anticipated discussion on methods for the development of a DSS and associated ROSETTA frameworks using mathematical relationships to account for the multi-attributes needed to assist the decision maker in choosing the blending mix.

Analysis of the requirements narrative was accomplished that produced new knowledge regarding the FTS problem. This knowledge was then used to develop the methodology, the systems architecture and the various assessments used within it accounting for all the attributes of concern permitting traceability to the requirement models to assist in management of the decisions involving the choice of which blending mix to use at which point in the student pilots lifecycle along with a high level functional verification of the proposed workflow process and tool with example outputs described throughout Chapter 7.

The research focussed on developing a methodology to assist in choosing the appropriate blending mix (training technology) for a specified mission scenario at a specific point in the pilots' pipeline lifecycle. Abstraction techniques have been used to concentrate on the

attributes of interest to provide full traceability between predefined training attributes and mission tasks. The training attributes are used to assess the capability / suitability of an available blending mix solution to a specified level of K&S delineated by the decision maker. The proposed conceptual model provides a closed loop feedback system to account for the performance outcome of the mission scenario using the chosen blending mix first to a baseline scenario (simulated by mathematical algorithms) to identify if any issues exist within certain tasks of the mission and identify which training attributes are involved in the particular task that might have caused performance deviations from predicted, and then by using subjective and objective feedback mechanisms to gather further data from the pilots using the mix.

Included in the thesis are methods relating to advanced SE techniques to assist in development of design solutions of any system where uncertainty exists. The requirements analysis phase concentrates on analysis of vague natural language sentences to obtain knowledge on the requirement specifics of the system, which are then transformed into requirement models to obtain knowledge not apparent or missing from a given requirements narrative for a clearer view of the problem space / domain. These logical semantic models are used for precise mathematical system requirements description to bring precision to the natural language sentences. This method is also used when there is a possibility of misinterpreting the natural language sentences. The modelling ensures robust mathematically provable arguments, i.e. true/false; that can be used to communicate to stakeholders the exact interpretation of the system requirements from the systems engineer's viewpoint and therefore create requirements that can form the system description using MBSE techniques. The Use Case modelling techniques describe a manner to model functional requirements but also to ensure that use case explosion is avoided. In addition, the method describes an approach for decomposition of the functional requirement to enable ease of communication to all stakeholders.

The mathematical based formalisms used within the conceptual model of the workflow are to assist in gaining knowledge regarding uncertainty within the system. The method is based on the RSM within the ROSETTA methodology for understanding of the system in context, using the approach of iteration and recursion through the systems design process and evolution over the lifecycle of the system to evolve the RSEs to optimise the relationships between parameters in a multi-attribute systems problem; thus, assist in gaining knowledge of

the relations between parameters of the system for the purpose of understanding how system entity attributes affect the overall system goals. The novel software information management structure and feedback systems, including the functionality of the workflow process and frameworks, offers new methods of database management and information analysis techniques to assess the choices made using the DSS. These methods, especially how the information is presented on screen, can be used when 'big data' analysis is required to simplify data using abstraction techniques to concentrate on relevant areas within the data set.

8.1.1 Summary of Satisfying Research Objectives

Humans are inherently irrational decision makers where the state of mind can change throughout the day and provide inconsistencies in actions and performance and hence liable to fallacies in decision making tasks, therefore, it seems a natural progression to quantify the type of decisions being made. Decision making often involves subjective judgement on aspects of the system where uncertainty exists and is the reason why SMEs are needed to populate attributes of concern with knowledge from an operational context, which is used to assess various expectancies from the performance of the system (technical and human). The naturalistic decision making approach (Chapter 3) is the description of a formalised decision making strategy that emphasises recognition over an analytical process for decision making tasks and for this reason why it has been used as the most representative model that has been adopted for this research project by virtue of the requirements analysis phase being used to recognise the 'full picture' before the development of the FTS conceptual architecture model used within the methodology (**Req1**).

Decisions with HITL systems require the understanding of physiology and psychology for the development of feedback systems to evolve the attributes of the DSS including the ROSETTA frameworks, to maintain a truthful / legitimate representation of current capability and experience levels thereby inheriting a recursive pattern to the decision making process that will affect the choice of blending mix (**Req2**). In complex organisations the decisions are multifaceted and have to be based on a number of decision levels, which sometimes involves contradictory objectives. The objectives were accounted for by quantifying the relationships between objectives and performing trade studies on them in a formal manner to allow for documentation and review that could add value to SMEs knowledge acquisition, understanding of technology suitability to student pilots and training (**Req2.1; Req3**).

The complexity involved in the FTS ensured that the decision making process would be far too convoluted to be accomplished within one DSS ROSETTA framework, hence, the workflow process is used to separate objectives for ease of understanding, i.e. training and organisational objectives are used independently in trade-studies ensuring that training objectives are satisfied (training technology/system suitability to training levels) before relevant organisational objectives (cost, availability, DOC) (**Req4**). This type of sequenced decision making ensures that the main training requirement of the FTS is satisfied before any organisational requirements are considered with respect to the FTS. The quantification of the result from the DSS ROSETTA frameworks can then be used for justification of the choices made embracing the context of continuous improvement of the RSEs for knowledge acquisition for the SMEs, training designers and technology developers (**Req2.1; Req6**).

8.2 Review of Research Problem and Novelty

An extensive literature review in the areas of flight training and decision support revealed current design practices involve the collection of requirement sentences within a spreadsheet style database (i.e. doors) that are very complicated to navigate through and efficiently map to system architecture elements, which require considerable familiarity training before individuals can effectively navigate them. It became apparent that management of requirement statements, communication to stakeholders, and database management techniques are important considerations for the development of the DSS for mathematical analysis of objective judgements. Therefore, MBSE techniques has been used with a basic ontology model to more formally model the requirement statements, which provided additional information regarding the interpretation of the sentences that can be used to communicate back to the stakeholders and assist in gaining additional knowledge to ensure that the ‘right’ system is being developed. The meta-models offer an additional advantage to obtain a bi-directional understanding of the problem domain; these are produced using formal semantic relationships with a combination of four invariants, which make up a scenario that describes the system in a context relating to robustness in design (**Req1;Req2**). Consequently, a common ontology is used to communicate the system description to assist in resolving interpretation issues and control the complexity of the architecture artefacts used in the model (**Req1**).

The workflow tool allows for directing the decision makers into entering relevant data for the production of the mathematical relationships between ROSETTA framework parameters

(Req3.1). This focusses the decision maker on the transformation of subjective judgement to an objective quantitative metric that incurs the ability to have an attribute with more than one value associated with it by the virtue of the RSE method used to perform trade-studies **(Req3.2; Req6)**. As these relationships are documented within specific directories of the database architecture (to allow for ease of review), the decision makers are fully aware that the progression and performance with an associated training system is fully documented and grading might require justification if there is an inherent disparity in predicted and actual performance outcomes. The identified relationships, developed by the SMEs, can also be used to gather further knowledge and understanding of how a particular technology suits an individual, thereby, providing evidence of suitability in quantitative form assisting in determining the correct blending mix to use at various points within the training pipeline **(Req3)**. The evaluations can also be used to further quantify fidelity characteristics based on a combination of subjective and objective assessments from both the SMEs and the users of the technology (student pilot) to further clarify aspects of technology that require improvement or is deemed unsuitable for K&S acquisition for practicing and procuring essential training attributes **(Req6)**.

The parametric approach used in the research involved various dimensions of attributes, vectorising experience hours along with a new method of calculating task load that required to be integrated into the software based process workflow tool. The difficulty within the FTS is quantifying the relationships between training attributes and technology characteristics along with non-functional worker oriented attributes (K&S) with functional oriented training attributes (MECs) relating to the available training systems. Using MDA techniques, the methods to separate the concerns of the two problems generated two different dimensions of K&S, one relating to the MECs mapped to the training systems and one with the technology characteristics mapped to organisational objectives **(Req1)**. The two dimensions permitted a natural way to evaluate the relationships, via mixed method approaches, and allowed a coherent representation of the relationships without merging concerns, thereby, offering a novel way to gain knowledge of how the K&S levels affect the attainment of MECs using a chosen blending mix for mission readiness. Assigning relationships between K&S and MECs for each training technology/system and performing trade studies on the suitability of each technology for evaluation of the required K&S levels, as per ROSETTA 1 stage of the workflow process, can provide propositions on which cost effective training system/technology can be used to satisfy training outcomes; thus, provide an enhanced and

traceable blending mix solution method with the advantage of gaining knowledge of how the non-functional worker oriented training attributes are related to the functional with respect to systems of the FoS.

The experience vector uses real simulator and flight experience hours to alter the rate of change of a curve to integrate within the ROSETTA framework to allow for interoperability with other attribute parameters (See Chapter 7.6). This method of measuring experience allows for visual representation of the level of experience the student pilot has; the curve can be used to encourage the decision maker to identify the correct levels of K&S and MEC the pilot should be capable of achieving before considering issues with suitability of technology to student pilots. Disparity between the experience curve and the training levels prompt for further investigations into either the capability of the technology or the ability of the pilot (**Req2.1**).

The designed questionnaires used to obtain student pilot feedback on the training system integrates the strengths of current SA and workload assessments but allows more focus to be directed at evaluating the technology used for training (**Req2**). Without the specific feedback data, the solution of which blended mix to use will be increasingly difficult to resolve. All aspects of the technology from an 'end users' point of view is assessed and the data stored in the databases are then used to check the decision maker's own judgements on the suitability of the training technology/system on training the required levels for ToT objectives (**Req3**). The evaluation of the data is used to evolve the relationships contained in the ROSETTA frameworks but also to allow the decision maker to re-evaluate their opinions on suitability based on the subjective and objective feedback data from the student pilots. This closed loop system presented by the methodology permits constant evolution of assumptions and decision making tasks at the individual level that is hoped will provide a more efficient blended mix training solution for efficient training outcomes (**Req4;Req6**).

8.3 Significance of Research

The design methods used should open up new frontiers in systems design methodology to account for the problem of imprecision in HITL systems by formalising feedback into the databases, which are then used to evolve the surrogate models to make more precise judgements about reality. In this way, the 'definitive' solutions will not be accurate for a number of generations and should be seen as an opportunity for solutions; nevertheless, the

results obtained from the DSS should assist in optimising blending solution choice irrespective of the information content stored within the ROSETTA frameworks (**Req3**). The constant evolution of the relationships contains data regarding lack of understanding and knowledge deficit in the organisation that is used to comprehend the areas of uncertainty that exist within the problem space. Identification of these areas of uncertainty provide additional research opportunities to develop methods and techniques to help manage and govern future DSS for large complex systems problems.

The goals of the research problem and proposed methodology can be exploited to define additional areas of interest to improve efficiency of the FTS involving requirements, such as:

1. A more robust and easy to understand method has to be developed to make possible prediction of how difficult planned mission scenarios will be to successfully completing all task goals, using all available training tools.
2. Using M&S, the decision maker should encounter the ability for rapid evaluation on workload demands of all pilots for a particular training system that will lead to realistic prediction of pilot's flight control accuracy throughout the planned mission scenario.
3. All relevant factors for design approach, which influence pilot difficulty, needs to be identified and placed as attributes within relevant classes / blocks within the M&S environment.

The models developed from the analysis incorporate levels of quality and rigour of design. The quality ensures that the design is suitable for addressing the research question and the levels of abstraction with the design method used captures meaning, effects, and relationships required to evaluate the data in addressing the research problem. Testing the validity / legitimacy of the model and associated data is incorporated into the theory of perceived reality. The model and data sets used in mixed methods research, for a real system, is seen for the purposes of this research project as a continuous process rather than being a fixed attribute used for the purpose of study. The quality of inferences and conclusions of the results might never be reached due to missing or misunderstood knowledge used as the basis for design; however, this type of approach requires iteration of the data and possibly recursion of the design method. The systematic evolution of the model(s) is expected to occur inviting new data from both approaches and thus new knowledge being acquired each time the system is evaluated.

8.4 Are the Developed Frameworks usable in Practice?

The extent of the amount of effort, cost and time to implement the decision support system (DSS) using the ROSETTA frameworks described in this thesis is dependent on a number of factors including to what level of knowledge and specification of technology is needed to assist the decision maker in the selection of the appropriate blending mix at a student pilot's current position in the pipeline. The perceived requirement to manage analysis and the measurement techniques could be based on a periodic assessment throughout the systems lifecycle since most of the mathematical relationships used for trade study analysis are updated from knowledge gained from the feedback systems. The initialisation of the ROSETTA methodology is time-consuming and human-resource intensive; the level of effort in conceptualising the DSS should be decided on based on how complex and ad-hoc the current system is. The constrained viewpoint described allows for quantitative and qualitative data to be integrated into a framework where knowledge and information is stored with the advantage of exploring aspects of uncertainty using RSEs; nonetheless, it may be apparent that for small FTS, level 0 of the trade-off analysis would be suitable for the decision maker to eliminate a number of systems as unsuitable for the current training mission. The evaluation of qualitative measurement methods includes monitoring or interviewing stakeholders (SMEs, pilots, designers, etc.) with information gathered being uploaded into the system model for analysis directed at the blending mix choice. Once qualitative methods have answered, the question is then of what to measure, at this point quantitative methods are then used to sample this space. Hence, data and metrics are gathered from models or meta-models, performing surveys, and controlled experiments. Obviously, to manage the proposed DSS effectively for various systems problems a number of constraints need to be met, these include:

- Adequate quality of resources;
- Adequate quantity of resources;
- Availability of resources.

This thesis has not included the availability of resources (people, economic, and technology) in any great detail within the analysis, nevertheless, this is by far one of the key factors that will determine to what degree and level the DSS using the ROSETTA frameworks and the methods described in the thesis will be used in a practical way to solve multi-faceted systems problems where uncertainty exist.

It is anticipated that the benefits of the DSS provided by the methodology include:

- The DSS should enable a better understanding of the use of technology within an organisation using feedback systems to enhance knowledge on systems.
- The DSS should provide a useful aid to navigate through complex systems, SoS or FoS to gain a better understanding of functionality and uncertainty within the system.
- The DSS is likely to provide time and cost saving of understanding uncertainty in systems for optimised decision choices in the long term (short term costs, however, is likely to be excessive for initial setting-up and populating the database and frameworks).
- The DSS should provide fast response to wicked problems within an organisation by tracing through the sensitivity relationships and performing trade study analysis for identified solutions.
- The models used to design the DSS can be considered as a living document which can be evolved upon changes in actors or systems within the FoS.
- The DSS conceptual model should lead to better communication amongst project members, using a common ontology that is easily understood.
- The DSS reduces the apparent process complexity, by giving focus on more important attributes in an abstract approach.

The methods described in the methodology are fully transferrable into solving multifaceted systems problems where uncertainty and lack of knowledge exists. The closed loop system proposed by the workflow should assist in evolving the understanding of complex systems within the FTS for all human stakeholders and thus lead to a more efficient blending mix elimination strategy based on formalising relationships between system parameters. The methodology provides the means of integrating a number of disparate methods into a seamless whole, whereby the loosely coupled design allows the decision maker to choose which assessment stage/method to use to solve issues regarding governance of a modern flight training programme and other HITL systems.

8.5 Conclusion of Research Domain & Methodology

From the research and the in-depth interview(s) with training contractors, there is and always has been a contention about the utility of certain VRTEs for acquiring the relevant levels of K&S. In addition, the training contractor interviewed for the research has made a conscious decision not to alleviate the cost and hours of real world flight training to the VRTEs but to add the VRTE training to the student pilot's workload and increase the total cost of training in real terms, due to the lack of confidence and knowledge of the ToT efficacy of such training technology/system. Though, it is perceived that this is only a short term issue that is

generational in nature; there are cultural, hierarchical and structural mind-sets that need to be addressed and most high level decision makers can be 1 to 2 generations behind the new target audience and may require convincing of the utility and effectiveness of new solution methods incorporating VRTE technology to training.

Instead of thinking of the problem as trading training hours, the training course should concentrate on training effectiveness using all the tools and resources available. A pilot can accomplish more complex tasks with a live flight hour if the VRTEs are available to do less complex things, thus, flight hours could be reduced (resulting, hopefully, in reduced training cost) whilst maintaining or improving readiness. The main issue to overcome with VRTEs is the inability to generate a 'fear factor' and as such the pilot may become complacent about mistakes. The lack of data that supports the use of virtual technology for training various levels of K&S to produce students who are ready to perform on the front line under severe stress conditions is the outstanding problem. The proposed methods described in this thesis provides a guide of how to gather enough data to produce a coherent argument, either for or against, the suitability and efficacy of advance blending mixes used to train pilots for readiness.

The degree of fidelity, learning stage of the student pilot, and the goals of the training technology/system are not mutually exclusive. As a result, the learning stage of student pilots must be continually clarified, controlled, and monitored before assessing and distinguishing a valid relationship between fidelity of the VRTE and ToT value to the real world. Furthermore, the relationships determined by SME's have to invite a degree of dynamicity, due to imprecision of human systems, through the learning stages of the student pilot. Once understood, prediction validity of student performance in an operational environment on real-world aircraft can be undertaken in a formalised manner; intuitively, the concept of learning, assessment and fidelity understanding must be viewed as complementary to training, if multiple disparate training systems are to be used. The subjective and performance feedback systems can be analysed as to the reasons why the chosen blending mix struggles to assist in gaining K&S for the ToT. In addition, the justification of the levels of fidelity actually required to train K&S levels can be questioned and aid in the discussions in the use of VRTEs and the inclusion of blended technology within real aircraft.

8.5.1 Discussions

Technology has advanced rapidly, whilst the training programmes supporting them have shown little change for the past two decades. The scientific literature has demonstrated the use of proficiency-based training over fidelity training, yet fidelity is still an institutional bias when choosing a VRTE tool to be used for training. The view that FSTDs, especially high fidelity ones, should replicate an aircraft is outdated (motion cues at present cannot be emulated to a high degree of mapping to real-world) and would definitely lead to a disappointing training outcome when balanced against the cost of such an investment. The adoption of new training strategies based on training proficiencies using the VRTEs combined with reduced real-world flight training would yield more effective training than using present training strategies based on number of hours flight (both in simulator and aircraft). Though, not everyone will accept the need for change, however, as new technology becomes available the need for change becomes apparent. Convincing people of the need for change is difficult to accomplish especially when effectiveness of the current processes is high.

The methods described in this thesis can be integrated within organisation's own workflow process and can offer substance to gather evidence to substantiate suitability of technology to work processes. The relationships within the ROSETTA frameworks focus the SMEs on capability of the technology/system to relevant organisational objectives; moreover, unlike the development of technology and how private companies sell the technology with the look, feel and functionality, the aims of the ROSETTA framework(s) concentrates on the effectiveness of socio-technical issues that relate to fundamental objectives and goals that should be at the forefront of the technology design and acquisition. This information can be used for debate before contracts are awarded for providing technology for training or other public related acquisitions.

8.5.2 Method Limitations

The assessment of the 'proof of functional concept' DSS methodology remains confined to the laboratory environment, however, the results ultimately lead back to the SE techniques described in the thesis that have been embedded in the programming and method of analysis in the DSS. The preliminary investigations of the output results of each stage pertained to the SE methods used in the design stage and hence improved knowledge of how DSS tools and techniques can be applied to much larger scale problems to improve and benefit large

organisations that are committed to solving complex problems. Nevertheless, it is recognised that there is significant latitude in the prescribed output attributes presented in Chapter 7 due to shortcoming in research time constraints and tool acquisition choice. Consequently, it can be concluded that the research approach and methodology, although applied correctly, was only partially successful in proving MBSE techniques in the context of development, utilisation and decision making to provide solutions for large complex decision making problems; although, the methods offer a solid foundation for developing a robust approach for understanding relationships between systems elements where uncertainty is involved (**Req6.1**).

To produce and use the ROSETTA frameworks requires substantial computer resources and in highly detailed systems the lag caused by the consistent updating of information in each slot of the framework may instigate a reduction of confidence and usability of the ROSETTA frameworks for performing analysis. This problem can be alleviated by modifying the functionality of the ROSETTA frameworks to give an option to show the graphs that are of main interest in the decision making activity and use discrete values in a matrix representation(s) of the framework to update values as the crosshairs in the prediction profiler are adjusted in the analysis. These discrete values identify the relationship value between the requirements and the design metrics; both the partial derivative and the current requirement level values are set depending on the location of the crosshairs within the displayed plots. This provides an easy remedy for the computer resource problem.

As a practical workflow process tool, the approach is limited in its current form due to presented methodology being of ‘functional prototype’ standing for applications in an industrial context; future research work is needed to address this issue, building upon the preliminary work outlined within this thesis.

8.6 Further Work

Additional research requires to be focused on establishing clear dependencies between human and organisational processes, human and technology factors, operational effectiveness and technological configurations; and allow for the capability to assess technology at a high level of abstraction. Specifying relationships between these causal links and thus generation of metrication practices should provide useful ‘performance indicators’ to guide the decision makers towards suitable trade-offs between training technologies for the achievement of specific organisational goals. The blending mixes along with the DSS should be developed

to enable automatic detection of the levels of experience and readiness of the pilots, via feedback techniques, and account for this in the planned training tasks using blending mixes so that each pilot is tested according to their ability levels, which optimizes training goals and minimises the time required to achieve that goal.

It is clear that a novel approach for measuring SA for each individual is needed within an exercised mission scenario using disparate training tool configurations. It is perceived that an integration of a multi-measurable toolkit is needed that will correlate with actual performance outcomes along with pilot's subjective assessments of using the chosen blending mix in the exercise for a more precise assessment of SA that will ensure reliability and robustness of decision choices. Documenting all assessments using software based tools and metrication of subjective judgements, can alleviate the potential for bias as data stored within the databases might need to be justified by the decision maker who made the assessments.

The methodology for gathering and modelling requirements is based on academic work which has been used sparingly in some industrial projects. There is a need for additional research based on industrial best practices and hard earned experience gained from modelling and developing a system using the UML or SysML tools, working with tight deadlines, profit margins and isolated departments to further see the valuableness and practicality of using such tools. Executable architectures that are fully simulate-able would be extremely beneficial if UML/SysML modelling approaches are to be directed to development of real systems. This will enable human readable models to aid communication between stakeholders and reduce the additional tools required for parametric simulations of system element behaviour that can trace directly to the requirement diagram. Unfortunately, current tools only allow for basic execution of models and simulation capability is extremely weak, consequently, expensive APIs or plugins are needed to integrate with the tools for the basic graph generation that causes configuration and timescale issues. Tool specification is an important aspect to consider when developing systems to solve complex problems; using tool(s) that adds 'accidental' complexity to the solution is equivalent to achieving undesirable outcomes for the system solution for the end user.

8.7 Final Remarks on Research Domain

It is very easy to become excited by something which looks and feels advanced and the potential capabilities various training technologies have. It is important not to become

carried away with the awe of how something looks and the decision makers should be directed to the task of determining what the key issues of the system should be, the nature of the HMI / UI and consideration of the type of technology that should be employed (to avoid vendor lock-in and costly future upgrades). Decision makers must be made to realise that the VRTE systems used for training, as well as others, are merely synthetic computer generated representations of a physical system that allows a user to interact with it as though it were real and that there are no real consequences of events, as such a critical part of performing a task in the real world is missing. Thus, the choice of which training tool and blending mix to use has to be based on the operational task the pilot needs to perform and which training attribute measures are being assessed at the time of execution.

There is limited data involving the training value of simulation for air combat manoeuvres although there is no shortage of opinions, generally, from SMEs who have differing views and who have been training primarily on real aircraft. A number of criteria can be used in the military context to gain more foundational data:

- First, the reactions gathered from feedback data about the specific objectives and goals given in the instruction can give an indication of what the purpose of the training technology is intended to achieve;
- Second, concentration of the specific learning material and grading contexts are important: is the main parameter successful completion of tasks in missions or a consistent improvement of performance in disparate missions using disparate blending mix configurations?
- Third, the measurement of transfer of learning from VRTE to live training to meet organisational objectives: this is generally associated to cost and proficiency.
- Fourth, produce a criteria and grading scheme that will assist in the iterative and recursive nature of training sorties and technology usage; this can be based on the measurement of performance and the ability for conflict resolution in the decision making process.

Advantages of some of the low fidelity VRTEs, such as performance measurements, offer data collection that the more expensive training environments do not provide. These performance measures can be used to adjust training programmes for individual student pilots to gain the best levels of readiness for each. Transparent feedback, when provided to student pilots, can allow them to adjust their own performance accordingly and even alter their own

mental models to assist them in decision making during the assessment stage. Simply providing a practice mechanism will not by itself improve the performance of the student pilot, but feedback and different methods of instructor led training may yield positive results.

One of the key issues with using VRTE for training is the lack of additional stressors to the participant during a training mission. Of extreme concern in current modern FSTDs is when an aircraft crashes (or 'hit') the aircraft appears in the sky undamaged, which reduces the experience of casualty or the appearance of risk. However, one of the unique features of most VRTEs is the ability to playback missions with varying views, including 3-dimensional views, to assist in gaining a more objective understanding of their behaviours and the results of those behaviours. The AMPA playback facility used by the RAF uses a 3d image of the tactical scenario that is difficult to see using a playback system, hence, the legacy approach (aircraft on sticks) is used to fully explain the tactical manoeuvres during post-mission briefings. This notwithstanding, in live and simulated flight there is an absence in viewing student pilot actions in the cockpit, as camera feedback is missing. This facility would enhance the post mission briefing and bring a clearer picture on how the student is reacting to external stimuli and add to the variables under consideration for measuring readiness.

It is important to be aware of the VRTE regarding various training tasks as they have training capability limitations. This leads to a decision of when the student pilot is ready to 'prove' that ToT has occurred, by real-world flight assessment. Thus, the fundament of transfer of training is the thorough analysis of current pilot, instructor and blending mix capabilities and limitations, and training or organisation needs. The analysis of the training scenario tasks should be undertaken with respect to the dimension of the training effects of the task in gaining and developing relevant K&S. Each task should be decomposed into relevant pilot actions so that transfer effects for different types K&S can be studied in more detail leading to more concise training scenarios directed more closely to transfer of training; to be developed through the training pipeline, i.e. training cockpit procedures does not require a high fidelity FSTD with detailed features and an accurate aircraft model, as the requirements for this training stage merely is to allow the pilot to carry out the procedures correctly.

It is perceived that the benefits to training lie in the ability of the mission scenarios in conjunction with appropriate VRTE configuration layouts to 'get the pilot's clock up to speed'. The PC-based VRTE provides a moderately realistic preview of the pace of the task; high fidelity simulations are designed to reflect the complexity of the real world tasks without

the associated cost and risk with real systems. On the surface, it appears an elegant strategy because it seeks to minimize the transfer of training problem by closing the gap between training and the real-world task. However, high fidelity training requires significant investment in dedicated facilities and hardware; it must be staffed by expert trainers and coaches all who need training on the systems; it must be supported and updated regularly. Moreover, the considerable investment in such facilities will invariably ensure that training demand will exceed the capacity of the training system. Thus, cost and efficiency become major factors for training systems constructed around high-fidelity simulations in a time of tight budgets and higher-level of skill in a rapidly technological changing environment.

REFERENCES

- Abma, H.J., & Maig, A., "The Selection of Realistic Training Environments for Tactical Aircrew Training", *National Aerospace Laboratory, IITSEC Conference*, Orlando, 28-1 December, 2011.
- Adolph, W.S. & Bramble, P., "*Patterns for Effective Use Cases*", Boston, London: Addison-Wesley, 2003.
- Ajao, I.O., Abdullahi, A.A., & Raji, I.I., "Polynomial Regression model of Making Cost Prediction in Mixed Cost Analysis", *Mathematical Theory and Modeling*, Vol. 2, No. 2, 2012.
- Alanen, M., & Porres, I., "Model Interchange using OMG Standards", *Software Engineering and Advanced Applications*, 31st Euromicro Conference, January, 2010.
- Alexander, A.L., Brunye, T., Sidman, J., Weil, S.A., "*From gaming to training: A review of studies on fidelity, immersion, presence, and buy-in and their effects on transfer in PC-based simulations and games*", DARWARS Training Impact Group, 2005.
- Alipour, F., Idris, K., & Karimi, R., "Knowledge Creation and Transfer: Role of Learning Organisation", *International Journal of Business Administration*. Vol. 2, No. 3, August, 2011.
- Alliger, G., Bennet, W., Colegrove, C., "Understanding Mission essential Competencies as a Work Analysis Method", Air Force Research Laboratory, AFRL-HE-AZ-TR-2007-0034, August, 2007.
- Alliger, G.M., Beard, R.L., Benner, W.Jr. & Colegrove, C.M., "*Understanding Mission Essential Competencies as a Job Analysis Method*", In the handbook of work analysis: The Methods, Systems, Applications, & Science of Work Measurement in Organisation, edited by M.J. Wilson et. al., New York: Routledge, 2012.
- Alliger, G.M., Beard, R.L., Bennet, W., Symons, S., & Colegrove, C., "A Psychometric Examination of Mission Essential Competency (MEC) Measures Used in Air Force Distributed Mission Operations Training Needs Analysis", *Military Psychology*, Vol. 25(3), pp. 218-233, 2013.
- Alvarez, K., Salas, E., Garofano, C.M., "An integrated model of training evaluation and effectiveness", *Human Resource Development Review*, Vol. 3, pp. 382-416, 2004.
- Angeli, G., & Manning, D.P., "Natural Logic Inference for Common Sense Reasoning", *EMNLP*, 25-29 October, Doha, Qatar, 2014.
- Aquino, H., & Kraiger, K., "Benefits of Training and Development for Individuals and Teams, Organisations, and Society", *Annual Review of Psychology*, 60, pp. 451-474, 2009.
- Armour, F. & Miller, G., "*Advanced use case modelling: Software systems*", Boston, London: Addison-Wesley, 2001.

ATSB (Newman, G), “*An Overview of Spatial Disorientation as a Factor in Aviation Accidents and Incidents*” Australian Transport Safety Bureau, ATSB Transport Safety Investigation Report B2007/0063, 2007.

Ayers, F.H., “The Application of Scenario Based Recurrent Training to Teach Single Pilot Resource Management (SRM) Under the FAA Industry Training Standards (FITS) Program”, *Journal of Aviation / Aerospace Education and Research*, Vol. 15, No. 2, 2006.

Bajraktarvic, N., Hall, W., & Fullick, P., “Incorporating Learning Styles in hypermedia Environment: Empirical Evaluation”, In P. de Bra, H.C. Davis, J. Kay & M. Schraefel (Eds.). *Proceedings of the Workshop on Adaptive Hypermedia and Adaptive Web-Based Systems*, Nottingham, UK, pp. 41-52, 2003.

Bamadu, O., Ye, X., “Virtual Reality and Virtual Reality ‘System Components’”, *Proceedings of the 2nd International Conference on Systems Engineering and Modeling (ICSEM-13)*, pp. 765-767, 2013.

Bares, D., David, C., Mills, R.F., “A Tactical Framework for Cyberspace Situational Awareness”, *15th ICCRTS Evolution of C2*, June 22-24, Santa Monica, California, 2010.

Barbe, J., Chatrenet, N., Mollard, R., Berard, P., Wolff, M., “Physical ergonomics approach for touch screen interaction in an aircraft cockpit”, *Proceedings of the Ergo’IHM 2012 conference*, ACM, New York, 2012.

Barile, S., & Saviano, M., “The Viable Systems Approach and Its Contribution to the Analysis of Sustainable Business Behaviours”, *Systems Research and Behavioural Science*, Vol. 31, Issue 6, pp. 683-695, December, 2014.

Baron, S., “A Control Theoretic approach to Modeling Human Supervisory Control of Dynamic Systems”, *Advances in Man-machine Systems Research*, Greenwich, CT: JAI Press, Vol. 2, pp. 1-47, 1984.

Bashir, A.M., “Subjective Measures to Human Performance in Virtual Environment with Kinaesthetic Feedback”, *Computer Science and Information Technology (CSIT)*, pp. 334-338, 27-28 March, Amman, 2014.

Beaubien, J.M., Webb, J.E., Wiggins, S.L., & Keeney, M.J., “Differentiating Measures of Learning (MOL) From Measures of Performance (MOP) During Aircraft Carrier Landing Practice”, *IITSEC Paper No. 15210*, 2015.

Beck, D., “*Spiral Dynamics, Mastering Values, Leadership, and Change.*”, Malden, MA: Blackwell Publishing. p. 31. , (2003), [ISBN 155786-940-5](#).

Benkamoun, N., Elmaraghy, W., Hayet, A.L., & Kduiss, K., “Architecture Framework for Manufacturing System Design”, *Proceedings of the 47th CIRP Conference on Manufacturing Systems*, Vol. 17, pp. 88-93, 2014.

Belanich, J., Sibley, D.E., & Orvis, K.L., “*Instructional characteristics and motivational features of a PC-based game*”, Research report 1822, Alexandria, VA: U.S. Army Research Institute for the Behavioural and Social Sciences, 2004.

Bell, A. “*Death by UML Fever*”, ACM Queue, March 2004.

Bell, B.S., & Kozlowski, S.W.J., “Advances in Technology-based Training”, *Managing Human Resources in North America*, Abingdon: UK, pp. 27-43, 2007.

Belling, P.K., Suss, J., Ward, P., “*Investigating Constraints on Decision Making Strategies*”, International Conference on Naturalistic Decision Making, Marseille, France, 2013.

Belvardi, G., Kiraly, A., Varga, T., “Monte Carlo Simulation Based Performance Analysis of Supply Chains”, *International Journal of Managing Value and Supply Chain*, Vol. 3, No. 2, June, 2012.

Bendansolli, P.F., “Theory Building in Qualitative Research: Reconsidering the Problem of Induction”, *Qualitative Social Research*, Vol. 14, No. 1, January, 2013.

Bennekom, K.T.P., & Tulnen Van J.L., “*Pilot task Demand Load Considerations and a Flight mechanical Assessment Tool for RNAV approaches with a Cessna Citation*”, Master thesis report, Delft University of Technology, 2010.

Bennet, J., Alliger, G., & Beard, R., “Understanding Mission Essential Competencies as a Work Analysis Method”, *Air Force Research laboratory*, Mesa, AZ., 2007, available at [Http://www.dtic.mil](http://www.dtic.mil) [accessed 26 October 2013].

Bennet, W.Jr., Schreiber, B.T., Andrews, D.H., “Developing Competency-based Methods or Near Real-time Air Combat Problem Solving Assessment”, *Computers in Human Behaviour*, vol. 18, pp. 773-782, 2002.

Benyon, D., “*Information and Data Modelling*”, 2nd edition, McGraw-Hill, London, 1997.

Berry, D.M., “Ambiguity in Natural Language Requirements Documents”, *Innovations for Requirements Analysis: 14th Monterey Workshop*, CA, 10-13 September, 2007.

Bestoso, A.J., “*ESG-I Missions and Assets for Advanced Modeling and Simulation of C2 Structures*,” Master’s Thesis, Naval Postgraduate School, Monterey, CA, September, 2005.

Bhattacharya, A., Geraghty, J., Young, P., “Supplier Selection Paradigm: An Integrated Hierarchical QFD Methodology Under Multiple-Criteria Environment”, *Applied Software Computing*, Vol. 10, Issue 4., September, 2010.

Bilotta, F.F., Werner, S.M., Bergose, S.D., Rosa, G., “Impact and Implementation of Simulation-based Training for Safety”, *The Scientific World Journal*, November, 2013.

BKCASE, “Body Of Knowledge and Curriculum To Advance Systems Engineering”, <http://www.bkcase.org/>, 2014.

Blevins, T., “The Architecture of Enterprise Architecture”, MITRE / INCOSE, 2006.

Blumel, E., & Haase, T., ”Virtual Reality Platforms for Education and Training in Industry”, *Advances in Databases and Information Systems*, 2(2), pp. 1-7, 2010.

Boff, K.R., “Revolutions and shifting paradigms in human factor & ergonomics”, *Applied Ergonomics*, Vol. 37(4), pp. 392-399, 2006.

Bolloju, N., Khalifa, M., & Turban, E., “Integrating knowledge Management into Enterprise environments for the next generation decision support”, *Decision Support Systems*, Vol.33, pp. 163-176, 2002.

Borgvall, J., Castor, M., Nahlinder, S., Oskarsson, P.A., & Svensson, E., “Transfer of training in military aviation”, *FOI-R-2378-SE. FOI METHODOLOGY REPORT*, Linkoping, Sweden, 2008.

Boardman, J., and Sauser, B., “System of Systems – the meaning of-“, *The Proceedings of rht 2006 IEEE/SMC Interaction Conference on System of Systems Engineering*, Los Angeles, CA, pp. 118-123, 2006.

Boehm, B., Lane, J., “21st Century Processes for Acquiring 21st Century Software-intensive systems of systems”, *Crosstalk- The Journal of Defence Software Engineering*, Vol.19, No.5, pp.4-9, 2006.

Boosman, F., “Simulation-Based Training: The Evidence is in”, Pseudorandom, http://boosman.com/blog/2007/06/simulationbased_training_the_e.html, [Accessed: 30/09/2014].

Boulding, K.E., “General systems theory: the skeleton of science”, *Management Science* 2, pp. 197-208, 1956.

Bowles, M.S., “*Relearning to e-learn: Strategies for electronic learning and knowledge*”, Melbourne: Melbourne University Press, 2004.

Brannen, J.J., “*Sustaining the US Lead in Unmanned Systems: Military and Homeland Considerations through 2025*”, Report of the CSIS International Security Program, February, 2014.

Brooks, F.P., “*The Mythical man month and other essays on software engineering*”, Addison-Wesley Longman, Inc., 1995.

Browning, T.R., “Applying the Design Structure Matrix to System Decomposition and Integration Problems: A review and New Directions”, *IEEE Transactions on Engineering Management*, Vol. 48, No. 3, August, 2001.

Buchanan, D., and Huczynski, A., “*Organisational behaviour: an introductory text (4th edition.)*”, Financial Times Management, 2001.

Buchelder, E., & Leveson, N., “*Describing and Probing Complex System Behaviour: A Graphical Approach*”, SEA Technical Paper 2001-01-2646, 2001.

Buhari, B.A., & Mohammed, M., “An Object-Oriented Model for Automatic Staff Promotion Management and Evaluation: Dynamic View”, *International Journal of Advanced Scientific and Technical Research*, Vol. 6, Issue 4, December, 2014.

Bull, S. & Ma, X., “Raising learner Awareness of Language Learning Strategies in Situations of Limited Resource”, *Interaction Learning Environments*, 9(2), pp. 171-200, 2004.

Bullemer, P.T., “Improving the Operation Team Situation Awareness: Lessons Learned from Major Process Industry Accidents”, Operator Team Situation Awareness, AFPM 2013AM, 2013.

Bullet, S., “Learning through Practice: Modeling Traditions, Orientations and Approaches”, *Professional and Practice-based Learning*, 2010, ISBN: 978-9048139392.

Burke, I.A., & Hutching, H.M., “Training Transfer: An Integrative Literature Review”, *Human Resource Development Review*, 6, pp. 263-269, 2007.

Burke, R., “*Project Management: Planning and Control Techniques*”, Chichester: Wiley, 2003.

Burgon, R., “Muscle Memory”, Tally-One, 12 January, 2014, <http://www.tally-one.com/muscle-memory/>, [Accessed: 23/08/2014].

Burns, A., “System Mode Changes – General and Critically-Based”, *Workshop on Mixed Criticality Systems: Real Time Systems Symposium (RTSS 2014)*, Rome, Italy, December, 2014.

Bushey, D.A., Forsythe, M.J., “The Recognition Primed Decision Model: An Alternative to the MDMP for GWOT”, *Field Artillery*, January-February, pp.10-13, 2006.

Buxbaum, P., “LVC for Integrated Training”, MT2 Responder Training Services, <http://www.MT2-KML.com>. [Accessed: 7/10/2014].

CAA, “*Flight Crew Reliance on Automation*”, CAA Paper 2004/10, December, 2004.

CAA, “*Monitoring Matters: Guidance on the Development of Pilot Monitoring Skills*”, Civil Aviation Authority: Loss of Control Action Group, 2013/02, 2013.

CAAP, “*Competency Based Training and Assessment in the Aviation Environment*”, Civil Aviation Advisory Publication, CAAP 5.59A-1(0), July, 2009.

Caird, J.K., "Persistent issues in the application of virtual environment systems to training", *Proceedings of HICS'96: 3rd annual symposium on human interaction with complex systems*, Los Alamitos, CA: IEEE Computer Society Press, pp. 124-132, 1996.

Campbell, J.P., & Kuncel, N.R., "*Individual and team training*", Handbook of work and organizational psychology, London: Blackwell, 2001.

Cannon-Bowers, J.A., Salas, E., "A Framework for developing team performance measures in training", *Team Performance assessment and measurement: Theory, Methods and applications*, Hillside, NJ, pp. 45-62, 2012.

CASA, "Competency Based Training and Assessment in the Aviation Environment", Civil Aviation Advisory Publications, CAAP 5.59A-1(0), July, 2009.

Carretta, T.R., "Predicting Pilot Training Performance: Does the Criteria Make A Difference?", AD-A248485, Human Resources Directorate, Texas, March, 1992.

Casner, M., Geven, R.W., Matthias, P.R., Schooler, J.W., "The Retention of Manual Flying Skills in the Automated Cockpit", *Human Factors: The Journal of the Human Factors and Ergonomics Society*, May, 2014.

Castor, M., Borgvall, J., Bennet, W., "Knowledge and skill-based evaluation of simulated and live training – From evaluation framework to field application", *Proceedings Of International Symposium for Aviation Psychology*, Dayton, OH: Wright State University, 2009.

Champney, R., Milham, L., Carroll, M.B., Stanney, K.M., Cohn, J., "A method to determine optimal Simulator Training Time: Examining Performance Improvement Across the Learning Curve", *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 50: 2654, October, 2006.

Chapman, R., & Colegrove, C., "Transforming Operational Training in the Combat Air Forces", *Military Psychology*, Vol. 25(3), pp. 177-190, May, 2013.

Chandra, D.C., & Grayhem, R.J., "Human Factors Research on performance-based Navigation Instrument Procedures for NextGen", *31st Digital Avionics Systems Conference*, 14-18 October, Williamsburg, VA, 2012.

Chawla, D., & Joshi, H., "Impact of Knowledge Management Dimensions on Learning Organizations: Comparison across Business Excellence Awarded and Non-awarded indian Organisations", *International Journal of Knowledge Management*, 7(2), pp. 68-85, 2011.

Checkland, P., "*Systems Thinking, Systems Practice*", Chichester: Wiley, 1999.

Cheng, E.W.L., & Hampson, I., "Transfer of Training: A Review and New Insights", *International Journal of Management Reviews*, 10, pp. 327-341, 2008.

Chiaburu, D.S., & Marinova, S.V., “What Predicts Skills Transfer?: An Exploratory Study of Goal Orientation, Training Self-efficacy and Organisational Support”, *International Journal of Teaching and Development*, 9, pp. 110-123, 2005.

Chialastri, A., “Human-Centred design in Aviation”, *CEUR Proceedings 4th Workshop HCP Human Central Processes*, 10-11 February, 2011.

Christopher, J.M, Mieczyslaw, K.M., Baclawski, K., Jerzy, A.L., Call, C., Hinman, M.L., Salerno, J.J., Boulware, D.M., “SAWA: An assistant for higher-level fusion and situation Awareness”, *Proceedings of the SPIE*, Vol. 5813, pp. 75-85, 2005.

Clarke, R. “*Appropriate Research Methods for Electronic Commerce*” URL: <http://www.rogerclarke.com/EC/ResMeth.html>, April, 2000, [Date Accessed: 05/06/11].

Clarke, R.C. & Mayer, R.M., “*E-Learning and the Science of Instruction*”, 3rd Edition, San Francisco, CA: Pfeiffer, 2011.

Cluni, R., Wansing, H., & Willkommen, C., “*Recent Trends in Philosophical Logic*”, Springer Science & Business Media, 2014, ISBN: 978-3-319-06079-8.

Cockburn, A., “*Writing effective use-cases*”, Reading, MA: Addison-Wesley, 2001

Coffield, F., Moseley, D., Hall, E., Ecclestone, K., “*Learning Styles and pedagogy in post-16 learning: A Systematic and Critical review*”, <http://www.hull.ac.uk/php/edskas/learning%20Styles.pdf>, [accessed: 25 July 2013].

Colegrove, C.M. & Alliger, G.M., “Mission Essential Competencies: Defining Combat Mission Readiness in a Novel way”, *NATO Research and Technology Organisation, Studies, Analysis and Simulation Panel, Conference on Mission Training Via Distributed Systems*, Brussels, Belgium, 2002.

Colegrove, C.M. & Alliger, G.M., “Mission Essential Competencies: Defining Combat Mission Readiness in a Novel Way”, *NATO Research & Technology Organisation, Studies, Analysis, and Simulation Panel, Conference on Mission Training via Distributed Simulation (SAS 38)*, Brussels, Belgium, 2002.

Colgrove, C.M. & Alliger, G.M., “Mission Essential Competencies: Defining Combat Mission Readiness in a Novel Way”, *In RTO SAS Symposium on air mission training through distributed simulation (MTDS) – Achieving and Maintaining Readiness*, Belgium: NATO research and Technology Organisation, Vol. 22, 2002.

Cohn, J., Kay, S., Milham, L, Bell, C.M., Jones, D., Sullivan, J., Darken, R., “*Training Effectiveness evaluation: from theory to practice*”, *The PSI Handbook of Virtual Environments for Training and Education*, pp. 157-172, 2009.

Corker, K.M., Verma, S., Gore, B.F., Jadhav, A., & Guneratne, E., “*SJSU/NASA Coordination of Air MIDAS Safety Development Human Error Modelling: NASA Aviation*

Safety Program Phase 1: Integration of Air MIDAS Human Visual Model Requirements and Phase 2: Validation of Human Performance Model for Assessment of Safety Risk Reduction through the Implementation of SVS Technologies", Hail Laboratory Report 2307-2005, San Jose, CA: SJSU, Human Automation Integration Laboratory, 2003.

Cormier, S.M. & Hagman, J.D., "*Transfer of Learning: Contemporary research and applications*", San Diego, CA: Academic, 1987.

Courtney, J.F., "Decision making and knowledge management in inquiring organisations: Towards a new decision-making paradigm for DSS", *Decision Support Systems*, No. 31, pp. 17-38, 2001.

Crocker, A., "*A Technique for the Assessment of Flight Operability Characteristics of Human Rated Spacecraft*", American Institute of Aeronautics and Astronautics, 2005.

Crooke, S., & Slack, N., "*Making Management Decisions*", 2nd edition, Prentice Hall International, London, 1991.

Cummings, C., Laing, K., Law, J., Mchaughlin, Papps, I., Todd, L., & Woolner, P.O., "*Can Changing A Aspirations and Attitudes Impact on Educational Attainment?: A Review of Interventions*", Joseph Rowntree Foundation, York, JRF: 2012, 2012.

Cummings, M.L., & Nehme, C.E., "*Modelling the Impact of Workload in Network Centric Supervisory Control Settings*", 2nd Annual Sustaining Performance Under Stress Symposium, College Park, M.D., February 25, 2009.

Damos, D.L., Schwartz, K.L., & Weissmuller, J.J., "Knowledge, Skills, Abilities and Other Characteristics for Military Pilot Selection: A Review of the Literature", *Strategic Research and Assessment*, Texas, 5 February, 2011.

DAU, "*Defense Acquisition Guidebook (DAG)*", Ft. Belvoir, VA, USA: Defense Acquisition University (DAU)/U.S. Department of Defense (DoD), 2010.

De-Brain, W.B., Parker, A.M., Fischhoff, B., "Individual Differences in a Adult Decision-Making Competence", *Journal of Personality and Social Psychology*, 92(5), pp. 938-956, 2007.

Defence Science Board, "*Report of the Defense Science Board Task Force on Training Superiority and Training Surprise*", Washington, D.C., January, 2001.

Delahaig, R., Van Dam, K., Gaillard, A.W.K., & Soeters, J., "Predicting Performance Under Acute Stress: The Role of Individual Characteristics", *International Journal of Stress Management*, Vol. 18(1), pp. 46-66, February, 2011.

Department of Defence, "*Systems engineering guide for system of systems*", version 1.0, 2008.

Department of the Army, “*Field Manual 101-5 Staff Organisation and Operations*”, Defence Technical Information Center, 31 May, 1997.

Denning, T., Bennet, W., Crane, P., “Mission complexity scoring in Distributed Mission Training”, *Proc. Of Industry/Inter-service Training Systems Conference*, Orlando, FL: National Security Industrial Association, 2002.

Deniaud, C., Honaet, V., Benoit, J., & Mestre, D., “The Concept of ‘Presence’ as A measure of Ecological Validity in Driving Simulators”, *Journal of Interaction Science*, 3(1), 2015.

Department of Defence (D0D), “DoD Modeling and Simulation (M&S) Glossary”, *DoD*, March, 2010.

Dickerson, C., “Towards a Logical and Scientific Foundation for Systems Concepts, Principles, and Terminology”, *3rd IEEE International Symposium on Service-Oriented Systems Engineering*, Monterey, CA: IEEE, 2008.

Dickerson, C.E., & Mavris, D.N., “A Brief History of Models and Model Based Systems Engineering and The Case for Relational Orientation”, *IEEE Systems Journal*, pp.99, 2013.

Dickerson, C.E., Valerdi, R., “Using relational model transformations to reduce complexity in SoS requirements traceability: preliminary investigation”, *Proceedings IEEE SOSE 2010*, Loughborough, June 2010.

Dismukes, R.K., Berman, B.A., Loukopoulos, L.D., “*The limits of expertise: rethinking pilot error and the causes of airline accidents*”, Aldershot, UK: Ashgate Publishing Co., 2007.

Doiron, M., “Aviation Risk Management – Accountabilities, Challenges, and Tools”, *Journal of Aviation Management*, Cirrus Aviation Safety Services, Canada, pp. 41-48, 2014.

Donmez, B.D., Nehme, C., & Cummings, M.L., “Modelling Workload Impact in Multiple Unmanned Vehicle Supervisory Control”, *IEEE Systems, Man, and Cybernetics, Part A Systems and Humans*, Vol. 40, Issue 6, pp. 1180-1190, 2010.

Drury, J.L., Richer, J., & Radcliffe, N., “Comparing Situation Awareness for Two Unmanned Aerial Vehicle Human Interface Approaches”, *IEEE International Workshop of Safety, Security and Rescue Robots (SSRR)*, 2006.

D'Souza, D., “*Model-Driven Architecture and Integration: Opportunities and Challenges*”, Version 1.1. <http://www.catalysis.org/publications/papers/2001-mda-reqs-desmond-6.pdf>, 2011.

Dumitresco, C., Tessier, P., Salinesi, C., Gerard, S., Davron, A., & Mazo, R., “Capturing Variability in Model Based Systems Engineering”, *Complex System Design and Management (CSDGM) Conference*, Paris, France, December, pp. 100-115, 2013.

Dunan, J., “Fidelity Versus Cost in Modelling and Simulation”, *12th ICCRTS: Adapting C2 to the 21st Century*, VA: USA, 2006.

Dunn, R.J., “The Impact of a Declining Defense Budgets on Combat Readiness”, #2828 National Security and Defense, 18 July, 2013.

EASA, “New Training Methods and New Teaching Technologies”, *European Aviation Safety Agency Proposed Amendments 2014-22*, RMT.0281 (MDM.082), October, 2014.

EASA, “Bridging Design and Training Principles”, *EASp Action EME 4.4: Internal Group of Personnel Training (IGPT)*, 28 May, October, 2013.

EASA, “*Loss of Control Prevention and Recovery Training*”, European Aviation Safety Agency, RMT.0582, Issue 2, March, 2015.

Edvinsson, L., & Malone, M.S., “*Intellectual Capital*” Great Britain: Piatkus, 1997.

Eickmeyer, K., & Gruhe, M., “Randomization and Derandomisation in Descriptive Complexity Theory”, *Proceedings of the 24th International Conference on Computer Science Logic*, pp. 275-289, 2010.

Elammari, M., & Issa, Z., “Using Model Driven Architecture to Develop Multi-Agent Systems”, *International Arab Journal of Information Technology*, Vol. 10, No. 4, July, 2013.

Elhag, A.A.M., Mohamad, R., Aziz, M.W., & Zeshan, F., “A Systematic Composite Service Design Modeling Method using Graph-Based Theory”, *PLoS ONE*, 10(4), 20 April, 2015.

Emre, A.M., “Analysis of the Benefits of Motion Simulators in 5th Generation Fighter Pilots’ Training”, *World Academy of Science, Engineering and Technology: International Journal of Social, Behavioural, Economics, Business and Industrial Engineering*, 10(2), pp. 332-335, 2016.

Endsley, M.R., “*Situation awareness in dynamic systems*”, *Situational awareness in dynamic systems (IAM Report 708)*”, Farnborough, UK: Royal Air Force Institute of Aviation Medicine, 1991.

Endsley, M.R., “Towards a theory of Situation Awareness in Dynamic Systems”, *Human Factors*, Vol. 37, pp. 32-64, 1995a

Endsley, M.R., Bolte, B., Jones, D.G. “*Designing for situation awareness*”, London, UK: Talyor and Francis, 2003.

Endsley, M.R., Bolte, B, Jones, D.G., “*Designing for Situation Awareness: An Approach to User Centred Design*”, New York: Taylor and Francis, 2003.

Endsley, M.R., Connors, E.S., “Situation Awareness: State of the Art”, *Power and Energy General Meeting – Conversion and Delivery of Electrical Energy in the 21st century*, IEEE, 2-24th July, 2008.

Eriksson, M., Borstler, J., & Borg, K., “*The PLUSS Approach – Domain Modeling with Features, Use Cases and Use Case Realizations*”, 9th International Conference of Software Product Lines, 2005.

ESA, “*Mission Analysis: Towards a European Harmonization*”, Paris, France: European Space Agency. Accessed 29/11/2014. Available at http://www.esa.int/esapub/bulliten/bulletin134/bul134b_schoenmackers.pdf, 2008.

Estock, J.L., Alexander, A.I., Nash, M., “A Model-based Approach to Simulator Fidelity and Training Effectiveness”, *Interservice/Industry Training Simulation, and Education Conference (I/ITEC)*, pp. 1-9, 2006.

Evans, S.E., “*Improving the Cost Efficiency and Readiness of MC-130 Aircrew Training: A Case Study*”, RAND Corporation, Santa-Monica, CA, 2015.

Everson, M.D., “*Training System Device Certification and Qualification Process*”, Naval Postgraduate School, September, 2013.

Evora, J., Hernandez, J.J., & Hernandez, M., “Decision Support for Complex Systems: A Smart Grid Case”, *COSMOS 2013: 6th Complex Systems Modelling and Simulation Workshop*, Italy, July, 2013.

Federal Aviation Administration, “*AC 120-51E – Crew Resource Management Training*”, AFS-210, January, 2004.

Federal Aviation Administration, “*Electronic Qualification Test Guide (eQTG) systems for Flight Simulation Training Devices (FSTDs)*”, 05-02, December, 2014.

Faisandier, A., “*Systems Architecture and Design*”, Belberaud, France: Sinergy’Com, 2012.

Fanning, R., “The role of Debriefing in Simulation-based Learning”, *Society of Simulation in Healthcare*, Vol. 2, No. 2, Summer, 2007.

Fehlmann, T., “The Impact of Linear Algebra on QFD”, *International Journal of Quality and Reliability Management*, 22(1), pp. 83-96, 2015.

Felder, R.M., & Silverman, L., “Learning and Teaching Styles in Engineering Education”, *Engineering Education*, 78(7), pp. 674-681, 1988.

Felder, R.M. & Spurlin, J.E., “A Validation study of the Index of Learning Styles. Applications, reliability, and validity of the Index of Learning Styles”, *International Journal of Engineering Education*, 21(3), pp.103-112, 2005.

Fernandes, A., & Braarud, P.O., “Exploring Measures of workload, Situation Awareness, and task Performance in the Main Control Room”, *6th International Conference on Applied Human Factors and Ergonomics (AHFE 2015)*, 2015.

Ferris, T.L.J, Cook, S.C., Sitnikova, E., “Design as Research Methodology for Systems Engineering”, *Proceedings of the Conference on Systems Engineering Research CSER2008*, Los Angeles, CA, USA: University of Southern California, 2008.

Finlay, P.N., “*Introducing Decision Support Systems*”, NCC Blackwell, Oxford, 1989.

Fletcher, C, “*Performance appraisal: assessing and developing performance and potential*”, Introduction to work and organisational psychology: a European perspective, Oxford: Blackwell, 2000.

Fletcher, G., Pethick, J, Muncie, H., “*Defining, Measuring and Assessing Preparedness*”, Defence Science and Technology Laboratory, 2011.

Fletcher, J.D., & Ward, A.P., “*The Evolving Definition of Readiness for Military Operations*”, Teaching and Measuring Cognitive Readiness, Springer Science & Business Media, New York, USA, 2014.

Flin, R., Lynne, M., Klaus-martin G., & Hans-Jurgen, H., “Development of the NOTECHS (Non-technical Skills) System for Assessing Pilots ‘CRM Skills’”, Human Factors and Aerospace Safety, 3(2), pp. 95-117, Ashgate Publishing, 2003.

Ford, T.C., Colombi, J.M., Jacques, D.R., Graham, S.R., “On the Application of Classification Concepts to Systems Engineering Design and Evaluation”, *Systems Engineering* 12, No. 2, pp. 141-154, 2009.

Forester, J., Bley, B., Cooper, S., Lois, E., Siu, N., “Expert Elicitation Approach for Performing ATHEANA Quantification”, *Reliability Engineering and System Safety*, 83(4), pp. 207-220, 2004.

Foster, M.E., Givliani, M., & Knoll, A., “Comparing Objective and Subjective Measures of Usability in a Human-Robot Dialogue System”, *Proceedings of the 47th Annual Meeting of the ACL and the 4th IJCNLP of the AFNLP*, Singapore, August, pp. 879-887, 2009.

Franks, P., Hays, S., & Mavin, T., “Can Competency-Based training Fly?: An Overview of Key Issues for Ab Initio Pilot Training”, *International Journal of Training Research*, August, 2014.

Freitas, S.F., & Jarvis, S., “A Framework for Developing Serious Games to Meet Learner Needs”, *Interservice/Industry Training, Simulation, and Education Conference (I/ITSEC)*, Paper No. 2742, pp. 1-11, 2006.

Friedman, T., “*The World is Flat: A brief history of the twenty-first century*”, Farrar, Straus and Giroux, 2005.

Friston, S., & Steed, A., “Measuring Latency in Virtual Environments”, *Journal IEEE Transactions on Visualisation and Computer Graphics*, Vol. 20, Issue 4, April, 2014.

Gabb, A., “Requirements Categorization”, INCOSE requirements working group, www.incose.org/rwg/01_req_categories/categories.html, 2001.

GAO, “Unmanned Aerial Systems: Actors Needed to Improve DOD Pilot Training”, United States Government Accountability Office, Report to Congressional Committees, May, 2015.

Ganguly, S., & Krishnamurti, R., “Algorithms and Discrete Applied Mathematics”, *1st International Conference (CALDAM, 2015)*, 8-10 February, Kanpur, India, 2015.

Gargantini, A., “Combining Formal Methods and MDE Techniques for Model-driven System Design and Analysis”, *International Journal on Advances in Software*, 3. 1. 2. pp.1-18, 2010.

Gawron, V.J., “Human Performance, Workload, and Situational Awareness Measures Handbook: 2nd Edition”, CRC Press, 2008.

Gee, J., “What video games have to teach us about learning and literacy”, New York, Palgrave Macmillan, 2003.

Gehr, S.E., Schreiber, B., Bennet, W., “Within-Simulator Training Effectiveness Evaluation”, *Proceedings of Interservice/Industry Training, Simulation and Education Conference (IITSEC)*, 2004.

Gelijns, A.C., Brown, L.D., Magnell C., Ronchi, E., Moskowitz, A.J., “Evidence, Politics, and Technology Change”, *Health Affairs*, 24, No. 1, pp. 29-40, 2005.

Gerlach, T., & Durak, U., “AVES SDK: Bridging the Gap Between Simulator and Flight Systems Designers”, *AIAA Modelling and Simulation Technologies Conference*, Dallas, Tx, 22-26 June, 2015.

Ghani, U., “Role of Feedback in Simulation-Based Learning”, *Journal of Social, Behavioural, Economics, Business and Industrial Engineering*, 10(1), pp285-292, 2016.

Gibbons, A.S., “An Architectural Approach to Instructional Design”, Routledge Taylor & Francis Group, 2014, ISBN: 978-0-415-80738-8.

Gillet, S., Nuic, A., & Mouillet, V., “Enhancement in Realism of ‘ATC’ Simulations by Improving Aircraft Behaviour Models”, *In 29th Digital Avionics Systems Conference*, 2010.

Goldberg, S., “Human Factor Issues in the Use of Virtual and Augmented Reality for Military Purposes”, USA, NATO RTA Report, Pentagon Reports: Fast. Definitive, complete, USA, 2005.

Gory, A., & Morton, S., “A Framework for Management Information Systems”, *Decision Support Systems*, Vol. 13, No. 1, pp. 55-70, 1971.

Gottesdiener, E., “Use Cases: Best Practices”, Rational Software, http://www.rational.com/media/whitepapers/usecases_bp.pdf, 2003, [Accessed: 03/05/2012].

Graessar, C.A., & King, B., “*Technology Based Training*”, Human Behaviour in Military Contexts, The National Academies Press, USA, 2008.

Graham, C.R., “*Blended Learning Systems: Definition, Current Trends, and Future Directions*”, Handbook of Blended Learning: Global Perspectives, San Francisco, CA: Pfeiffer Publishing, 2004.

Gray, C., “Improving Flight Condition Situational Awareness through Human Centered Design”, *Work* 41, pp. 4523-4521, 2012.

Gray, D., “*Intellectual Capital: Measuring and Enhancing the truth Value of your Business*”, Financial Times Executive Briefing. Great Britain: Prentice Hall, 2001.

Grogan, P.T., deWeck, O.L., Ross, A.M., & Rhodes, D.H., “Interactive Models as A Systems Design Tool: Applications to Systems Project Management”, *13th Conference on Systems Engineering Research*, Hoboken, NJ, March, 2015.

Grote, G., and Kunzler, C., “Diagnosis of Safety culture in safety management audits”, *Safety Science*, No.34, pp. 131-150, 2000.

Grossman, R., & Salas, E., “The Transfer of Training: What Really Matters”, *International Journal of Training and Developments*, 15(2), pp. 1360-1376, 2011.

Guenov, M.D., & Barker, S.G., “Application of Axiomatic Design and Design Structure Matrix to the Decomposition of Engineering Systems”, *Systems Engineering*, Vol. 8, No. 1, 2005

Gundert, D., Oubaid, U., Oechsle, S., Zinn, F., “Determining the Workload of Ability in a Sub Selection Process”, *International Journal of Psychology*, Vol. 47, Issue. 51, pp. 42-45, January, 2012.

Gustaussen, P.M., Bjorkman, V.I., & Lundmark, S., “Agile Live, Virtual, and Constructive Training and Experimental Network”, *Conference: NATO RTO MSG-076 Symposium*, Vol. MSG-076, paper 14, 2013.

Haar, R.T., “Virtual Reality in the Military: Present and Future”, *In 3rd Twente Student Conference on IT*, 2005.

Hamblin, C.J., “*Transfer of Training from Virtual Reality Environments*”, Doctor of Philosophy Dissertation, May, 2005.

Hanson, C., Schnefer, J., Burken, J.J., & Larson, D., “*Complexity and Pilot Workload Metrics for Evaluation of Adaptive Flight Controls on a Full Scale Piloted Aircraft*”, NASA/TM-2014-216640, February, 2014.

Hardison, C.M., Zaydman, M., Olowatola, T., Saavedra, A.R., Bush, T., Peterson, H., & Strauss, S.G., “360-Degree Assessments: Are they the Right Tool for the U.S. Military”, RAND Corporation, Santa Monica, CA, 2015, ISBN: 9780-8330-8905-2.

Harrison, T., “*Rethinking Readiness*”, Strategic Studies Quarterly, Fall, 2014.

Hart, S., “NASA-Task Load Index (NASA-TLX): 20 Years Later”, *Proceedings of the Human Factors and Ergonomic Society 50th Annual Meeting*, Santa Monica: HFES, pp. 904-908, 2006.

Hartley, K., “*The Future of Outsourced Services in the UK Defence Sector*”, University of York, 3 June, 2015.

Haslbeck, A., Kirchner, P., Schubert, E., Bengler, K., “A Flight Simulator Study to Evaluate Manual Flying Skills of Airline Pilots”, *Proceedings of the Human Factors and Ergonomics Society 58th Annual Meeting*, Chicago, 2014.

Hays, R.T. & Singer, M.J., “*Simulation Fidelity in Training System Design: Bridging the Gap between Reality and training*”, New York, NY: Springer-Verlag, 1989.

Hays, R.T., “*The Effectiveness of instructional games: A Literature review and discussion*” Technical report 2005-004, Naval Air Warfare Training System Division: Orlando, FL, 2005.

Hazard, B., “Legacy Systems Integration with MDA: An MDA Success Story”, <http://omg.org/mda-files/UNextMDA4.pdf>, [Accessed: 12/10/2014].

Heena, C., & Rangna, B., “A Comparative Study of UML Tools”, *ACAI 11th Proceedings of the International Conference on Advances in Computing and Artificial Intelligence*, pp 1-4, NY: USA, 2011.

Henderson, E., Patray, J., & Breaux, R., “Performance Measures for Training Spatial Knowledge in Virtual Environments using Division Inc. and VEGA Marine Geometric Models”, *IEEE International Conference on Information Visualisation*, 17-21 June, pp. 524-528, 2000.

Henning, C., “Mathematical Models and Reality: A Constructivist Perspective”, University College London, 19 June, 2009, <http://www.ucl.ac.uk/statistics/research/pdfs/rr304.pdf>. [Accessed: 31/10/2013].

Hernandez, J.E., Zarate, P., Dargam, F., Delibastic, B., & Liu, S., “Decision Support systems – Collaborative Models and Approaches in Real Environments”, *Euro Working Group Workshops, EWGS-DSS*, London, June, 2011.

Hernandez, J.E., Liu, S., Delibastic, B., Zarate, P., Dargam, F., & Ribero, R., “Decision Support Systems II – Recent Developments Applied to DSS Network Environments”, *Euro Working Group Workshop, EWG-DSS*, Liverpool, April, 2012.

Hess, R.A., & Marches, F., “Analytical Assessment of Flight Simulator Fidelity Using pilot Models”, *Journal of Guidance, Control and Dynamics*, Vol. 32, No. 3, May-June, 2009.

Higgins, A.G., & Howell, K., “VMAS – A Proposed Educational Framework, Validation Methodology for Medical Simulation Training (VMAS)”, Committee Educational Framework, 2004.

Higgins, S., Xiao, Z., & Katsipataki, M., “*The Impact of Digital Technology on Learning: A Summary for the Education Endowment Foundation*”, Education Endowment Foundation, November, 2012.

Hjelmfelt, A.T., Pokrant, M.A., “*Coherent tactical Picture*”, CAN RM 97-129, March, 1998.

HoC, “*Defence Acquisition: Seventh Report of Session 2012-13*”, House of Commons Defence Committee: HC9, 29 January, 2013.

Hodgson, A., Siemeiniuch, C.E., Hubbard, E., “Culture and the Safety of Complex Automated Sociotechnical Systems”, *IEEE Transactions of Human-Machine Systems*, Vol. 43, No. 6, November, 2013.

Holden, J. & Westfall, P., “Learning Styles & Generational Differences: Do they matter?” *Proceedings of the Interactive technologies Conference 2010*, Arlington, VA, 18-20 August, 2010.

Holden, T., "Model driven Aviation Training Enterprise Architecture", *Loughborough Internal Conference*, Loughborough University, Hollywell Park, UK, 2012.

Holden, T., Dickerson, C.E., "A ROSETTA Framework for Live / Synthetic Aviation Trade-offs: Preliminary Report", *8th Annual International Conference on SoSE*, 4th June, Wailea-Makena, HI, USA, 2013.

Holden, T., & Dickerson, C.E., "Relation Oriented Systems Engineering Framework for Flight Training", *The Art and Science of Using Live-Virtual-Constructive Simulations for Analysis*, Journal for Defense Modeling and Simulation, pp. 1-13, 2013.

Holt, J., Perry, S., Bryans, S., Payne, R., Hansen, F., & Hellerstede, S., “A Model-based Approach for Requirements Engineering for Systems of Systems”, *IEEE Systems Journal*, 22(4), 2014.

Hooey, B.L., Foyle, D.C., Andre, A.D., “A Human Centred Methodology for the Design, Evaluation, and Integration of Cockpit Displays”, *Proceeding of the NATO RTO SCI and SET Symposium on Enhanced and Synthetic Vision Systems*, NATO, 2002.

Hohmann, M., & Orlick, T., “Examining the Psychological Skills Used by Elite Canadian Military Pilots”, *Journal of Excellence*, Vol. 4, Issue No. 16, 2014.

Horsink, S.A., Meyer, R., Wagner, S., Buchty, R., & Berekovic, M., “On RTL to TLM Abstraction to Benefits Simulation Performance and Modelling Productivity in NoC Design Exploration”, *Proceedings of the 2014 International Workshop of Networking on Chip Architectures*, pp/ 35-44, NY, 2014.

Horvath, L., & Rudus, I.J., “Systems Engineering Methods for Multidisciplinary Product Definition”, *12th International Symposium on Intelligent Systems and Informatics (SISY)*, 11-13 September, pp. 293-298, 2014.

Hosack, B., & Papadice, D., “Increasing Personal Value Congruence in Computerized Decision Support using System Feedback”, *Axioms*, Vol. 3, pp. 84-108, 2014.

Hosemann, M.S., “*The Proficiency Puzzle: Maintaining Airmanship in America’s Mobility Force Since 9-11*”, School of Advanced Air and Space Studies, Maxwell Air-Force Base, Alabama, June, 2013.

Hou, M.H., Sobieral, S., Pronovost, S., Roberts, S., Banbury, S., “*Suitable learning styles for intelligent tutoring technologies*”, Defence Research and Development Canada, 2010.

Hughes, F.E., “Training and Education for Future Human Space Flights”, *Space 2004 Conference and Exhibit*, San Diego, California, 28-30 September, 2004.

Hurley, D., “Defence Simulation Strategy and Roadmap”, Commonwealth of Australia, 2011, ISBN: 978-0-642-29741-9.

Hussain, M., & Senere, D.T., “Virtual Control System Development Platform with the Application of PLC Device”, *Proceedings of the International Multi-Conference of Engineers and Computer Scientists*, Vol. 2, 13-15 March, Hong-Kong, 2013.

IATA, “*Guidance Material and Best Practices for Pilot Aptitude Testing*”, International Air Transport Association, Geneva, March, 2012, ISBN:978-92-9233-779-7.

IATA, “*Evidence-Based Training Implementation Guide*”, International Air Transport Association, Montreal, Geneva, July, 2013.

IATA, “*Guidance Material and Best Practices for the Implementation of Upset Prevention and Recovery Training: 1st Edition*”, International Air Transport Association, Montreal – Geneva, June, 2015

ICAEA, “*Guidelines for Aviation English Training Programmes*”, International Civil Aviation Organisation, 2009, ISBN: 978-92-9231-403-3.

ICAO, “*Manual of Evidence-based Training*”, International Civil Aviation Organisation, DOC 9995 AN/497, 2013.

ICAO, “*Aviation Training Going Forward*”, ICAO Training Report, Vol. 2, No. 2, 2014.

ICAO, “Integration of Human Factors in Research, Operation, and Acquisition”, *Second Meeting of the APANPIRG ATM Sub-Group*, Hong-Kong, China, 4-8 August, 2014.

ICF, “*Optimizing The Mix Between Virtual and Live Military Training*”, Office of the Under Secretary of Defence for Personnel and Readiness, ICF International, 12 September, 2013.

IEEE, “*Standard Glossary of Software Engineering Terminology*”, IEEE Standard 610.12, 1990.

IEEE, “*US International Traffic in Arms Regulations*”, IEEE Advancing Technology for Humanity, <https://www.ieee.org/about/corporate/compliance/legal/itar.html>, [Accessed: 12/12/2014].

IFALPA, “*The Future of Flight Training*”, The Global Voice of Pilots, POSITION, 11POS04, March, 2011.

IFALPA, “*Pilot Training Standards: Guide For Best Practices*”, The International Federation of Airline Pilot’s Association, 2012.

INCOSE, “*Systems Engineering Handbook: A Guide for Systems Lifecycle Processes and Activities: 4th Edition*”, John Wiley & Sons, 2015, ISBN: 978-1118999-400.

Iqbal, Z., Grigg, N.P., Govinderaju, K., Campbell-Allen, N., “Statistical Comparison of Final Weight Scores in Quality Function Deployment (QFD) Studies”, *International Journal of Quality & Reliability Management*, Vol. 31, Issue:2, pp. 184-204, 2014.

ISO, “*Ergonomic of Human-system Interaction - Part 210 – Human-centred design for interactive systems*”, ISO 9241 – 210:2010, International Standards Organisation, Geneva, 2010.

ISO/IEC/IEEE, “*Systems and Software Engineering System Lifecycle Processes*”, International Standards Organisation Document, 2015.

ISO/IEC/IEEE, “*Systems and Software Engineering – Life cycle processes – Requirements Engineering*”, Geneva, Switzerland: International Organisation for Standardization (ISO)/International Electrotechnical Commission (IEC), Institute of Electrical and Electronics Engineers. ISO/IEC/IEEE: 29148:2011, 2011.

ISO/IEC/IEEE, “*Systems and Software Engineering – Architecture Descriptions*”, Geneva, Switzerland: International Organisation for Standardization (ISO)/International Electrotechnical Commission (IEC)/Institute of Electrical and Electronics Engineers (IEEE), ISO/IEC/IEEE: 42010, 2011.

ISO/IEC, “*Systems and Software Engineering – System Life Cycle Processes*”, Geneva, Switzerland: International Organisation for Standardization (ISO)/International Electrotechnical Commission (IEC), ISO/IEC/IEEE 15288:2008 E, 2008.

ISO/IEC/IEEE 15288, “*Systems and Software Engineering – System Life Cycle Processes*”, International Standards Organisation for Standardisation / International Electrotechnical Commissions / Institute of Electrical and Electronic Engineers, ISO/IEC/IEEE 15288:2015, Geneva, Switzerland, 2015.

Iyengar, R., Mahal, A.R., Aklilu, L., Sweetland, A., Karim, A., Shin, H., “The Use of Technology for Large-scale Education Planning and Decision Making”, *Information Technology for Development*, Routledge Taylor & Francis Group, July, 2014, DOI: 10.1080/02681102.2014.940267.

Jaiswal, E.S., “A Case Study on Quality Function Deployment (QFD)”, *IOSR Journal of Mechanical and Civil Engineering (IOSR-JMCE)*, Vol. 3., Issue 6, pp. 27-35, Nov-Dec, 2012.

Jacko, J.A., “*The Human-Computer Interaction Handbook: Fundamentals, Evolving Technologies, and Emerging Applications, 3rd Edition*”, CRC Press, Taylor & Francis Group, 2012, ISBN: 978-1-4398-2944-8.

Jean, V.G., “*Success of Simulation-Based Training in Tough to Measure*”, National Defence, National Defence Industrial Associates, December, 2008.

Jeannot, E., Kelly, C., Thompson, D., “*The Development of Situation Awareness Measures in ATM Systems*”, HRS/HSP-005-REP-01, Bretigny-sur-Orge, France: Eurocontrol Experimental Centre, 2003.

Jenkins, D.F., Stanton, N.A., Walker, G.H., Salmon, P.M., & Young, M.S., “Using Cognitive Work Analysis to Explore Activity Allocation within Military Domains”, *Ergonomics*, 51(6), pp. 798-815, 2008.

Jirgl, M., Haulikam, M., & Bradane, Z., “The Dynamic Pilot Behavioural Models”, 25th DAAM International Symposium of Intelligent Manufacturing and Automation, Vol.100, pp. 1192-1197, 2014.

Johnson, C.E., “*Organisational Ethics: A Practical Approach*”, Sage Publications, 2012, ISBN: 9781412987967.

Johnson, P.W., “*Evaluating systems of systems against mission requirements*”, Ph.D. dissertation, Loughborough University, UK, 2010.

Johnson, P.A., Estrada, A., Braithwaite, M.G., Manning, J.C., “*Assessment of Simulated Spatial Disorientation Scenarios in Training US Army Aviators*”, USAARL Report No. 2000-06, Aircrew Health and Performance Division, US Army Aeromedical Research Laboratory, November, 1999.

Johnson, R.B. & Onwuegbuzie, A.J., “Mixed Methods research: A research paradigm whose time has come”, *Educational Researcher*, 33(7), pp. 14-26, 2004.

Jou, Y.T., Lin, W.T., Lee, W.C., & Yeh, T.M., “Integrating the Taguchi Method and Response Surface Methodology for Process Parameter Optimization of the Injection Molding”, *Applied Mathematics & Information Sciences*, 8(3), pp. 1277-1285, 2014.

JSP 822, “*Training Management Policy: The Governance of Joint Training and Education Requirements*”, Part3 – Chapter 1, Training Educational Skill and Resettlement Division, Crown, 2012.

JSP 822, “*Part 4: DSAT QS 001:2008*”, British Standards Institute, Director General Training and Education, December, 2008.

Karlsen, J.E., “Design and Application for a Replicable Foresight Methodology Bridging Quantitative and Qualitative Expert Data”, *European Journal of Futures Research*, Vol. 2, No. 1, pp. 1-12, 2014.

Katajavouri, N., Lindblom-Ylanne, S., & Hirvonen, J., “The Significance of Practical training in Linking Theoretical Studies with Practice”, *Higher Education*, 51, pp.439-464, 2006.

Kavyaganga, K., Tweedale, J.W., Thatcher, S., & Jain, L.G., “Recent Advances in ‘Soft’-computing: Theories and Application”, *Journal of Intelligent & Fuzzy Systems*, Vol. 24, No. 3, pp. 457-466, 2013.

Kecea, N., Garbajosa, J., Bourque, P., “Modeling Functional Requirements to Support Traceability Analysis”, *IEEE International Symposium on Industrial Electronics*, Montreal, Quebec, 9-13 July, pp. 3305-3310, 2006.

Keen, P.G.W., “Value Analysis: Justifying Decision Support Systems”, *Decision Support Systems: Putting Theory into Practice*, 3rd edition, R.H. Sprague & H.J. Watson, Prentice Hall, New Jersey, pp. 57-73, 1993.

Keller, J.W., Leiden, K., & Small, R., “Cognitive Task Analysis of Commercial Jet Aircraft Pilots during Instrument Approaches for Baseline and Synthetic Vision Displays”, *Proceedings of the 2003 NASA Aviation Safety Program Conference on Human Performance Modelling of Approach and Landing with Augmented Displays*, NASA/CP-2003-212267, Moffett Fields, CA: NASA Ames Research Center, 2003.

Keller, M., Schnell, T., Lemos, K., “*Pilot performance as a function of display resolution and field of view in a simulated terrain following flight task using a synthetic vision system*”, University of Iowa, Operator performance Laboratory, Langley, VA, 2003b.

Kelly, T. & Weaver, R., “The goal structuring notation – a safety argument notation”, *In Dependable Systems and Networks 2004 Workshop on Assurance Cases*, 2004.

Kennedy, P., Hartholt, A., Gratch, J., Swartar, W., Traum, D., Marsella, S., & Piepol, D., “Building Interactive Virtual Humans for Training Environments”, *Proceedings of the Interservice/Industry Training, Simulation, and Education Conference (I/ITSEC)*, 2007.

Kennedy, Q., Taylor, J.L., Reade, G., & Yesavage, J.A., "Age and Expertise Effects in Aviation Decision Making and Flight Control in a Flight Simulator", *Aviation Space Environment Medicine*, 81(5), pp. 489-497, May, 2010.

Kerkering, J.C., "Eliciting and Analyzing Expert Judgment, A Practical Guide", *Technometrics*, 44(2), pp. 183-194, January, 2012.

Kilpatrick, D.L., "Evaluation of training", *Training and development handbook: A guide to human resource development*, New York: McGraw-Hill, 1976.

Kirby, B., Fletcher, G., & Dudfield, H., "Live Virtual Constructive Training Blend Optimisation Study", BAE Systems, Quintec, 2011.

Kirkwood, A., & Price, L., "Technology-enhanced Learning and Teaching in Higher Education: What is 'Enhanced' and How do we Know?: A Critical Literature Review", *Learning, Media, and technology*, 39(1), pp. 6-36, 2014.

Kleer, M., Gizatuilin, A., Drebler, K., & Muller, S., "Real Time Human in the Loop MBS Simulation in the Fraunhofer Robot-based Driving Simulator", *Archive of Mechanical Engineering*, Vol. 61, Issue 2, pp. 270-285, August, 2014.

Klein, G.A., "The recognition-primed decision (RPD) model: Looking back, Looking Forward", *Naturalistic Decision Making*, Mahwah, NJ: Lawrence Erlbaum Associates, pp. 285-292, 1997.

Klir, G.J., "IPO Dictionary / Methodology Committee draft", Facets of Systems Science, Plenum Press, September, 1991.

Korteling, J.E., Oprins, E.A., & Kallen, V.L., "Measurement of Effectiveness for Training Simulations", North Atlantic Treaty Organisation (NATO), RTO-SAS-095, 2013.

Kossiakoff, A., Sweet, W., "Systems Engineering: Principles and Practice", Hoboken, NJ, USA: John Wiley and Sons, 2009.

Kozlowski, S.W.J., "Training and Developing adaptive teams: Theory, Principles, and Research", *Decision making under stress: Implications for training and simulation*, pp. 115-153, Washington, DC: APA Books, 1998.

Kozlowski, S.W.J., Gully, S.M., Brown, K.G., Salas, E., Smith, E.A., Nason, E.R., "Effects of training goals and goal orientation traits on multi-dimensional training outcomes and performance adaptability", *Organisational Behaviour and Human Decision Processes*, No. 85, pp.1-31, 2001.

Kozuba, J., & Bondaruk, A., "Flight Simulator as an Essential Device Supporting the Process of Shaping Pilot's Situational Awareness", *International Conference of Scientific Paper: AFASES 2014*, Brasov, 22-24 May, 2014.

Krausert, A., “The Timing of Training Effects”, *IRRU Research Seminar*, Warwick Business School, 18 November, 2015.

Kriovile, A., Addamssiri, N., & Gadi, T., “An MDA Method for Automatic Transformation of Models from CIM to PIM”, *American Journal of Software Engineering and Applications*, Vol. 4, No. 1, pp. 1-14, 2015.

Kronqvist, A., Jokinen, J., & Rousi, R., “Evaluating the Authenticity of Virtual Environments: Comparison of Three Devices”, *Advances in Human-Computer Interaction*, Volume 2016, pp. 14, 2937632, 2016.

Kruse, S., “Co-Evolution of Metamodels and Model Transformations: An Operator-based, Stepwise Approach for the Impact Resolution of Metamodel Evolution on Model Transformations”, Books on Demand, Hersteing and Verlag, 2015, ISBN: 978-3-7386-2894-4.

Lane, J. & Valerdi, R., “Synthesising system-of-systems concepts for use in cost estimation”, *Systems Engineering* 10, No.4: pp. 297-307, 2007.

Laudeman, I.V, Palmer, E.A., “Quantitative Measurement of Observed Workload in the Analysis of Aircrew Performance”, *International Journal of Aviation Psychology*, 5(2), pp/. 187-197, 1995.

Lamsweerde, A.V, “*Requirement Engineering: From System Goals to UML Models to Software Specifications*”, Wiley, 2009.

Lateef, F., “Simulation-based Learning: Just Like the Real Thing”, *Journal of Emergencies, trauma, and Shock*, 3(4), pp. 348-352, Oct-Dec, 2010.

Lee, A.T., “*Flight Simulation: Virtual Environments in Aviation*”, Ashgate Publications, 31 May, 2005. ISBN: 978-0759642879.

Levenchuck, A., “Towards A Systems Engineering Essence”, *99th INCOSE Russian Chapter Meeting*, 28 January, 2015.

Leveson, N., Dulac, N., Marias, K., & Carroll, J., “Moving Beyond Normal Accidents and High Reliability Organisations: A Systems Approach to Safety in Complex Systems”, *Organisation Studies*, Vol. 30, No. 2-3, pp. 227-249, 2009.

Levy, C.P., “*A Comparison Study of F-15C Fighter Squadron Ready Aircrew Program Flying Hour Schedule vs. the RAND Corporation Flying Hour Scheduling Linear Program*”, Air Force Institute of Engineering and Management, ADA450490, 13 June, 2006.

Leung, Y., “*Spatial Analysis and planning under Imprecision*”, North Holland, 1988.

Li, W., & Harris, D., "Identifying Training Deficiencies in Military Pilots by Applying Human Factors Analysis and Classification System", *International Journal of Occupational Safety and Ergonomics (JOSE)*, Vol. 19, No. 1, pp. 3-18, 2013.

Liang, R., Zhang, M., Lui, Z., & Krokes, M., "Cognitive and Synthetic Behaviour of Avatars in Intelligent Virtual Environments", *Virtual Reality*, 1 March, 2008.

Lilly, S., "How to Avoid Use-Case Pitfalls", *Software Development Magazine*, 2000.

Lippi, V., Avizzano, C.A., Mattet, D., & Ruffaldi, E., "Effect of Delay on Dynamic Targets Tracking Performance and Behaviour in Virtual Environments", *RO-MAN*, 13-15 September, pp. 446-451, 2010.

Littlewood, B. & Wright, D., "The use of multilegged arguments to increase confidence in safety claims for software based systems", *IEEE Transactions on Software Engineering (TSE)*, 33(5), pp. 347-365, 2007.

Liu, S., & Zarate, P., "Knowledge based Decision Support Systems: A Survey on Technologies and Applications Domain", *Joint International Conference between INFORMS and EWG-DDS*, Toulouse, France, Vol. 180, pp. 62-66, 31 July, 2014.

Liu, S., Smith, M., Tuck, S., Pan, J., Alkuraiji, A., & Jayawickrama, O., "Where can Knowledge-based Decision Support Systems Go in Contemporary Business management", *Journal of Economics, Business and Management*, Vol. 3, No. 5, May, 2015.

Londsberry, C.R., Van-Biskirk, W.L., Astwood, R.S., Mercado, A.D., Aatere, A.J., "Adoptive Training Considerations for Use in Simulation-based Systems", *Naval Air Warfare Center Training Systems*, ADA535421, October, 2010.

Longridge, T., Burki-Cohen, J., Go, T.H., Kendra, A.J., "Simulator Fidelity Considerations For Training and Evaluation of Today's Airline Pilots", *Proceedings of the 11th International Symposium of Aviation Psychology*, Columbus, OH, 5-8 March, 2001.

Madahar, M., Cleary, P., & Ball, D., "Categorisation of Spreadsheet Use Within Organisations, Incorporating Risk: A Progress Report", *Mathpubs Scientific Publication Data*, 2008.

Madni, A.M., & Sievers, M., "Systems of Systems Integration: Key Considerations and Challenges", *Systems Engineering Journal*, Vol. 17, Issue 3, pp. 330-347, 2014.

Magableh, B., Rewashdeh, B., & Barret, S., "A Framework for Evaluating Model-driven Architecture", *American Journal for Software Engineering and Applications*, Vol. 1, Issue 1, pp. 10-22, December, 2012.

Magnusson, S., Berggen, P., "Dynamic assessment of Pilot Mental Status", *Proceedings of the Human Factors and Ergonomics Society 46th Annual Meeting*, Baltimore, MD: Santa Monica, CA, 2002.

Magnusson, S., “Similarities and differences in psychophysiological reaction between simulated and real air-to-ground missions”, *International Journal of Aviation Psychology*, Vol. 12(1), pp. 49-61, 2002.

Maguire, M., “*Usability*”, Ergonomics and Safety Systems, Carl Sandon & Rodger Harvey, ESRI, Loughborough University, Loughborough, 2003.

Maier, M., & Rechtin, E., “*The art of Systems Architecting: 2nd Edition*”, Boca Raton, CRC Press LLC, 2000.

Maier, M., & Rechtin, E., “*The Art of System Architecting: 3rd Edition*”, Boca Raton, FL, USA: CRC Press, 2009.

Malmin, O.K., & Reibling, L.A., “*The Contribution of Aircraft Simulators to the Training and Readiness of Operational Navy Aircraft Squadrons*”, CRM-143, Alexandria, VA: Centre for Naval Analysis, 1995a.

Malimin, O.K., & Reibling, L.A., “*Using Simulators to Train Fleet Aviators*”, CRM-95-50, Alexandria, VA: Centre for Naval Analysis, 1995b.

Maltby, J., Day, L., & Macaskill, A., “*Personality, Individual Differences and Intelligence: 2nd Edition*”, Pearson Education Limited, ISBN: 978-0-273-72290-8, 2010.

Mantovani F., & CastelNouve, G., “*Sense of Presence in Virtual Training: Enhancing Skills Acquisition and transfer of Knowledge through Learning Experience in Virtual Environments*”, Being There: Concepts, Efforts and Measurement of User Presence in Synthetic Environments, IOS Press, Amsterdam, 2003.

Marken, R.S., Taylor, W.W., Ausink, J.A., Hanser, L.M., Anderegg, C.R., & Wickman, L., “*Absorbing and Developing Qualified Fighter Pilots: The Role of the Advanced Simulator*”, RAND: Project Air Force, 2007, ISBN: 978-08330-4154-8.

Martin, B.O., Kolomitro, K., & Lam, T.M., “Training Methods: A Review and Analysis”, *Human Resource Development Review*, Vol. 13, No. 1, March , 2014.

Martin, W.L., Murray, P., & Bates, P., “The Effects of Startle on Pilots During Critical Events: A Case Study Analysis”, *30th EAAP Conference: Aviation Psychology & Applied Human Factors – Working Towards Zero Accidents*, 2012.

Martin, W., Murray, P., & Bates, P., “Flight Training: Pathological Reactions”, *Aero Safety World*, May, 2013.

Mason, M.A., “EASA Automation Policy: Bridging Design and Training Principles”, EA Sp Action EME4.4 – Develop an Automation Policy, 28 May, 2013.

Mastashari, A., McComb, S.A., Kennedy, D.M., Cloutier, R., & Kurfiatis, P., “Developing a Stakeholder-assisted Agile CONOPS Development Process”, *INCOSE Systems Engineering*, Vol. 15, Issue 1, pp. 1-13, Spring, 2012.

Matthews, M.D. & Beal, S. A., “*Assessing Situation Awareness in Field Training Exercises*”, U.S. Army Research Institute for the Behavioural and Social Sciences. Research Report 1795, 2002.

Mavin, T.J., Roth, W.M., & Dekker, S., “Understanding Variance in Pilot Performance Ratings”, *Aviation Psychology and Applied Human Factors*, Vol. 3(2), pp. 53-62, 2013.

Mavin, T.J., & Roth, M.W., “Between Reflection on Practice and the Practice of Reflection: A Case Study from Aviation”, *Reflective Practice: International and Multidisciplinary Perspectives*, Vol. 15, Issue 5, 2014.

Mavin, T.J., & Roth, W.M., “Optimizing a Workplace Learning Pattern: A Case Study from Aviation”, *Journal of Workplace Learning*, Vol. 27, Issue 2, pp. 112-127, 2015.

Mavris, D.N., & Kirby, M.R., “Technology Identification, Evaluation, and Selection for Commercial Transport Aircraft”, *Presented at the 58th Annual Conference of Society of Allied Weight Engineers*, San Jose, CA, May 24-26, 1999.

Mavris, N. & Dickerson, C.E. “Relational Orientated Systems Engineering and technology trade-off Analysis (ROSETTA) Environment”, *System of Systems Engineering (SoSE)*, 6th International Conference, June, 2011.

McAllister, R.P., “*Overcoming KC-10 Formal Training Unit Pilot Production Challenges*”, Air-Force Institute of Technology, Wright-Patterson Air-Force Base, Ohio, June, 2013.

McGrath, S.R., “*Leveraging Simulation Against the F-16 Flying Training Gap*”, Centre for Strategic and Technology, Air University, November, 2005.

McGuinness, B. & Foy, L., “A subjective measure of SA: the Crew Awareness Rating Scale (CARS)”, *Presented at the Human Performance, Situational Awareness and Automation Conference*, Savannah, Georgia, pp. 16-19, October, 2000.

McGuinness, B., “Quantitative Analysis of Situational Awareness (QUASA): Applying Signal Detection Theory to True/False Probes and Self-Ratings”, *Command and Control Research and Technology Symposium: The Power of Information Age Concepts and Technologies*, San Diego, California, 15th – 17th June 2004.

McHugh, J.M., & Casey, Jr., G.W., “*US Army Posture Statement*”, http://secureweb2.hqda.pentagon.mil/VDAS_ArmyPostureStatement/2010/, 2010, [reviewed Dec 2013].

McLean, G., Wise, L., & Williams, B., "Game-Like Digital Training Tools: Transfer of Cognitive and Perceptual Skills From Static to Dynamic Interfaces", *ASCILITE 2012*, Wellington, New Zealand, 25-28 November, 2012.

Meehan, M., Razzaque, S., Whitton, M.C., Brooks, F.P., "Effect of Latency on Presence in Stressful Virtual Environments", *Proceedings VR03 of IEEE Virtual Reality*, p. 141, 2003.

Mehler, B., Reimer, B., Zec, P.M., "Defining Workload in the Context of Driver State Detection and HMI Evaluation", *Proceedings of the 4th International Conference on Automotive User Interfaces and Interactive Vehicular Application*, New York, USA, pp. 187-191, 2012.

Meyer, M. & Booker, J., "*Eliciting and Analysing expert Judgement: A practical Guide*", Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, USA, 2001.

Meyer, G.F., Wong, L.T., Timson, E., Perfect, P., & White, M.D., "Objective Fidelity Evaluation in Multisensory Virtual Environments: Auditory Cue Fidelity in Flight Simulation", *PLoS ONE*, 7(9), 5 September, 2012.

Meyers, R.H., Montgomery, D.C., Christine M., "*Response Surface Methodology: Process and Product Optimisation using Design Experiments*", Hoboken, NJ: John Wiley & Sons, 2009.

Meystel, A.M. & Albus, J.S., "*Intelligent Systems: Architecture, Design and Control*", New York: Wiley, 2001.

Miller, L.E., Ruiz, J.R., Sharp, S.E., "Student and Instructor Perceptions of Effective Instructional methods in an University Aviation Human Factors Course", *Journal of Aviation / Aerospace Education & research*, Vol. 8, No.1, JAAER, 1997.

Miller, S., "Workload Measures: Literature Review", National Advanced Simulator: University of Iowa, N01-006, August, 2001.

Miragila, S., "*Systems Architecture and Innovation: The Modularity-Integrality Framework*", Cambridge Service Alliance, University of Cambridge, 2014.

Mitchell, H.F., "The Hawk Story", *Journal of Aeronautical History*, Paper No., 2013/01, 2013.

MOD, "Defence Issues: Acquisition and Procurement", <http://webarchive.nationalarchives.gov.uk/+/mod.uk/issues/>, 2004. [Accessed: 13/06/2014].

MOD, "Flight Training Report – First Pilots Complete New Fast Jet Training", Armed Forces and Ministry of Defence Reform, 13 June, 2013.

MOD, "Military Flight Training", HC81, Session 2015-2016, 12 June, 2015.

Moehlenbrunt, C., Jakobi, J., Paperfuss, A., “*The Role of Workload for Work Organisation in a Robotic Tower Control Centre*”, 9th USA/Europe AIT Traffic Management Research and Development Seminar (ATM 2011), Berlin, Germany, 14-17 June, 2011.

Mohammed, S. & Ringsies, E., “Cognitive diversity and consensus in Group decision making: The role of inputs, processes, and outcomes”, *Organizational Behaviour and Human Decision Processes*, 85, No.2, pp. 310-335, July, 2001.

Morie, J.F., Verhulsdonck, G., Lauria, R.M., & Keeton, K.E., “*Operational Assessment Recommendations: Current Potential and Advanced Research Directions for Virtual Worlds as Long-Duration Space Flight Countermeasures*”, Report for the Centre for Aerospace Information, NASA/TP-2011-216164, October, 2011.

Mossé, F.G., "Modeling Roles - A Practical Series of Analysis Patterns", *Journal of Object Technology*, vol. 1, no. 4. pp. 27–37, 2002.

Mowbray, C.T., Holter, M.C., Teague, G.B., Bybee, D., “Fidelity Criteria: Development, Measurement, and Validation”, *American Journal of Evaluation*, 24(3), 2003.

Mulalic, A., Mohd, P., Ahmad, F., “Perceptual Learning Styles of ESL students”, *European Journal of Social Sciences*, 7(3), pp. 101-113, 2009.

Muller, G., “Systems Engineering Research Methods”, *Procedia Computer Science*, Conference on Systems Engineering Research, Vol. 16, pp. 1092-1101, 2013

Muller, J., & Mukerji, J., “*MDA Guide Version 1.0.1*” Object Management Group, 12 June 2003.

Murray, J., “Model Based Systems Engineering (MBSE) Media Study”, 2 May, 2012, <http://syse.pdx.edu/program/portfolios/julia/MBSE.pdf>, [Accessed: 10/09/2015].

Naderpour, M., Lu, J., & Zhang, G., “A Safety Critical Decision Support Systems Evaluation Using Situation Awareness and Workload Measures”, *Reliability Engineering & Systems Safety*, Vol. 150, pp. 147-159, June, 2015.

Nahlinder, S., “Flight Simulator Training: Assessing the Potential”, Linköping Studies in Science and Technology Dissertations, No. 1250, Linköping University Institute of Technology, Sweden, 2009.

NAP, “*Opportunities for the Employment of Simulation in U.S. Ait-Force Training Environments: A Workshop Report*”, The National Academies Press, July, 2015.

NAUMC, “Aviation Training and Readiness Program”, Document No. NAVVMC 3500.14C, 23 August, 2011.

NDCAF, “*Evaluation of Air Force Training and Readiness Part 1 – Air Force Initial Occupational Training*”, National Defence and the Canadian Armed Forces, CRS1258-188, November, 2012.

Nemanti, H.R., Steiger, D.M., Iyer, L.S., & Herschel, R.T., “Knowledge Warehouse: an architectural Integration of Knowledge Management, Decision Support, Artificial Intelligence and Data Warehousing”, *Decision Support Systems*, Vol.33, pp. 143-161, 2002.

Neumann, V.J., “The Mathematician”, *Works of the Mind*, by R.B.Heywood, Chicago: University of Chicago Press, pp. 180-196, 1947.

Neville, B, “*Report on the Effectiveness of the WINGS – Pilot Proficiency Program*”, FAASTeam Outreach Program Accident Analysis, 15 September, 2011.

Nisangala, A., Weerasinghe, M., Dlos, G.K.A., & Sandarvnay, D., “Flight Simulator for Serious Gaming”, *European Journal of Computer Science and Information Technology*, Vol. 2, No. 1, pp. 53-63, March, 2015.

Nofi, A.A., “*Defining and Measuring Shared Situational Awareness*”, Center for Naval Analysis, CRM D00002895.A1/Final, November, 2000.

Northrop, L., Fieler, P., Klien, M., “*Ultra-large scale systems: The Software Challenge of the Future*”, Software Engineering Institute, Pittsburgh, 2006.

Numrich, S.K., & Tolk, A., “*Challenges for Human, Social, Cultural, and Behavioural Modelling*”, SCS M&S Magazine, pp. 1-9, January, 2010.

Nyssen, A.S., Larbuisson, R., Janssens, M., Pendeville, P., & Mayne, A., “A comparison of the training value of two types of anaesthesia simulators: computer screen-based and mannequin-based simulators”, *Anaesthesia & Analgesia*, Vol. 94, pp. 1560-1565, 2002.

Nyssen, S., Larbuisson, R., Janssens, M., Pendeville, P., Mayne, A., “A comparison of the training value of anaesthesia simulators: Computer screen-based and mannequin-based simulators”, *Anaesthesia & Analgesia*, Vol. 94, pp. 1560-1565, 2002.

O’Hagan, A., Buck, C., Daneshkhah, J., Eiser, P., Garthwaite, D., Jenkinson, D., Oakley, J., Rakow, T., “*Uncertain Judgements: Eliciting Experts’ Probabilities*”, Wiley, West Sussex, England, 2006.

Olson, G.A. & Austin, J., “A step towards early PC-based training that reduces risk: The effects of practicing an “instrument referenced” skill pattern on “visual referenced” performance of beginning flight students”, *Journal of Aviation/Aerospace Education and Research*, 14(2), pp. 47-51, 2005.

Olson, J., “Tactile display technologies as an enabler for space exploration operations”, *IEEE Aerospace Conference*, Big Sky, Montana, 3-10 March, pp. 1-12, 2007.

OMG SysML, “Systems Modelling Language Specification”, <http://www.omg.org/spec/sysml/current>, 2012.

Orasanu, J.M., “*Shared Mental Model and Crew Decision Making*”, Cognitive Science Laboratory Report #46, Princeton, NJ: Princeton University, 1990.

Oxford University Press (OUP), “*Oxford Dictionaries Online*”, 2013

Panteli, M., & Kirschen, D.S., “Situation Awareness in Power Systems: Theory, Challenges and Applications”, *Electric Power Systems Research*, Vol. 122, pp. 140-151, May, 2015.

Parasuraman, R., Hansman, J., & Brussolari, S., “*Framework for Evaluation of Human-System issues with ASDE-X and Related Surface Safety Systems*”, White paper for AAR-100, Washington DC: Federal Aviation Administration, 2002.

Park, D., Allen, R.W., Rosenthal, T.J., Fiorentino, D., “Training Effectiveness: How Does Driving Simulator Fidelity Influence Driver Performance”, *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, Vol. 49, No. 25, pp. 2201-2205, 2005.

Parvar, H., Fesharaki, M.N., Moshiri, B. “Shared Situation Awareness Architecture (S2A2) for Network Centric Disaster Management (NCDM)”, *International Journal of Computer Science*, Vol. 9, Issue 4, No. 2, July, 2012.

Pawar, B.S., “*Theory Building for Hypothesis Specification in Organisational Studies*”, SAGE Publication, 2009, ISBN: 978-81-321-0244-1.

Pease, G.F., “*Aircrew Management: Flying Operations*”, Air Force Instruction, 11-412, 10 December, 2009.

Peignat, J., Peneranda, A., Amabile, S., & Marcel, G., “Strategic Decision Support Systems for Local Government: A Performance Management Issue?”, *International Business Research*, Vol. 6, No. 2, 2013.

Perey, P., “Future Simulation Technologies”, *Proc. Royal Aeronautical Society Annual International Flight Crew Training Conference*, Sept 25-25, London, UK, 2008.

Perrota, C., Featherstone, G., Aston, H., & Houghton, E., “*Game-based Learning: Latest Evidence and Future Directions*”, The NFER Research Programme, Slough, NFER, April, 2013.

Persing, T.R., Bellish, C.A., “An investigation of the impact of aerodynamic model fidelity on close-in combat effectiveness prediction in piloted simulation”, *Paper presented at the AIAA Modelling and Simulation Techniques Conference and Exhibit*, pp. 15-18, San Francisco, California, August, 2005.

Petford, C., & Frank, S., “*Synthetic Environment Interoperability – Barriers and Enablers*”, NATO Science and Technology Organisation Collaboration Support Office, October, 2009, ISBN: 978-92-837-0076.

Petitt, K., “*Structural Redesign of Pilot Training for Aviation Safety in Automated Aircraft*” Emberry-Riddle Aeronautical University, 2013, http://www.kiwi-cms.nl/bureau/graphic-dynamics/mps/upload/files/Automation_Safety_Article.pdf, [Accessed: 15/04/2014].

Petrie, J., Basson, L., Stewart, M., Notten, P., & Alexander, B., “Decision Making for Design of Cleaner Processes: A Life Cycle Management Perspective”, *Proceedings of the First International Conference on Life Cycle Management: Bridging the Gap between Science and Application*, Copenhagen, Denmark, 2001.

Pirella, F.D. & Kappler, W.D., “*Identification and evaluation of scale design principles in the development of the Sequential Judgement, extended range scale*”, Wachtberg: Forschungsinstitut für Anthropotechnik, FAT Report No. 80, 1988.

Pirro, G. & Talia, D., “UFOME: An Ontology Mapping System with Strategy Prediction Capability”, *Data and Knowledge Engineering*, 69(5), pp.444-471, 2010.

Plat, P.A., Dahn, D.A., & Amburn, P., “Low-Cost Approaches to Virtual Flight Simulation”, *Proceedings of the IEEE National Aerospace and Electronics Conference*, NAECON, Vol. 23, pp. 940-946, 1991.

Plioutsias, A., & Karanikas, N., “Using STPA in the Evaluation of Fighter Pilots Training Programs”, *Proceedings of the 3rd European STAMP Workshop*, Vol. 128, pp. 25-34, Amsterdam, 5-6 October, 2015.

Pokoradi, L., “Sensitivity Investigation of Fault Tree Analysis with Matrix-Algebraic Method”, *Theory and Application of Mathematics & Computer Science*, pp 34-44, 2011.

Pool, D.M., Zaal, P.M.I., Damveld, H.S., Reason, M.M.J., Mulder, M., “Evaluating Simulator Motion Fidelity Using In-flight and Simulator Measurements of Roll Tracking Behaviour”, *AIAA Modeling and Simulation Technology Conference*, Minneapolis, Minnesota, 13-16 August, 2012.

Pourshadid, A., Johari, I., Richards, G., Amyot, D., & Akhigbe, O.S., “A Goal-Oriented, Business Intelligence-Supported Decision-making Methodology”, *Decision Analytics*, 1(9), December, 2014.

Prensky, M., “*True Believers: Digital Game-Based Learning in the Military*”, Digital Game-Based Learning, McGraw-Hill, 2001.

Pritchett, A.R., “Pilot Situation Awareness and Alerting Systems Comments”, *Journal of Aerospace: Advances in Aviation Safety Conference & Exposition*, SAE Technical Paper 981224, pp. 321-354, 2012.

Probasco, L., “*The Ten Essentials of RUP*”, The Rational edge, 2000.

Quadri, I.R., Gamatie, A., Boulet, P., Meftali, S., & Dekeyser, J.L., “Expressing Embedded Systems Configurations at High Abstraction Levels with UML MARTE Profile: Advantages,

Limitations and Alternatives”, *Journal of Systems Architecture*, Vol. 58, Issue 5, pp. 178-194, April, 2012.

Raistrick, C., Frances, P., Write, J., Carter, C., & Wilkie, I., “*Model Driven Architecture with Executable UML*”, Cambridge University press, 2004, ISBN: 0521-53771-1.

RAF, “RAF Recruitment”, <http://www.raf-mod.uk/recruitment/>, 2013, [Accessed: 7/6/2013].

Rantz, W.G., Dickerson, A.M., Sinclair, G.A., & Houten, R.U., “The Effect of Feedback on the Accuracy of Checklists Completion During Instrument Flight Training”, *Journal of Applied Behaviour Analysis*, 42(3), pp. 497-509, 2009.

RAS (Royal Aeronautics Society), “*The Impact of Flight Simulation in Aerospace*”, Royal Aeronautical Society Simulation Group, March, 2009.

Rajkoomar, M., “Blended Learning in Library and Information Science (LIS) Education and Training”, *Proceedings of the IATUL Conferences*, Paper 1, 2013.

Ratwani, K., “*Game-based Training Effectiveness Evaluation in an Operational Setting*”, United States Army Research institute for the Behavioural and Social Sciences, Study Report 2010-02, 2010.

Rayneri, L., Gerber, B.L., & Wiley, L.P., “The Relationship between Classroom Environment and the Learning Style Preferences of Gifted Middle School Students and the Impact of Levels of Performers”, *Gifted Child Quarterly*, 50(2), pp. 104-118, 2006.

Regli, W.C., Mayk, I., Cannon, C.T., Kopena, J.B., & Lass, R.N., “Development and Specification of a Reference Architecture for Agent-Based Systems”, *IEEE Transactions of Systems, Man, and Cybernetics*, Vol. 44, No. 2, February, 2014.

Reichardt, C.S. & Rallis, S.F., “Qualitative and Quantitative Inquiries are not Incompatible: a call for a new partnership”, *The qualitative-quantitative debate: New Perspectives*, San Francisco: Joey-Bass, pp. 85-92, 1994.

Rewete, S., “PC-Based Aviation Training Devices for Pilot Training in Visual Flight Rules Procedures: Development, Validation and Effectiveness”, Doctor of Philosophy Thesis, Massey university, New Zealand, 2014.

Ribino, P., Ausello, A., Giuseppe, C.R., & Gaglio, S., “A Knowledge Management and Decision Support Model for Enterprises”, *Advances in Decision Science*, Vol. 2011, p.16, 2011.

Ricci, N., Rhodes, D.H., & Ross, A.M., “Evolvability-Related Options in Military Systems of Systems”, *Conference on Systems Engineering Research*, Vol. 28, pp. 314-321, 2014.

Ridha, M.B., Alnaja, L., & Kalal, M.A., “Decision Support Systems and Their Role in Rationalising the Production Plans: A Case Study on a Plant in Nayaf”, *Industrial Engineering Letters*, Vol. 3, No. 4, 2013.

Robert.C, & Casella, G., "*Monte Carlo Statistical Methods*", Springer-Verlag New York, Inc., Secaucus, NJ, USA, 2005.

Roessingh, J.J.M., & Verheaf, G.G., "*Training Effectiveness of Embedded Training in a (Multi-)Fighter Environment: A Discussions Paper*", NATO-RTO-MP_HFM-169, 2009.

Rolland, C., "Conceptual modelling and Natural Language Analysis", *Seminal Contributions to Information Systems Engineering*, pp. 57-61, May, 2013.

Rosenberg, D. & Scott, K., "*Driving Design: The Problem Domain*", Software Development magazine, 2001.

Rostler, B.D., & Nemfakes, C., "*Building Towards an Unmanned Aircraft System Training Strategy*", RAND National Security Research Division, RAND Corporation, 2014.

Rothchild, I., "*Induction, Deduction, and the Scientific Method: An Eclectic Overview of the Practice of Science*", The Society of the Study of Reproduction, 2006.

Sage, A. & Cuppan, C., "On the systems engineering and management of systems of systems and federations of systems", *Information, knowledge and systems management*, Vol. 2, pp.325-345, 2001.

Sakarri, T., Palachi, E., Cohen, C., Hatsutori, Y., Shimizu, J., & Migashita, H., "Model Based Control System Design Using SysML, Simulink and Computer Algebra System", *Journal of Control Science and Engineering*, Vol. 2013, p. 14, 2013.

Salas, E. & Cannon-Bowers, J., "The Science of training: A decade of Progress", *Annual Review of Psychology*, 52, pp.471-499, 2001.

Salas, E., & Dietz, A.S., "*Situational Awareness: Critical Essays on Human Factors in Aviation*", Ashgate Publishing, September, 2011, ISBN: 9780754629733.

Salas, E. & Milham, L.M. "Training Evaluation in the Military: Misconceptions, opportunities, and Challenges", *Military Psychology*, 15(1), pp.3 -16, 2003.

Salmon, M., Stanton, N.A., & Walker, G.H., "Situation Awareness Measurement: A Review of Applicability for C4i Environments", *Applied Ergonomics*, 37(2), April, 2006.

Salud, E., "*Developing a Library of Display Effects On Pilot Performance: Methods, Meta-analyses, and Performance Estimates*", Master's Theses and Graduate Research, SJSU Scholar Works, San Jose State University, Paper 4363, 2013.

Sargent, R.G., "Verification and validation of simulation models", *Proceedings of the 30th winter simulation conference*, Los Alamitos: IEEE Computer Society Press, pp. 121-130, 2004.

Schank, J.F., Thie, H.J., Graf, C.M., Beel, J., & Sollinger, J., "Finding the Right Balance", National Research Institute, RAND, 2002, ISBN: 0-8330-3104-X.

Schmid, K., "A Formal Approach to Technical Debt Decision Making", *QoSA 13th Proceedings of the 9th International ACM SIGSOFT Conference on Quality of Software Architectures*, pp. 153-162, NY: USA, 2013.

Schurr, A., & Selic, B., "Model Driven Engineering Languages and Systems", *12th International Conference (MODELS, 2009)*, Denver, CO, October, 2009.

Schutte, P., Goodrich, K., & Williams, R., "Towards an Improved Pilot-Vehicle Interface for Highly Automated Aircraft: Evaluation of the Haptic Flight Control System", *4th AHFE International Conference on Applied Human Factors and Ergonomics*, 20-25 July, San Francisco: CA, 2012.

Scott, S.G., "Factors Impacting the Selection of Training Delivery Systems and Training Methodology of Virginia Training Professionals", *Virginia Polytechnic Institute and State University*, Doctor of Philosophy Dissertation, December, 2010.

Scott, D., Griendling, K., & Mavris, D., "An Assessment of ROSETTA for Smart Electricity Grid System-of-Systems Design", *6th Annual System of System Engineering Conference*, Albuquerque, New Mexico, June, 2011.

SDI, "The Global Military Simulation and Virtual Training Markets 2015-2025", *Strategic Defence Intelligence Report*, p. 187, March, 2015.

Senior, B., "Organisational Change (2nd edition)", London: Prentice Hall, 2002.

SEP, "Tarski's Truth Definition", *Stanford Encyclopedia of Philosophy*, 15 August, 2014, <http://plato.stanford.edu/entries/tarski-truth/>, [Accessed: 7/10/2014].

Shaw, N.D., Mian, A., Yadav, S.B., "A comprehensive agent-based architecture for intelligent information retrieval in a distributed heterogeneous environment", *Decision Support Systems*, Vol. 32, pp. 401-415, 2002.

Shove, E., "Submission to the House of Lords Science and Technology Select Committee call for evidence on behaviour change", <http://www.lancs.ac.uk/staff/shove/transitionsinpractice/papers/shoveHouseofLords.pdf>, 2010a. [Accessed: 20/10/2011].

Shy, K.S., Hageman, J.J., & Le, J.H., "The Role of Aircraft Simulation in Improving Flight Safety Through Control Training", *National Aeronautics and Space Administration, NASA/TM-2002-210731*, August, 2002.

Siemieniuch, C.E., and Sinclair, M.A., "On complexity, process ownership and organisational learning in manufacturing organisations, from an ergonomics perspective", *paper submitted for Special issue of IJHFEM*, April, 2000.

Silmon, J., Evans, R., Brownsword, N., & Nicholson, D., “An Architectural Approach to Requirements Engineering for Infrastructure Projects”, *Systems Engineering Journal*, Vol. 18, Issue, 3, May, 2015.

Simperl, E., “Reusing Ontologies on the Semantic Web: a Feasibility Study”, *Data and Knowledge Engineering*, 68(10), pp. 905-925, 2009.

Skaff, M., “F-35 :Lightening II Cockpit Vision”, SAE International, 2010-01-2530, October, 2010.

Smaili, H., Breeman, J., Lombaerts, T., Stroosma, O., “A Benchmark for Fault Tolerant Flight Control Evaluation”, *Progress in Flight Dynamics, GNC, and Avionics*, Vol. 6, pp. 333-346, 2013.

Smit, M.C., “A Cost Analysis of Military Training”, NATO Science and Technology Organisation, *RTO System Analysis and Studies Panel (SAS) Workshop*, Amsterdam, Netherlands, 5-6 June, 2012.

Smith, E., McIntyre, H., “*Simulation and collective training – creating a total experience*”, The Royal Aeronautics Society, Flight Training DERA/AS/FMC/CP010229, June, 2000.

Smith, M.K., “*Howard Gardner and Multiple Intelligences*”, The Encyclopaedia of Informal Education, <http://www.infed.org/thinkers/gardner.htm>. [Retrieved: 20/04/2015]

Smith, R., “Technology Disruption in the Simulation Industry”, *Journal of Defense Modeling and Simulation: Applications, Methodology, Technology*, Vol. 3, No. 1, pp. 3-10, January, 2006.

Soley, R.M., “*Leveraging Model Driven Architecture for Model Driven Engineering*”, Object Management Group ©, March 2004.

Sony, J., Zhuang, D., Sang, G., Wanyan, X., “Pilot Mental Workload Measurement and Evaluation Under Dual Task”, *4th Annual Conference on Biomedical Engineering and Informatics (BMEI)*, Vol. 2, pp. 809-812, 15-17 October, 2011.

Sottolare, B., “The Human Interface to Embedded Virtual Simulation”, Improving Human Effectiveness Through Embedded Virtual Simulations, Findings of Task Group HFM-165, 30 October, 2004.

Sparks, A.L., Burki-Cohen, J., Go, T.H., “Transfer of Training from a Full-flight Simulator vs. A High Level Flight Training Device with a Dynamic Seat”, *American Institute of Aeronautics and Astronautics, AIAA Modeling and Simulation Technologies Conference*, 2-5 August, Canada, 2010.

Staal, M.A., “*Stress, Cognition, and Human Performance: A Literature Review and Conceptual Framework*”, NASA/TM-2004-212824, AMES Research Center, August, 2004.

Stanton, N.A. & Young, M., “The application of ergonomics methods by novices”, *Applied Ergonomics*, 34, pp. 479-490, 2003.

Stewart, J.E. & Dohme, J.A., “Automated hover trainer: Simulator based intelligent flight training system”, *International Journal of Applied Aviation Studies*, No.5, pp. 25-40, 2002.

Stewart, J.E., Johnson, D.M., & Howse, W.R., “Fidelity Requirements for Army Aviation Training Devices: Issues and Answers”, U.S. Army Research Institute for the Behavioural and Social Sciences, Research Report 1887, April, 2008.

Stevens, J.A., & Kincaid, P., “The Relationship between Presence and Performance in Virtual Simulation Training”, *Open Journal of Modeling and Simulation*, 3, pp. 41-48, 2015.

Stevenson, J. & Dunn, R., “Knowledge Management and Learning Styles: Prescriptions for future teachers”, *College Student Journal*, 35(4), pp.483-490, 2001.

Stevens, J.A., & Kincaid, P., “The Relationship between Presence and Performance in Virtual Simulation Training”, *Open Journal of Modelling and Simulation*, 3, pp. 41-48, 2015.

Stevens, J., Ortiz, E., Reinermann-Jones, L., & Maxwell, D., “Approach to Examine Efficacy of Game-based and Virtual Simulation Training”, *SummerSim '15 Conference on Summer Computer Simulation*, pp. 1-7, San Diego, CA, USA, 2015.

Stimpson, A.J., Buinhog, L.S., Bezek, S., Boussemont, Y., Cummings, M.L., “A Model-based Measure to Assess Operator Adherence to Procedure”, *Proceedings of HFES 2012*, 56th Annual meeting of the Human Factors and Ergonomics Society, Boston, M.A., October, 2012.

Stotz, R.W., “System Definition: Defining the Intended Use for a System”, *Journal of Validation Technology*, pp. 54-67, 2007.

Strachan, I.W., “Visual and Motion Cues In the Real World and In Simulators”, *Visual and Motion Cues*, ETSA Newsletter, September, 2014

Strybel, T., Minakata, K., & Kraft, J., “Assessing the Situation Awareness of Pilots Engaged in Self Spacing”, *Human Factors and Ergonomics Society Annual Meeting*, 52(1), September, 2008.

Suh, N.P., “Axiomatic Design Theory for Systems”, *Research in Engineering Design*, Vol. 10, pp. 189-209, Springer-Verlag, London, 1998.

Sujda, R., “Empirical Evaluation of Decision Support Systems”, *Environmental Modelling and Software*, 20(2), pp. 269-277, 2006.

Summers, J.E., “Simulation-based Military training: An Engineering Approach to Better Addressing Competing Environmental, Fiscal, and Security Concerns”, Washington Academy of Sciences, Spring, 2012.

Sumwalt, R.L., Thomas, R.J., Dismukes, K., “Enhancing Flight-Crew Monitoring Skills Can Increase Flight Safety”, 55th international Air Safety Seminar, Flight Safety Foundation, 4-7 November, 2002.

Suskie, L., “*Theories and instruments for Identifying Student Learning Styles*”, National Resources Canter, 2011.

Symons, S., France, M., Bell, J., Bennet, W., “Linking knowledge and skills to mission essential competency-based syllabus development for distributed mission operations”, 25th *Interservice/Industry Training Simulation & Education Conference*, Orlando, FL, 2003.

Teantzer, G., & Bordeleau, F., “Modeling Foundations and Applications”, 11th *European Conference (ECMFA 2015)*, Italy, 20-24 July, 2015.

Tamarskar, S., Neema, K., & DeLaurentis, D., “*Framework for Measuring Complexity of Aerospace Systems*”, Research in Engineering Design, February, 2014.

Tariq, A., & Rofi, K., “Intelligent Decision Support System – A Framework”, *Information and knowledge Management*, Vol. 2, No. 6, 2012.

Tashakorri, A. & Teddie, C., “*Validity issues in mixed methods research: Calling for an integrative framework*”, American Education Research Association, San Francisco, CA, 2006.

Taylor, H.L., Talleur, D.A., Emanuel, T.W., Rantanen, E.M., “*Effectiveness of flight training devices used for instrument training*” (Final technical Report AHFD-05-9/FAA-05-4), August, 2005.

Teachout, M.S., Barto, E.L., Carreta, T.R., King, R.E., & Michaels, C.F., “Consistency of Pilot Trainee Cognitive Ability, Personality, and Training Performance in Undergraduate Pilot Training”, Airforce Research Laboratory, AFRL-RH-WP-TR-2013-0081, 2013.

Thatcher, S., “Scenario-based Learning and Assessment for Second Year Aviation Students”, *Global Journal of Engineering Education*, Vol. II, No. 2, 2007.

The Object Management Group (OMG), “*MDA Guide Revision 2.0*”, OMG Document ormsc/2014-06-01, 2014.

Thomas, R., & Lee, C.C., “Development of Training Scenarios in the Flight Training Device for Flight Course at Embry-Riddle Aeronautical University”, *Journal of Aviation/Aerospace Education & Research*, Vol. 24, No. 3, Spring, 2015.

Thunholm, P., “A new model for tactical mission planning for the Swedish Armed Forces”, *Command and Control Research and Technology*, San Diego, C.A, 2006

TNO/NLR, “*Programme of Requirements (PVE) for Mission Training Center Fighter Aircraft*”, The Netherlands: TNO/NLR, 2007.

Tokunaga, T., Fujimura, S., “Equivalence of Design Structure Matrix and Axiomatic Design”, *Proceedings of the 15th International DSM Conference*, Melbourne, Australia, 29-30th August, 2013.

Tsui, A.S., Nifadkar, S.S., Ou, A.Y., “Cross-National Cross-Cultural Organisational Behaviour Research: Advances, Gaps, and Recommendations”, *Journal of Management*, June, 2007.

Tucker, R., “Southern drift: The learning styles of first- and third-year students of the built environment”, *Architectural Science Review*, 50(3), pp. 246-255, 2007.

Turban, E., Aronson, J.E., & Liang, T.P., “*Decision Support Systems and Intelligent Systems*”, Prentice Hall, 2005.

Tubsree, C., & Tubsree, N.Y., “The Application of Instructional Design Principles in E-Learning Design and Development: The Case Study at the University of Trento, Italy”, *Proceedings Society for Information Technology & Teacher Education International Conference*, Austin, Texas, 5 march, 2012.

Unsworth, K.L., and West, M.A., “*Teams, the challenge for cooperative work*”, Introduction to work and organisational psychology: a European Perspective, Oxford: Blackwell, 2000.

Vafadar, M., “Virtual Reality: Opportunities and Challenges”, *International Journal of Modern Engineering Research (IJMER)*, Vol. 3, Issue 2, March-April, pp. 1139-1145, 2013.

Valbuena, G., “Decision Making Process – A Value-Risk Trade-off Practical Applications in the Oil and Gas Industry”, *Scientific & Academic Publishing*, 3(3), pp. 142-151, 2013.

Van Der Pal, J. & Ligthart, V.H.M., “*EuroTraining TNA training needs analysis: NLR’s Synthesised Contribution*”, NLR Report NLR-CR-2003-134, Amsterdam: NLR, 2003.

Van Der Pal, J., Fickweiler, A., Boot, E.W., & Pouw, F.C.M., “*Effective and Efficient Training Design Part 1: Training Design models in the Dutch Ministry of Defence*”, National Aerospace Laboratory, Amsterdam, 2010.

Velada, R., Caetano, A., Michel, J.W., Lyons, B.D., & Kavanaugh, M.J., “The Effects of Training Design, Individual Characteristics and Work Environment on Transfer of Training”, *International Journal of Training and Development*, 11, pp. 282-294, 2007.

Venkatesh, V., Brown, S.A., & Dyla, H., “Bridging the Qualitative-Quantitative Divide: Guidelines for Conducting Mixed Methods Research in Information Systems”, *MIS Quarterly*, Vol. 37, Issue 1, pp. 21-54, 2013.

Verbeek, J., “*A Systems Engineering Approach for Development of Aerospace Production Systems*”, SAE Technical Paper 2013-01-2162, 2013.

Vermeulen, M.I., Tromp, F., Zuithoff, N.P., Pieters, R.H., Damoiseaux, R.A. & Kuyvenhoven, M.M., “A Competency Based Selection Procedure for Dutch Postgraduate GP Training: A Pilot Study of Validity and Reliability”, *PubMed*, pp. 307-313, December, 2014.

Vincenty, T., “Direct and Inverse Solutions of Geodesics on the Ellipsoid with Application of Nested Equations”, *Survey Review*, Vol. 23, No. 176, pp. 88-93, April, 1975.

VRS, “Virtual Reality in the Military”, Virtual Reality Site, <http://www.vrs.org.uk/virtual-reality-military/>, [Accessed: 24/10/15].

Vogel-Heuser, B., Schutz, D., Frank, T., & Legat, C., “Model-Driven Engineering of Manufacturing Automation Software Projects – A SysML-based Approach”, *Model-Based mechatronics System Design 2: Model Based Engineering*, Vol. 24, Issue 7, pp. 883-897, October, 2014.

Wagenerf, F., Ison, D.C., “Crew Resource Management Application in Commercial Aviation”, *Journal of Aviation Technology and Engineering*, 3(2), pp. 2-13, 2014.

Wagener, F., & Ison, D.C., “Crew Resource Management Application in Commercial Aviation”, *Journal of Aviation Technology and Engineering*, 3(2), pp. 2-13, 2014.

Wagner, E., “Essay: In Search of the secret handshakes of ID”, *The Journal of Applied Instructional Design*, 1(1), pp. 33-37, April, 2011.

Wall, J.D., Stahl, B.C., & Salom, A.F., “Critical Discourse Analysis as a Review Methodology: An Empirical Example”, *Communication of the Association for Information Systems*, Vol. 37, Artidell, 2015.

Wang, Y., White, M., Owen, I., Hodge, S., Barakos, G., “Effects of Visual and Motion Cues in Flight Simulation of Ship-borne Helicopter Operations”, *CEAS Aeronautics Journal*, Vol. 4, Issue 4, pp. 385-396, December, 2013.

Ward, P., Williams, M., & Hancock, P.A., “*Simulation for Performance and Training*”, The Cambridge Handbook of Expertise and Expert Performance, Cambridge University Press, June, 2012, ISBN: 97805118416796.

West, D.M., “Using Technology to Personalize Learning and Assess Students in Real-Time”, Centre for Technology Innovation at Brookays, 6 October, 2011.

Wey, Y., White, M., Owen, I., Hodge, S., & Barakas, G., “Effects of Visual and Motion Cues in Flight Simulator of Ship-borne Helicopter Operations”, *CEAS Aeronautical Journal*, Vol. 4, Issue 4, pp. 385-395, December, 2013.

Wickens, C.D. & Holland, J.G., “*Engineering Psychology and human performance: Third edition*”, Upper Saddle River, NJ: Prentice Hall, 2012.

Wiggins, M., “*An Assessment of General Aviation Pilot Performance during Simulated Flight*”, ATSB Research and Analysis Report, Australian Transport Safety Bureau, April, 2006.

Wihl, L, “*Training for the Combined Cyber / Kinetic Battlefield*”, MODSIM World 2015, Paper No. 9, pp. 1-11, 2015.

Wilber, F.R., "A System of Systems approach to E-enabling the commercial airline applications from an air-framers perspective", Keynote Speech, *IEEE SoSE Conference*, USA, 2007.

Williams, G.W., Lawrence, K., Weeks, R., "Modelling and Simulation Technologies: Reconfigurable Flight Simulators in Modelling and Simulation", *In Proc. AIAA Modelling and Simulation Technologies Conference*, Reston, VA: American Institute of Aeronautics and Astronautics, 2004.

Wojciechowski, J., "Validation of a task network human performance model of driving", Master's Thesis, Virginia Polytechnic Institute and State University, 2006.

Woods, S., "Flight Crew Reliance on Automation", CAA Paper 2004/10, Civil Aviation Authority, December, 2004.

Xuelin, H., Jiangyun, W., Liang, H., Shidong, M., "Situation Awareness Enhancement of Cockpit Synthetic Vision System: Visualisation of Tunnel-in-the-sky", *9th International Conference on Electronic Measurement and Instrumentation – ICEMI*, 16-19 August, pp. 3-657-3-668, 2009.

Xui, W., "Learning Styles and Their Implications in Learning and Teaching", *Theory and Practice in Language Studies*, Vol. 1, No. 4, pp. 413-416, April, 2011.

Yanardoner, E., Kiziltepe, Z., Siggie, F.N., & Sekerler, S.A., "The Learning Styles and Personality Traits of undergraduates: A Case at a State University", *Anthropologist*, 18(2), pp. 591-600, 2014.

Yang, H., Sackett, P., Arvey, R., "Statistical power and cost in training evaluation: Some new considerations", *Personnel Psychology*, 49, pp.651-668, 1996.

Yang, C., Yin, T., Zhao, W., Huang, D., & Fu, S., "Human Factors Quantification via Boundary Identification of Flight Performance Margin", *Chinese Journal of Aeronautics*, Vol. 27, Issue 4, pp. 977-985, August, 2014.

Yeh, M., Yo, Y.J., Donovan, C., & Gabree, S., "Human Factors Considerations in the Design and Evaluation of Flight Deck Displays and Controls", *The National Transportation Systems Center*, Cambridge, MA: US DOT Volpe National Transportation Systems Centre, 2013.

Young, J.G., Trudea, M., POdell, D., Marinelli, K., Dennerlein, J.T., "Touch-screen tablet user configuration and case-supported tilt head and neck flexion angles", *Work*, 41(1), pp. 81-91, 2012.

Young, S.D., Adelston, B.D., & Ellis, S.R., "Demand Characteristics of a Questionnaire used to Assess Motion Sickness in a Virtual Environment", *Virtual Reality Conference*, 25-29 March, pp. 97-102, 2006

Zadeh, A., "Outline of a new approach to the analysis complex systems and decision processes", *IEEE Trans. Systems, Man and Cybernetics*, pp28-44, Jan, 1973.

Zanetti, M.S., "A *Complex Systems Approach to Software Engineering*", ETHZurich, Diss. ETH Nr. 21653, 2013.

Zarghooni, S., "Situation Awareness: A Scientifically Viable Construct?", Institute of Psychology, University of Oslo, Term Paper: PSY4401, December, 2009.

Zhang, L, Yuan, X, Wang, L, Dong, D, "Design and Implementation of Ergonomics Evaluation System of 3D Airplane Cockpit", *Digital Human Modelling*, Springer-Verlag, pp.1047-1055, 2007.

Zheng, Y., & Tangwen, Y., "Using NASA-TLX to Evaluate the Flight Desk Design in Design Phase of Aircraft", School of Aeronautics and Astronautics, Shanghai, Jiao-Tong University Proceda Engineering, Vol. 17, pp. 77-83, 2011.

Zoughbi, G, Briand, L., Labiche, Y., "Modeling safety and airworthiness (RTCA DO-178b) information: Conceptual model and UML profile", *Software and Systems Modeling*, pp. 1-31, 2010.

APPENDIX.....SEE VOLUME II