



Takashi Shibui

Value-Based Planning for Capability Evolution

SCHOOL OF AEROSPACE, TRANSPORT AND
MANUFACTURING
PhD in Aerospace

PhD
Academic Year: 2023 - 2024

Supervisor: Prof. Marin D. Guenov
Associate Supervisor: Dr. Atif Riaz
December 2023



SCHOOL OF AEROSPACE, TRANSPORT AND
MANUFACTURING
PhD in Aerospace

PhD

Academic Year 2023 - 2024

TAKASHI SHIBUI

Value-Based Planning for Capability Evolution

Supervisor: Prof. Marin D. Guenov
Associate Supervisor: Atif Riaz
December 2023

© Cranfield University 2023. All rights reserved. No part of this
publication may be reproduced without the written permission of the
copyright owner.

ABSTRACT

Systems such as military equipment can be viewed as capability components constituting capability systems alongside others (e.g., interoperating systems, personnel, etc.). Capability systems are developed to perform the tasks of their acquirer organisation and realise the capabilities identified in the organisation's capability planning. Also, they are continuously engineered for the capabilities' evolution.

In this time-dependent context, individual systems need to sustain their value to the acquirer. Their design ought to be better aligned with the overarching capability development. Two challenges were identified regarding this goal:

1. Evaluation of design alternatives based on their contribution to the organisational capability, corresponding to the top-level organisational task. It should account for the organisation's priorities concerning different capability areas.
2. Exploration of design evolution alternatives considering the long planning horizons, expecting future changes and uncertainties regarding the capability needs and the system's context (e.g., adversaries, environment, etc.).

To address them, a novel framework was developed. It incorporates:

- Value-based evaluation of design alternatives.

It is founded on a value metric formally defined to represent the contribution to the organisational capability. Two methods were developed to estimate the "value": The first method determines task importance reflecting the acquirer's needs. The second method estimates the achieved capability levels by identifying the trade-off relationships of factors regarding the system design and context.

- Scenario planning method.

It allows to model different futures using codified change trajectories of needs and contextual factors. By calculating value trajectories of the evolving system alternatives for the scenarios, what-if analysis to aid decisions is made possible.

- A consistent domain ontology underpinning the framework.

It defines and relates concepts from relevant industrial domains.

The proposed framework and methods were tested and evaluated with representative test cases. Evaluation by industrial practitioners was also performed. The results showed the framework's usefulness and relevance to real-world practices.

Keywords:

Capability-Based Planning, Capability Systems Engineering, Value-Centric Design Methodologies, Decision Analysis, Structural Equation Modelling

ACKNOWLEDGEMENTS

Looking back, this PhD has certainly been a great challenge for me in many respects. Returning to academia after years at work, bringing family to a different environment, with a different language, and the pandemic on top of that... In the face of them, the completion of this research was unthinkable without the contribution of numerous people who have helped, supported, and motivated me.

First and foremost, I would like to express my deepest gratitude to my supervisors, Professor Marin D. Guenov and Dr. Atif Riaz, who guided me through this journey with wisdom, thoughtfulness, patience, and patience. I am also thankful to brilliant colleagues in the Advanced Engineering Design Group, Dr. Xin Chen, Dr. Sergio Jimeno Altelarrea, and Dr. Soufianne El-Fassi. It was my privilege that I had somebody who generously shared the first-hand experiences of PhD and academic research. I would like to thank Dr. Yifan Zhao and Dr. Stevan van Heerden, for the valuable expert advice in the midst of the project. Very special thanks go to the experts at BAE Systems, who kindly devoted time to provide insightful feedback, which greatly helped me to improve my work.

I have been truly blessed to be a part of Cranfield University and I would like to extend my appreciation for those wonderful people. In particular, I am thankful to Ms. Joanne Holden and Mr. Abdullaah Ali for the help with academic English skills. Numerous friends, from every part of the world, thank you for making my student life so rich. Additionally, on behalf of my family, I would like to appreciate your warm kindness to my loved ones.

I am indebted to leaders and colleagues in my home organisation, for everything that made this journey possible. It has indeed been my privilege.

Last not least, I would like to explicitly express my gratitude to my family, who have always supported me with tremendous consideration and unconditional love. Although we have enjoyed, those three years have undoubtedly been no less challenging to you than to me. I cannot be prouder of you that you have gone through. *Otsukaresama, Arigato!*

TABLE OF CONTENTS

ABSTRACT	i
ACKNOWLEDGEMENTS.....	iii
LIST OF FIGURES	vii
LIST OF TABLES	ix
LIST OF ABBREVIATIONS	xi
SYMBOLS	xiii
VOCABULARY	xv
1 Introduction.....	1
1.1 Background.....	1
1.1.1 Systems and Surrounding Changes.....	1
1.1.2 Evolution of System Design	2
1.1.3 Systems in Capability Development Context.....	2
1.1.4 Evolving System as Capability Component.....	5
1.2 Motivation, Aim and Objective	7
1.3 Research Methodology	10
1.4 Thesis Structure.....	12
2 Literature Review	15
2.1 Value-Centric Design Methodologies.....	15
2.2 Use of Effectiveness in Design	21
2.3 Design Exploration in Dynamic Context.....	22
2.4 Capability as a subject of systems engineering	24
2.4.1 Concept and Perspectives of Capability.....	24
2.4.2 Capability Development Activity	27
2.4.3 Capability Systems and Systems-of-Systems Engineering	31
2.4.4 Analytical Methodologies for Mission Execution.....	33
2.4.5 Evolutionary Development	36
2.5 Operation Modelling.....	37
2.6 Critical Discussion.....	39
2.7 Research Scope	41
2.8 Summary and Conclusions	43
3 Framework Definition	45
3.1 Capability Planning Ontology.....	45
3.2 Framework Formalisation	50
3.2.1 System Boundaries Definition	50
3.2.2 Capability Levels	54
3.2.3 Contribution of Sol to Organisational Capability	58
3.2.4 Value Definition Based on Contribution.....	60
3.2.5 Value-Based Evaluation among Temporal Changes.....	62
3.3 Summary and Conclusions	68
4 Capability Needs Prioritisation.....	71

4.1 Introduction	71
4.2 Pairwise Comparisons within Value Hierarchy.....	72
4.3 Demonstration Test.....	76
4.4 Conclusions	80
5 Mission Effectiveness Estimation	81
5.1 Introduction	81
5.2 Structural Equation Modelling	82
5.2.1 Model Specification without Latent Variables	83
5.2.2 Model Specification with Latent Variables	85
5.2.3 Model Estimation.....	89
5.3 Demonstration Test.....	91
5.4 Conclusions	95
6 Evaluation.....	97
6.1 Introduction	97
6.2 Software Implementation	98
6.3 Case Study	102
6.4 Industrial Evaluation.....	110
6.4.1 Questionnaire	112
6.4.2 Feedback	113
6.5 Conclusions	116
7 Summary and Conclusions.....	119
7.1 Research Summary	119
7.1.1 Objective 1	120
7.1.2 Objective 2	123
7.1.3 Objective 3	125
7.1.4 Evaluation	126
7.2 Research Findings.....	127
7.3 Research Contributions	131
7.4 Future Work	133
REFERENCES.....	135
APPENDICES	145

LIST OF FIGURES

Figure 1 Design Loop	17
Figure 2 A Taxonomy of Decision Analysis Practice	19
Figure 3 Design loop under dynamic system perspective.	23
Figure 4 Generic Process Chart of Capability-Based Planning Error! Bookmark not defined.	
Figure 5 Systems of Systems Domains in Defence.....	33
Figure 6 Mission-System Analysis.....	34
Figure 7 SoS Systems Engineering Wave Model.....	37
Figure 8 Hierarchy of Models Used within U.S. Department of Defense	38
Figure 9 Ontology Model as a UML Class Diagram	49
Figure 10 Boundaries of Sol and Capability Systems	52
Figure 11 System Solution Contexts	53
Figure 12 Breakdown of Organisational Tasks into Missions	58
Figure 13 Calculating Relative Contribution to Organisational Capability	59
Figure 14 Formation of Sol Value Tree	61
Figure 15 Characterisation of Capability System, Sol, Context and Operational Environment as a UML Class Diagram.....	63
Figure 16 Temporal Representation of the Proposed Framework.....	66
Figure 17 Scenario Characterisation as a UML Class Diagram	67
Figure 18 Task Hierarchy and Pairwise Comparisons.....	77
Figure 19 Icicle Chart of Value Hierarchy	78
Figure 20 Icicle Chart of Value Hierarchy (Changed)	79
Figure 21 Stacked Bar Chart of Mission Importance Weights	79
Figure 22 Hypothetical Path Diagram with All Direct Paths.....	84
Figure 23 Three Types of Latent Variable Construct.....	86
Figure 24 Example of Proposed Hypothetical Model.....	87
Figure 25 Fitted Model of a DCA Mission.....	94
Figure 26 Scenario Planner Page of Demonstration Software	101
Figure 27 Design Browser Page of Demonstration Software	102

Figure 28 Task Hierarchy and Pairwise Comparisons of a Use Case 104

Figure 29 Needs and Context as of epoch 0 in “Pessimistic” Scenario 108

Figure 30 Needs and Context as of epoch 1 in “Pessimistic” Scenario 108

Figure 31 Needs and Context as of epoch 0 in “Optimistic” Scenario 108

Figure 32 Needs and Context as of epoch 1 in “Optimistic” Scenario 108

Figure 33 Visualisation of Value Trajectories 109

LIST OF TABLES

Table 1 Fundamental Scale..... 73

Table 2 Evaluators’ Levels of Agreement Regarding Framework’s Underlying Assumptions 114

Table 3 Evaluators’ Levels of Agreement Regarding Individual Methods’ Usefulness..... 114

Table 4 Evaluators’ Levels of Agreement Regarding Usefulness and Validity of Overall Framework 115

LIST OF ABBREVIATIONS

AAM	Air-to-Air Missile
AFT	Alternative-Focused Thinking
AHP	Analytic Hierarchy Process
AI	Artificial Intelligence
CADMID	Concept, Assessment, Demonstration, Manufacture, In-Service, Disposal
CAS	Close Air Support
CBP	Capability-Based Planning
DCA	Defensive Counter-Air
DfE	Design for Evolvability
DLoD	Defence Lines of Development
DoE	Design of Experiments
DOTMLPF	Doctrine, Organization, Training, Materiel and equipment, Leadership and education, Personnel, Facilities
DRM	Design Research Methodology
DS	Descriptive Study
DSE	Design Space Exploration
EBD	Effectiveness-Based Design
EEA	Epoch-Era Analysis
EWR	Early Warning Radar
GUI	Graphical User Interface
IEC	International Electrotechnical Commission
IEEE	Institute of Electrical and Electronics Engineers
ISO	International Organization for Standardization
JCA	Joint Capability Areas
JCIDS	Joint Capabilities Integration and Development System
M&S	Modelling and Simulation
MATE	Multi-Attribute Tradespace Exploration
MAU	Multi-Attribute Utility
MAUT	Multi-Attribute Utility Theory
MCDM	Multi-Criteria Decision Making
MDO	Multidisciplinary Design Optimisation

ME	Mission Engineering
MET	Mission Engineering Thread
MIM	Mission Integration Management
MIMIC	Multiple Indicator, Multiple Cause
MoE	Measure of Effectiveness
MoP	Measure of Performance
MT	Mission Thread
PS	Prescriptive Study
RC	Research Clarification
SAM	Surface-to-Air Missile
SEM	Structural Equation Modelling
SoI	System-of-Interest
SoS	System-of-Systems
SWaP	Size, Weight, and Power
TEPID OIL	Training, Equipment, Personnel, Information, Concepts and doctrine, Organisation, Infrastructure, Logistics
UK	United Kingdom
UJTL	Universal Joint Task List
US	United States
VCDM	Value-Centric Design Methodology
VDD	Value-Driven Design
VFT	Value-Focused Thinking

SYMBOLS

A	Pairwise Comparison Matrix
C	Context Vector
D	Design Vector
E	MoE Vector
P	MoP Vector
V	Sol Value
V_o	Organisational Capability Level
V_i	Capability Level of Mission i (in Organisational Task Spectrum)
v_i	Capability Level of Mission i (in Sol Task Spectrum)
W_i	Importance Weight of Mission i (in Organisational Task Spectrum)
w	Needs Vector
w_i	Importance Weight of Mission i (in Sol Task Spectrum)

VOCABULARY

Capability	The ability to complete a task under specified conditions and level of performance (Adopted from (OUSD(R&E) USDoD, 2020))
Capability Component	A system required to build and realize a capability (Henshaw et al., 2011)
Capability Level	The level of performance achieved when a capability system performs a task, as assessed by the acquirer
Capability System	A system which performs a task
Context (of a system)	The environment in which the system operates (Ricci, 2014, p. 78)
Context Variable	A variable which differentiates one context of Sol from another
Design Variable	A variable representing a decision which differentiates one design of Sol from another
Effect	A set of changes in the world which is created as the consequence of an activity
Epoch	A discrete time period with fixed context and needs (Ricci, 2014, p. 81)
Needs (of a system)	What is desired of the system (Ricci, 2014, p. 78)
Needs Variable	A variable which differentiates one needs of Sol from another
Operational Environment (of a Capability System)	Entities outside the system boundary of a capability system
Organisational Task	A task to which an organisation is assigned
Scenario	The expression of one possible future regarding factors exogenous to the Sol
(Design, Context, Needs) Space	The space spanned by a finite set of (design, context, needs) variables
System-of-Interest (Sol)	System whose life cycle is under consideration (ISO/IEC/IEEE, 2023)
Task	A clearly defined action or activity (Adopted from (United States Department of Defense, 2013), as cited in (OUSD(R&E) USDoD, 2020))
Task Priority	The prioritisation of one task over another when capability systems are considered for multiple tasks

Value (of a system)	The metric which represents the desirability of the system, defined to reflect the preferences of a decision maker
(Design, Context, Needs) Vector	A vector representing an instance of (design, context, needs), formed by a finite set of (design, context, needs) variables

1 Introduction

This chapter begins with a background (Section 1.1), resulting from initial research. Then, the motivation behind the research is stated, which leads to the definition of the general aim and objective of the work (Section 1.2). This is followed by the description of the research methodology (Section 1.3). Finally, the layout of the thesis is outlined (Section 1.4).

1.1 Background

1.1.1 Systems and Surrounding Changes

Engineering systems are developed to provide benefit to their stakeholders. The benefit may vary depending on the different stakeholders. The benefit for the supplier may be the profit made through selling the system. For the acquirer, it may be the profit made through operating the system, if the system is meant for a commercial purpose. In defence domain, which is fundamentally not intended for profit, the benefit to the acquirer is usually not straightforward to establish, as it will be discussed later in the thesis. In any case, the desirability of a system is usually driven both by cost and benefit (e.g., cost-to-benefit ratio). The stakeholders need to take them into consideration, and eventually, make decisions.

Systems are usually expected to operate for a long operational time. The benefit needs to be sustained, as well as the cost needs to be kept low, for a better cost-to-benefit throughout the system's lifecycle. However, both benefit and cost are influenced by the circumstances surrounding the system, which are often volatile. For example, in civil aviation, airliners recently suffered from sheer drops in travel demand because of random events such as terrorism, economic recession, and pandemics. As the consequence of such changes, the cost-to-benefit may worsen and in extreme cases, the system operation may be terminated earlier than initially expected. While adverse circumstances are generally difficult to predict, it is desirable to account for possible changes and uncertainties and to consider their impact on the system explicitly in the development process.

1.1.2 Evolution of System Design

Design evolution is one possible way to mitigate the negative impact of the exogenous changes and to keep the systems beneficial to the stakeholders for a longer period. This approach leverages partial re-designs retaining some parts of the original design. In comparison to clean-sheet design, there is usually a limit in achievable performance increment because of the constraint imposed by the legacy part. On the other hand, the overall cost can be lower because of the commonality with the original design. An important consideration is that the original system needs to have a “room” for accommodating the evolution. This comes in various forms depending on the nature of evolution to be applied. For example, integration of additional equipment requires margins of Size, Weight, and Power (SWaP), and suitable interface, which provides modularity. The enabling quality of system design is referred to as evolvability, which is defined as “the ability of a system whose design to be inherited and changed across generations (over time)” (O. L. de Weck et al., 2012). Because of the benefit of design evolution, it was proposed to take it into account in the initial design of a system lifecycle. This principle can be called Design for Evolvability (DfE) (Juzgado & Bonnema, 2008).

1.1.3 Systems in Capability Development Context

The aspiration of this research is to improve the method for carrying out design evolution planning of systems which are developed and operated within the context of *capability development*. While the concept of capability is prevalent in some domains such as defence (and potentially applicable to other domains), it has not necessarily reached the state of common understanding in the wider engineering community.

Organisations such as the nation’s armed forces are built around the tasks for which they ought to be prepared and ready to perform. The fundamental *organisational task* can be as broad as “national defence”, and the entire (defence) organisation exists in order to undertake it. High-level tasks are broken down into smaller pieces with accordingly clarified details. As the result, so-called

missions, such as “air defence” and “anti-ship strike”, are derived as the “building blocks” of the organisational task.

What the organisation needs are *capabilities* to perform these tasks and missions. It obtains capabilities through acquiring *solutions* which realise the needed capabilities. Since their concern is capabilities as the properties of solutions rather than solutions themselves, they are generally open to diverse solution alternatives as long as they have sufficient capabilities. The development activity of these solutions is referred to as CBP (Capability-Based Planning).

CBP starts from needs, which are expressed as capabilities defined in a top-down, solution-independent manner. If properly implemented, CBP potentially offers three advantages over the traditional approach, where requirements are generated with more specific solution in mind and less consideration into the broader (e.g., organisation-wide) capability:

- First, CBP facilitates the exploration of a wide range of solution alternatives involving broader organisation. For example, when a defence ministry needs to fill the gap created by the retirement of existing fighter aircraft, if they simply start developing another fighter aircraft as the replacement with little exploration of other options, it is the traditional approach. By contrast, in CBP, they first need to consider what needs to be done in the future and identify capability gaps. In such a case, the scope of exploration can be much broader than just a fighter aircraft. It can possibly include ground-based surface-to-air missile systems and combat ships, or just re-alignment of non-materiel, such as personnel, and so forth.
- Second, CBP stimulates needs consideration from a broad perspective because of the top-down needs definition. In the previous example, fighter-for-fighter approach may be a local optimal solution for the air force, but it may be globally suboptimal for the entire armed forces, including army and navy, who are responsible for a wider spectrum of missions. This problem is commonly referred to as “organisational stovepipes”, which is considered to be a source of inefficiency (Joint Systems Analysis Group Technical Panel 3, n.d.).

- Third, CBP advocates focusing on the outcome of missions, rather than specific (system) performance. That is, from the individual systems perspective, CBP emphasises the need for validating systems in accordance with what needs to be done, rather than meeting requirements, defined using detailed performance parameters. From defence planning perspective, the emphasis is on the specification of tasks, without excessively presupposing particular situations (e.g., particular adversaries).

Because of these potential benefits, CBP has become the internationally preferred approach for defence planning (Taliaferro et al., 2019). Therefore, it is advocated here that the design (including evolution) of a particular System-of-Interest (Sol) needs to be done in the context of CBP. In such a case, the benefit which the acquirer expects can be assumed to be capability realisation. While it does not necessarily generate monetary return by itself, it inevitably comes with a certain cost. Therefore, the acquirer is likely to consider how the Sol will contribute to the organisational capability realisation, in comparison with the associated costs.

For individual missions, the acquirer's concern is how well missions are performed and how well capabilities are realised by corresponding solutions. Such solutions are often considered to be systems in their own rights, termed *capability systems*, because of the existence of multiple interacting constituent *capability components* within them and interactions among them. Such components include people, processes, equipment, and organisations, which are commonly abbreviated by defence organisations as TEPID OIL (Training, Equipment, Personnel, Information, Concepts and doctrine, Organisation, Infrastructure, Logistics) or DOTMLPF (Doctrine, Organization, Training, Materiel and equipment, Leadership and education, Personnel, Facilities). For each mission, these materiel and non-materiel capability components are integrated together to realise a capability.

1.1.4 Evolving System as Capability Component

Evolutionary development is so prevalent in capability systems (which are *System-of-Systems (SoS)* by definition) that it is considered as one of their key characteristics. Once they are formed, it is common to continually evolve without being replaced entirely. Evolution of capability systems is realised by evolution or replacement of *capability components*. However, another key characteristic of capability systems is that components have their own lifecycles, which may be managed independently from the capability system and other capability components. Thus, there is a possibility that some changes of these components may act negatively on the performance of the capability systems (e.g., retirement with no replacement). Therefore, capability components whose lifecycles are beyond control can be one of the sources of future uncertainties, alongside changing needs, adversary systems' evolution, and so forth.

From the viewpoint of each capability component, the possibility of design evolution is crucial, as it has a significant impact on the long-term benefit it can provide. A capability component creates this benefit by contributing to capability realisation in a mission as a part of their corresponding capability system. Multi-role systems, such as modern multi-role fighter aircraft, are involved in multiple capability systems, thus their value comes from multiple sources.

An example which illustrates the nature of evolving systems is the history of the product family of the F-15 fighter aircraft in the U.S. Air Force. The initial F-15A (operational from 1975) (The United States Air Force, 2019a) was designed as an air superiority fighter which can outperform its competitors in air-to-air combat missions. Later, it evolved into F-15E Strike Eagle (operational from 1989) (The United States Air Force, 2019b), which is a fighter-bomber, capable of performing air-to-ground missions as well as air-to-air. It leveraged a large part of the existing F-15 airframe, rather than designing a new system from scratch. That was supposed to lower the development and production cost. Additionally, because of the commonality between the two types, expected impact on logistics can be also lower. Subsequently, the end of the cold war made a significant change in the capability needs. At that time, F-22 (operational from 2005)(The United States

Air Force, 2022), which is the follow-on aircraft to supersede F-15, was already “on the horizon”. However, the development and production of F-22 were adversely affected by the shift of needs, making the reliance on F-15 in air-to-air missions largely unchanged for longer than initially planned. This was despite the superiority of F-22 over F-15 in air-to-air combat performance. At a later time, reflecting the then change of strategic environment, air-to-air missions were once again prioritised. It gave rise to the need for a fighter which complements the newer F-22 and F-35 fighters in air-to-air missions, and is still able to conduct air-to-ground missions. As a result, the procurement of a further modified F-15 variant, F-15EX Eagle II (not yet operational as of September 2023), was initiated in 2019. This indicated that F-15EX with modern avionics and weapons was expected to be sufficiently effective in future environment and thus beneficial to the organisational capability, despite the lack of many features found in fighters of newer generation (e.g., F-22 and F-35), such as low-observability. It was also envisioned that F-15EX would fill the niche of ground / maritime attack role with hypersonic missiles (The United States Air Force, 2021), which are not suitable for F-22 and F-35.

Overall, F-15 family can be considered as a successful example in maintaining the value of the system, though its successful evolution was largely a fortunate consequence of changing needs, triggered by strategic-level changes, which acted in favour of it. While its main source of value was initially air-to-air missions, later it was diversified into air-to-ground missions, which were relatively important in the post-Cold War years. The multi-role features enabled contribution to the organisational task (national defence) in multiple ways. This is in sharp contrast to F-22, which is highly optimised for high-end air-to-air missions. These, however, turned out to be of lower priority when the type became operational. Also, intended or not initially, the original F-15 design had sufficient “room” for evolvability to enable transition into F-15E and F-15EX.

As seen in this example, the realisation of a higher capability level is certainly important for the capability acquirer. Also very important is the contribution to a wider range of missions, especially those which are seen as a priority in the

specific timeframe. Mission priorities are determined by the acquirer's needs, which (as seen from the example) may change during the system lifecycle. This emphasises the need for an alignment of individual capability component development to the larger CBP activity, where capability needs are determined, and missions are prioritised.

1.2 Motivation, Aim and Objective

Evolution planning in the context of CBP can be seen as a decision-making process aiming to choose between alternatives. For systematic and traceable decision making, it is crucial that consistent supporting information is available to the decision makers. Therefore, this is one of the motivations behind the research presented in this thesis.

Methods broadly termed Design Space Exploration (DSE) are well-established means for information generation particularly for the early concept design phase of system design. A design space is first defined to characterise different alternatives of the System-of-Interest (Sol), then the design space is populated with design points, each of which represent a design alternative. These alternatives are systematically evaluated, allowing "exploration" of the design space for preferable alternatives. The evaluation metric can be defined arbitrarily according to the purpose of the analysis. If it is defined so that the preference of a decision maker is appropriately reflected, it is expected to directly aid their decision making.

Recently, DSE was extended to enable considering multiple timeframes, thus accounting for temporal changes (Ross, 2006; Ross & Hastings, 2006). These include both endogenous and exogenous changes to the Sol in question. The former refers to the changes to design (i.e., evolution of design), while the latter refers to those regarding the situation (environment, circumstances, etc.). One widely used generic framework for the DSE's expansion is Epoch-Era Analysis (EEA) (Ross & Rhodes, 2008), which has been applied for investigation into change-related properties ("ilities") of systems such as changeability, evolvability, and so forth. However, its application to the "capabilities-based" context has not

been sufficiently developed in existing works, in that some important characteristics of capability systems and capability components are not reflected.

Specifically, the scope of the methods is confined to single mission and arbitrarily selected metrics of mission outcome, which are commonly termed, Measures of Effectiveness (MoE). For example, with existing methods, one can investigate the impact of radar range on the number of intercepted enemy missiles in an air defence mission. This is supposed to be useful as a means of investigating the relationships between system design alternatives and the outcome within a single mission. The approach is partly in line with the capability-based concept in that it could account for:

- High-level metric of mission outcome, rather than low-level (system) performance, which can be indicators of capability.
- Integration of different capability components (as long as they can be characterised with a design space, possibly a high-dimensional one).

However, it must be noted that each mission metric represents only one aspect of a capability to perform a mission. There should be an even higher-level aggregation structure of multiple mission metrics into a single indicator metric of capability. It is likely to be determined subjectively according to the acquirer's intention.

Moreover, the aforementioned lack of consideration above the MoE-level leads to a potential overlook of the multiple-to-multiple relationships between capability systems and capability components. As described above, many capability components are multi-role, involving multiple capability systems as their constituents. This indicates that such capability components cannot be simply evaluated based on one particular mission capability. Balancing between missions needs to be introduced, as it is likely that there are performance trade-offs between missions. Also, the appropriate balance can change over time reflecting the changing stakeholder needs. To enable expressing trade-offs across missions, standardised indicator metric of mission capability is required. Considering the above described characteristics of capability components and capability acquirer needs, the overall aim of the research was defined as follows.

Aim: Enable the evolution planning of capability components from the perspective of capability acquirer.

This highlights a gap to be filled, i.e., a method for aggregating realised capabilities in different missions in accordance with the acquirer needs (priorities). Here, needs are assumed to be determined top-down from the organisation-wide, strategic level priorities. Mission-level trade-offs are influenced by high-level priorities, which are determined in upstream planning activity, possibly carried out separately by different decision makers. For example, the needs for capabilities in air-to-air and air-to-ground emerges as the result of the consideration into the larger armed conflict which the acquirer needs to prepare for. If the more imminent adversary is perceived to be terrorist organisations, the air-to-air mission would be deemed of lower priority because of terrorists' limited air assets. By contrast, if "symmetric" conflict against other nations is recognised as a prime concern, then the air-to-air mission should be given higher priority. With this hierarchy of needs, cause-effect relationships between high-level and low-level can be described. This is especially important when temporal changes are considered, as some changes may occur at high-level such as changes of national strategies, and they can be included in the model if this hierarchy is considered. These complex multi-tier relationships of concepts highlighted a need for a novel analytical framework as the enabler of the research aim. By using a consistent framework for design and evolution planning activities, systematic and traceable decisions are expected to be possible.

Therefore, the primary objective of this research was determined as the formulation of such a framework.

Objective: Devise a framework which aids evolution planning of capability components.

The expected practical benefit is seen as the establishment of a link between capability planning, which is performed top-down in the acquirer's organisation,

and individual system-level design evolution planning which needs to involve the supplier.

Further specification of the objectives follows from the state-of-the-art review in Chapter 2.

1.3 Research Methodology

As stated in the aforementioned research objective, this research pertains to a development of a framework which can be seen as a design support tool for capability system, which is a system in its own right. Such a tool is required to be able to appropriately reflect the nature of the capability development practices, actually support the practices, and create an industrial impact. To investigate the extent of fulfilment of these requirements, verification and validation of the developed tool were also carried out as a part of this research.

As a guidance for an approach to the problem, Design Research Methodology (DRM) (Blessing & Chakrabarti, 2009) was referred to. According to it, this “Development of Support” type of research falls into “Type 3”, which is recommended to consist of the following four stages: review-based research clarification (RC), review-based descriptive study-I (DS-I), comprehensive prescriptive study (PS), and initial descriptive study-II (DS-II).

The RC stage of this research employed literature review as the method to achieve these goals. The review covered the current state-of-the-art of system design methodologies for evolving systems, aiming to identify challenges and opportunities within these areas of practice. As the result of the review, the research aim and the initial objective of this research were formulated.

The subsequent DS-I stage was carried out following the formulation of the aim and objective. While literature review was continued as the method for this stage, its scope of the review was expanded to investigate the characteristics of capability, the current practices of capability systems engineering, and specific analytical methodologies such as decision analysis, mission analysis, operation modelling, and so forth. The purposes of this stage were to understand the identified challenges from a broader perspective, as well as to identify candidate

solutions for these challenges. As the result, the initial objective was revisited and reformulated into 3 refined, more specific objectives, which are stated in Section 2.6. Hence, an iteration took place between the RC and the DC-I stages.

In the next PS stage, a design support tool was developed in response to the formulated research objectives. As the overarching effort, an analytical framework was formulated. Considering the characteristics of capability and practices related to capability, which were identified in the literature review, assumptions were made to underpin the framework. Specific methods, which are to enable analysis based on the framework, were also developed. Additionally, a part of the test and evaluation effort of the developed tool was also carried out during the PS stage. Specifically, it is the Support Evaluation phase, where the functionality of the tool was tested to determine whether it can produce the intended result. To enable the evaluation, the framework was implemented in an interactive computer programme.

The rest of the test and evaluation work of the developed design support tool was carried out as the DS-II stage of the research. This leads to the verification, as well as the validation of the entire tool, which in turn indicates the successfulness of the overall research. As the DRM (Blessing & Chakrabarti, 2009) recommends, the evaluation in DS-II stage was divided into Application Evaluation and Success Evaluation phases. Each of these evaluation phases were carried out from different perspectives.

The Application Evaluation tested the potential applicability of the tool to an intended use case. This was carried out by applying the framework to a use case and investigating the output. Because of the lack of real-world data available for this research, an application to real use case was not possible. Instead, a hypothetical use case was developed within this research, with an intention to make it sufficiently realistic so that the challenges the framework addressed were appropriately represented.

In contrast, the Success Evaluation tested the overall successfulness of the design support tool. This includes the usability and usefulness of the tool from the viewpoint of the potential users in relevant industrial practices. In other words,

whether the developed framework can successfully deliver the intended benefit to the intended user was examined. In this case, the benefit is a provision of information for decision making, and the beneficiaries are defence acquisition planners and decision makers. In light of this, industrial evaluation was employed as an additional evaluation method to collect subjective opinions on the framework from a viewpoint sufficiently similar to its intended users. This also complemented the lack of a real-world application use case, as the evaluators were considered to be able to judge the potential ability of the framework when it is applied to real use cases, because of their knowledge on the real-world problems. The evaluators were also asked to evaluate the potential industrial impact of the framework, which indicates the validity of the research. Additionally, throughout the entire evaluation, the validity of the underpinning assumptions was critical for the framework's validity. Therefore, this industrial evaluation opportunity was utilised also to validate these assumptions against the real practices in industry, by leveraging the knowledge of the evaluators. Through this evaluation process, the verification of the framework, as well as the validation of the research against the overall aim was attempted.

1.4 Thesis Structure

This thesis is organised into 7 chapters. This Chapter 1 provides an introduction to the entire research, including the background and the motivation behind it. The overall aim and initial objective, which were formulated as the output of RC stage of the research, are also described. In Chapter 2, the summary of the literature review, which were conducted as RC and DC-I stages of the research, is presented. In the chapter, the breadth of the review is covered area by area in each section. It also includes a critical discussion as Section 2.6, where research gaps were identified, and the initial objective of the research was further clarified into 3 more specific objectives.

In Chapters 3, 4, and 5, works conducted as PS stage are presented. They are the responses to the three clarified research objectives. Chapter 3 pertains to the definition of a novel analytical framework, while Chapters 4 and 5 pertain to the development of specific analytical methods to be used under the framework.

Chapters 4 and 5 also cover the initial evaluation efforts made to verify the functionality of the methods. These evaluations were carried out by applying the methods to specially developed use cases. Descriptions of these use cases also constitute these chapters.

Chapter 6 is entirely devoted to a discussion of the evaluation effort regarding the overall framework, as well as the whole research. It corresponds with the DS-II stage of the research. An application of the framework to a comprehensive use case, which was carried out as Application Evaluation, is outlined in Section 6.3. Also, the outcome of industrial evaluation is reported in Section 6.4.

In the final Chapter 7, the conclusion of the thesis is presented. The summary of the research is given in the first section, followed by separate sections for the list of research findings, research contributions, and suggestions for future work.

2 Literature Review

Presented in this chapter is a critical literature review of state-of-the art in the fields relevant to this work. The objective is to establish potential gaps, leading to a further refinement of the objectives of the research.

The chapter is organised as follows: In Sections 2.1 to 2.3, the review of existing generic design exploration methods, together with their expansion to account for temporal changes, is presented. These methods are expected to underpin the envisaged framework as stated in Chapter 1. In Section 2.4, capability and related concepts are discussed. In Section 2.5, approaches to models which supply input data for the framework are described. The refined objectives are stated in Section 2.6, which is followed by defining the scope of this research in the subsequent Section 2.7. In Section 2.8, concluding remarks are made.

2.1 Value-Centric Design Methodologies

Satisfaction of stakeholder needs is the primary objective of systems engineering. According to the international standard ISO/IEC/IEEE 24765, systems engineering is defined as follows:

Interdisciplinary approach governing the total technical and managerial effort required to transform a set of stakeholder needs, expectations, and constraints into a solution, and to support that solution throughout its life. (ISO/IEC/IEEE, 2017)

Design is an essential activity carried out as a part of systems engineering, and therefore the needs satisfaction should be treated as its primary objective as well. In the traditional approach, requirements are first formulated by translating the needs. Then, these requirements are used as an input to the subsequent system design. When designs are formulated, their needs satisfaction is assessed by judging whether these requirements are met. Alternatives which do not meet all the requirements are considered unsatisfactory.

While this so-called requirement-centric design has been widely practiced, its suitability to capture the true stakeholder needs has been questioned. This

problem is illustrated in (Hastings & McManus, 2004, pp. 70–71). Requirements essentially “reduce” a design space into binary, making “all designs that meet requirements are equally good. All designs that fail to meet requirements are equally bad” (Brown et al., 2009). Additionally, requirements are usually defined individually on each attribute (criteria) and trade-offs between each other are not expressed. This problem is illustrated in (Collopy & Hollingsworth, 2012, p. 756). Hence, “Requirements only tell what you do not want” (Collopy & Hollingsworth, 2012). These shortcomings of requirements-centric design are similarly discussed in (Hupman et al., 2015).

In (Ricci, 2014; Ricci et al., 2014), conceptual design activity of complex systems is illustrated as a design-performance-value loop, or “design loop” (Figure 1). This can be viewed as a generic representation of DSE (Design Space Exploration) activity, where design alternatives are discovered and evaluated. Design alternatives populating a design space are first mapped onto a performance space spanned by attributes of interest, then mapped onto a value space. These two mappings are provided by a performance model and a value model, respectively. (Ricci, 2014) states that performance assessment has a certain degree of objectivity and measurability, while stakeholder satisfaction, which value models are representing, is inherently subjective and difficult to measure. While performance model is named as such, cost and other non-performance attributes of interest can be its output unless they are affected by subjective stakeholder preference. In other words, cost model can be a part of performance model. A requirement set can be viewed as a value model which outputs binary, indicating whether satisfactory or unsatisfactory.

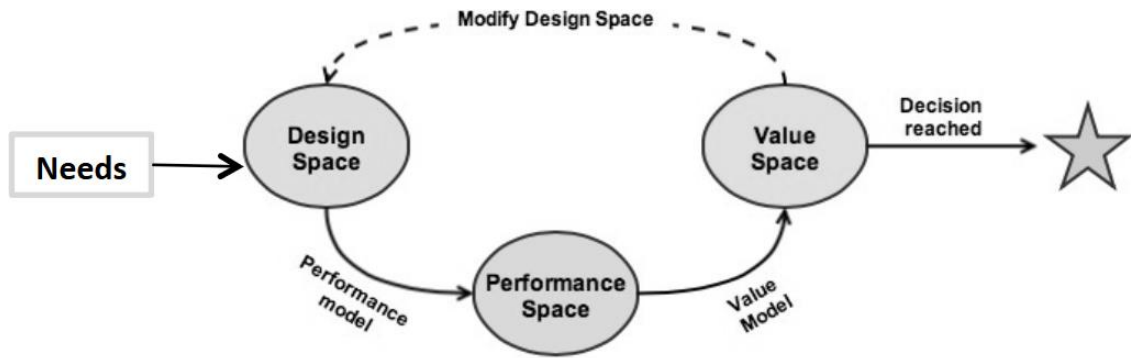


Figure 1 Design Loop (Ricci et al., 2014)

As an alternative to requirement-centric design, Value-Centric Design Methodologies (VCDMs) (Brown et al., 2009) was proposed. These can be viewed as a part of decision-based engineering design (Hazelrigg, 1998), which incorporates the methodology of decision analysis into engineering design. Value function and utility functions, which are preference representation function under certainty and uncertainty, respectively, are used for comparing design alternatives. VCDMs have been practiced in different ways, under different names: Value-Driven Design (VDD) (Collopy & Hollingsworth, 2012), Decision-Based Design, Value-Based Design (Abbas & Cadenbach, 2018), and so forth. Here, *value* refers to the metric which represents the desirability of alternatives in accordance with the preferences of decision makers and it does not necessarily reflect the monetary worth, as the uses of the word in more general context suggest.

A review of various perspectives and methods of VCDMs was provided in (Collopy, 2009; Ross et al., 2010). While the use of value or utility function is a commonality among them, their scope is not entirely same; for example, proponents of VDD emphasise that VDD need to be compatible with numerical optimisation techniques, hence its output need to be scalar as it can serve as their objective function. They argue that MATE (Multi-Attribute Tradespace Exploration) cannot be a part of VDD because value (or utility) and cost are separately calculated, yielding 2-D output (Collopy & Hollingsworth, 2012). In this thesis, VCDMs will be used as an umbrella term collectively referring to these methods. In line with the concept of design loop (Figure 1), The characteristic of

VCDMs is that their value models output one or multiple cardinal numbers to capture the stakeholders' preference better, in contrast to the binary output in the requirement-centric design.

A methodology of decision analysis is outlined in (Parnell & Tani, 2013, Chapter 3) with a taxonomy of practice (Figure 2). From the perspective of decision analysis, the approach of VCDMs is in line with Value-Focused Thinking (VFT) (Keeney, 2008), which is characterised by four major ideas: “start first with values, use values to generate better alternatives, create decision opportunities, and use values to evaluate the alternatives” (Parnell & Tani, 2013). This approach is in contrast to Alternative-Focused Thinking (AFT), which is intended to differentiate alternatives which are identified beforehand. As the individual objectives (criteria, attributes) against which alternatives are evaluated, VFT suggests making distinction between *fundamental objectives* and *means objectives*. The former are criteria which are valued directly by the decision makers as ultimate goals, while the latter contribute to fundamental objectives, thus their contribution to the ultimate goals are indirect. The use of fundamental objectives is recommended as the components of value and utility function.

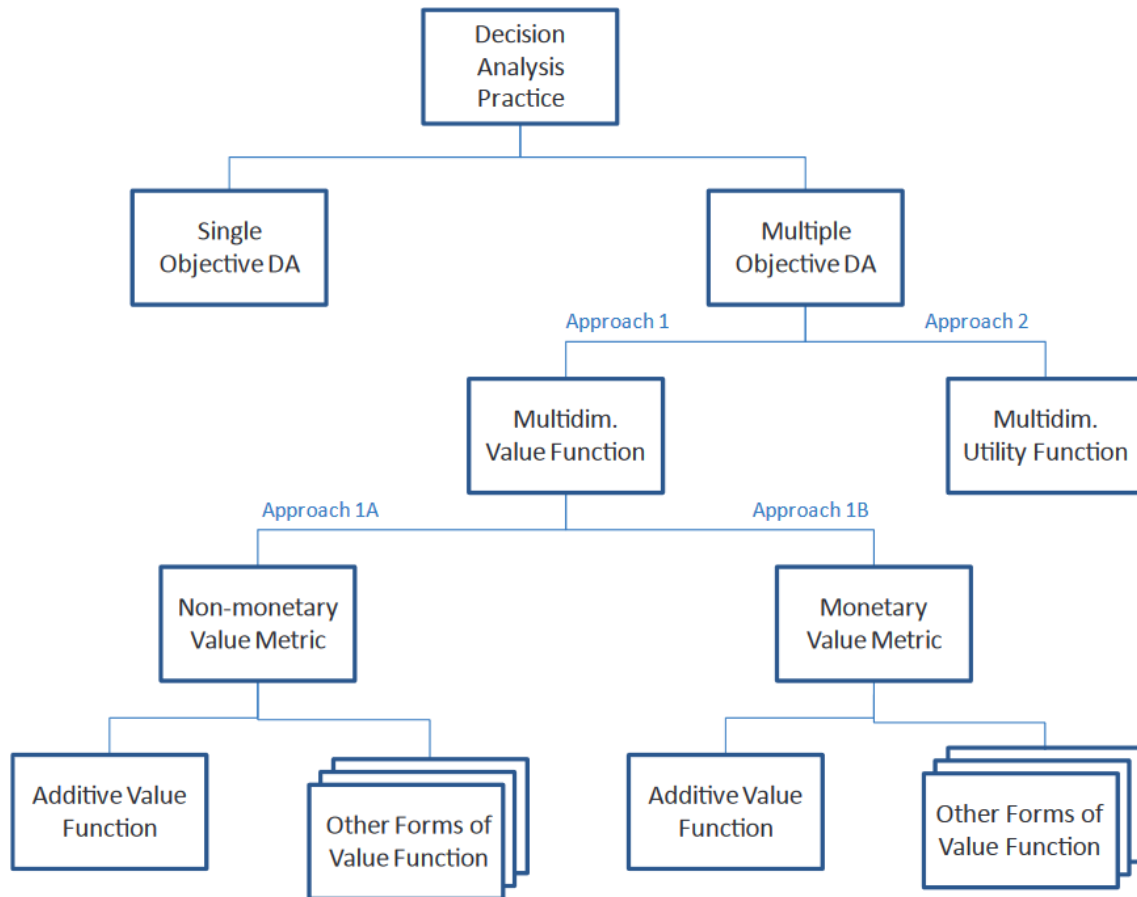


Figure 2 A Taxonomy of Decision Analysis (DA) Practice (Parnell & Tani, 2013, p. 55)

The taxonomy of practice can be applied to different approaches of VCDMs. While value function represents stakeholders' preference on the trade-offs among the multiple objectives (or attributes), utility function can represent stakeholders' preference regarding risk-taking as well. Within the taxonomy in Figure 2, distinction between Approach 1 and Approach 2 is made depending on how this risk preference is treated when utility function is defined. In Approach 1, components of value function are first aggregated across objectives, yielding a multi-dimensional value metric. Then it is transformed into utility considering the risk preference. In Approach 2, on the other hand, components of utility function regarding different objectives are calculated accounting for the risk preference, then aggregated considering the trade-off preference. If the risk preference can be considered as invariant across attributes, then consideration into uncertainties

is sufficient in relation to the aggregated multidimensional value function, and therefore Approach 1 can be used. Otherwise, Approach 2 needs to be used.

Another taxonomy division is made between Approach 1A and 1B, based on the use of monetary or non-monetary value metric. In light of the importance of considering cost in system design, monetary value metric directly comparable to cost is “practical and generally convenient” (Collopy, 2009), especially for the systems whose goal is making profit. Therefore, if stakeholders can establish trade-off between cost and each benefit, Approach 1A would be desirable. Otherwise, Approach 1B will be chosen. Multi-Attribute Utility Theory (MAUT) is a method which uses one such value model, where aggregation of non-monetary benefits is quantified as Multi-Attribute Utility (MAU), to be related to monetary cost (Ross et al., 2010). In either approach, it is crucial to capture the stakeholder’s preference.

It should be noted that there is a confusion regarding the use of terms *value* and *utility* within the literature on VCDMs, as stated in (Abbas & Cadenbach, 2018). In decision theory, on which VCDMs are based, they are distinguished based on the consideration into uncertainty. Generally, VCDMs are capable of handling uncertainty and risk, therefore they are formalised around utility. Some applications of VCDMs only consider deterministic cases, and according to the decision theory, their preference metrics are *value* in fact. Nevertheless, their preference metrics are often called *utility* in literature. In this thesis, the distinction will be made in line with decision theory.

Value modelling is central to VCDMs, and various models have been proposed and discussed (Collopy, 2009; Ross et al., 2010). In some literature, principles of VCDMs are simply applied to allow trade-offs, ranking alternatives, and optimisation, using value or utility functions which are defined rather arbitrarily. Advanced value models and modelling methods have been proposed as well. For example, (Papageorgiou, 2018; Papageorgiou et al., 2016) employed the Analytic Hierarchy Process (AHP) and game theory to model decisions across multiple stakeholders.

MATE or Multi-Attribute Tradespace Exploration (Hastings & McManus, 2004, sec. 4.3; Ross & Hastings, 2005) is a notable example of VCDM which saw a wide range of applications. According to the definition, tradespace refers to “the space spanned by the completely enumerated design variables” and it is populated by finite number of possible design options. The value metric is defined with MAUT, therefore the output of value model is a pair of MAU and cost. While one may further aggregate MAU and cost to obtain scalar metric (e.g., utility per cost), Utility-Cost plot is commonly used to visualise the result for exploration. While MATE is a comprehensive MCDM methodology in itself, it also served as a basis for Dynamic MATE and Epoch-Era Analysis (EEA), which are introduced in Section 2.3. Glossary of terms regarding MATE is provided in (Ross & Hastings, 2005).

2.2 Use of Effectiveness in Design

VCDMs, which were discussed in the previous section, can be seen as methodologies distinguished by the employed value model in the design-performance-value loop (Figure 1). On the other hand, as the whole design activity aims to reflect the acquirer needs more directly, the use of “high-level” performance measures for the inputs of the value model is being suggested. Here, measures which can be defined, measured, and verified within a relatively small context are referred to as “low-level”, while the opposite are called “high-level”. Taking fighter aircraft as an example, instances of low-level measures are maximum speed or turn rate, while those of high-level are loss exchange ratio in 4-on-4 air combat (i.e., ratio of losses of enemy force to friendly force). While the former is only related to single aircraft, the latter involve interactions with other friendly and enemy aircraft. Using the terminology of VFT, high-level measures are likely to pertain to the acquirer’s fundamental objectives, while low-level are to means objectives at best. In other words, high-level measures are likely to be able to validate the system in terms of the needs. In contrast, low-level measures are relatively simple to define, measure, and verify, therefore likely to be used in system requirements. In such a case, the validity of requirements against the needs may be questioned.

Traditionally, the linkage between the high-level and low-level measures are established separately outside the design loop, through operational analysis activities. In such case, in the previous example, the acquirer first needs to gain confidence that the fighter can perform sufficiently well in combat if certain speed and turn rate are achieved. Typically, they conduct parametric studies to investigate the sensitivities of speed and turn rate on the exchange ratio. Then, based on the results, they set requirements for speed and turn rate for use in the design loop.

Recently, design methodologies which utilise complex performance model embedded in the design loop were developed. These performance models can output measures for fundamental objectives, which are often referred to as effectiveness or capability. In (Biltgen, 2007; Biltgen & Mavris, 2007), a quantitative evaluation framework of system-of-systems in a mission was proposed. Effectiveness-Based Design or EBD (Allison & Kolonay, 2018) is a recently developed concept which combines aircraft design synthesis and operational analysis in single loop, enabling Multidisciplinary Design Optimisation (MDO) of a combat aircraft against operational outcome in a mission. This trend is underpinned by the development of M&S (Modelling and Simulation) techniques, together with increased computational power to run through the calculation. Validity of the performance model is requisite for these approaches to produce reliable results, and high-fidelity models commonly dubbed as “digital twins” are potential enablers (Harper, 2021). Additionally, when the system being designed is operated in complex operations such as combat, Machine Learning and Artificial Intelligence (AI) often need to be inserted in their models to control each simulated system and provide tactics for each side.

2.3 Design Exploration in Dynamic Context

The DSE technique was expanded to enable evaluation of changing (including evolving) system alternatives. A prominent example a derivative of MATE. Dynamic MATE was applied for exploring system alternatives in terms of changeability (Ross, 2006; Ross & Hastings, 2006). It explores alternatives within “tradespace network” in contrast to ordinary tradespace exploration. Tradespace

network contains design points (or nodes) of each point in time and transition paths connecting one design point to another (transitioning during two points in time). By using this, value increment (or decrement) by design changes, as well as required cost, can be considered in the exploration of alternatives regarding changing system. Dynamic MATE was combined with EEA (Ross & Rhodes, 2008), which addresses consideration into factors exogenous to a system. These factors may change in the future, and these changes may, in turn, affect the value of the system. Therefore, it is crucial to take them into account when value delivery in a dynamic world is of concern.

In (Ricci, 2014, pp. 78–79), it was proposed that factors influential on a system to be classified into three spaces. Design space pertains to the system itself, whose design is controlled, or the “inner environment”. Context space pertains to the environment in which the system operates, which is exogenous to the control of designers, or the “outer environment”. Needs space pertains to what is desired of the system, which is also exogenous to the designers’ control.

In the terminology of EEA, an epoch is defined as “discrete time period with fixed context and needs (i.e., value expectations)” (Ricci, 2014, p. 81). A sequence of epochs forms an *era*. Fixed context and needs within an epoch enable evaluating system designs per each epoch. With the use of epoch and era, system design in a dynamic circumstance was made possible. In (Ricci, 2014, p. 82), Figure 1 was modified to illustrate the concept of epoch as shown in Figure 3.

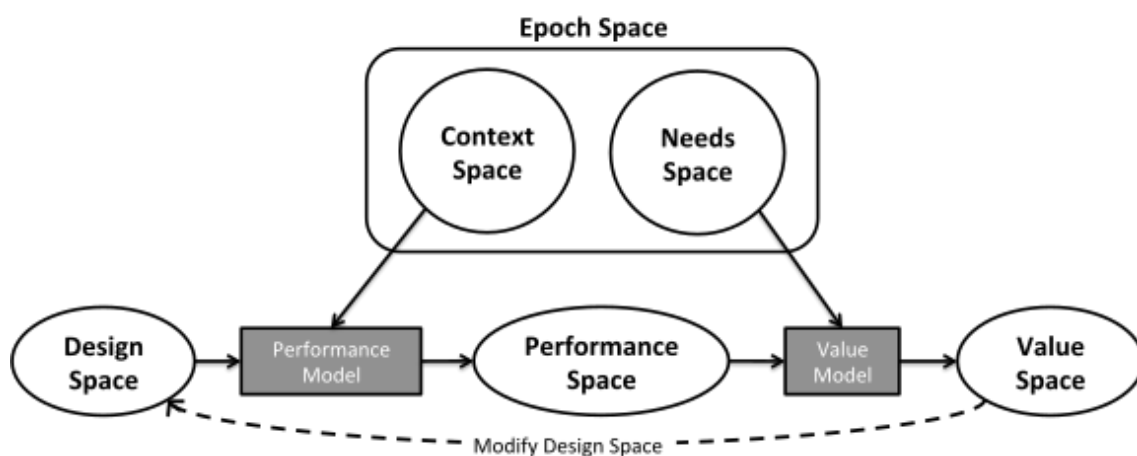


Figure 3 Design loop under dynamic system perspective (Ricci, 2014, p. 82).

EEA can be viewed as a “computational scenario planning methodology” (Roberts et al., 2009), as it characterises one plausible unfolding of future with one scenario. Scenario in this context can be specifically referred to as *developmental scenario* (Bakken, 2003, sec. 2.2), in contrast to *situational scenario*, whose focus is on a particular future situation. Characteristics of scenario planning in uncertainty modelling are discussed in (O. L. de Weck et al., 2012).

The combination of MATE and EEA saw a wide range of applications for exploring alternatives of evolving systems. For example, in (van Heerden, 2018; Van Heerden et al., 2018), it was integrated with a model which predicts the redesign effort associated with a transition from one design to another. Termed “evolvability exploration”, the proposed technique enabled identifying promising alternatives in a long-run, which are expected to offer good long-term cost-to-benefit by leveraging evolvable design.

2.4 Capability as a subject of systems engineering

2.4.1 Concept and Perspectives of Capability

This research aims to propose a framework which informs individual systems’ stakeholders regarding the alternatives of their System-of-Interest (Sol) from a broader perspective of capability development. It is assumed that customer needs for individual systems, which is an input to the framework, can be derived top-down from the results of capability development activities. Therefore, to take customer needs into account appropriately in the design, it must be related to the upstream activity of capability development.

Prior to using the concept of capability in this research, there is a need to clarify what is implied with *capability* in terms of semantics. According to the *Oxford Advanced Learner's Dictionary*, the core meaning of *capability* is “the ability or qualities necessary to do something” (Lea & Bradbery, 2020). However, even without a departure from this core meaning, there is a shared recognition that *capability* is only a meaningful concept when it is defined within a context (Yue & Henshaw, 2009)(Anteroinen, 2012). In particular, its characteristic is dependent on the level at which capability is being considered (Yue & Henshaw, 2009). It

was also observed that the term *capability* is used interchangeably to refer to the eventual objectives, the supporting tasks that need to be accomplished in support of these objectives, and the wherewithal (or solution) to conduct those tasks (Fitzsimmons, 2007). Therefore, when the concept of capability is used, the viewpoint of the user needs to be clarified for consistency. In the literature review of this research, existing works which investigated the nature of capability and its multitude in practices are explored first.

In (Henshaw et al., 2011), eight world-views (WX, X=1...8) on which practices regarding capability and capability engineering are based were identified from across several different areas:

- W1 – Equipment has Capability
- W2 – Capability used to describe solution independent requirements
- W3 – Deciding / balancing in which capability to invest
- W4 – Service Provision
 - W4a – Delivering specific business services
 - W4b – Developing fallback services to be used at a future date
- W5 – Reconfiguring available assets, people and processes quickly to meet current circumstances
- W6 – Developing a capability (all Components of Capability) solution
- W7 – Managing the interdependencies between capabilities
- W8a – Organizations have capability
- W8b – Capabilities emerge through processes of interaction between individuals, groups and organizations

All these world-views were treated as valid, but at the same time, some contradictions between them were pointed out as well.

In a military-specific context, (Anteroinen, 2012, 2013) described several different perspectives of military capability:

- 1) Military capability as an instrument of foreign policy
- 2) Capability as an effect or a function to execute tasks
- 3) Capability as systems

- 4) Capability as weapons system or a platform
- 5) Capability as fighting power through military units

As can be seen, the world-views and the perspectives are interrelated. Based on these, it was possible to identify the characteristics of capability which need to be considered in the context of individual system design. From low level to high level:

Characteristic 1: An equipment has a capability, which is synonymous to function or performance.

This is based on the W1 and 4) above, and such capability can be referred to as equipment capability. These capabilities are likely to be associated with low-level tasks, whose outcomes are not directly required by stakeholders, but may turn out to be necessary to form building blocks thereof. For example, the “capability” to achieve supersonic flight is not necessarily required, but it is rather required to perform some larger tasks, such as air-to-air combat, with a sufficient effectiveness.

Characteristic 2: A capability is realised through a solution, which is constituted by capability components and treated as a system in itself.

This is based on the W6 and 3) above, and the focus is on the internal structure of capability solutions.

Characteristic 3: Stakeholders require capabilities to perform a set of tasks through achieving desired effects.

This is based on the W2 and 2) above, and states that what stakeholders need are capabilities to perform high-level tasks, which are needed directly by the stakeholders and regardless of solutions. These are in contrast to low-level tasks, which were discussed in Perspective 1.

Characteristic 4: The purpose of capabilities is the achievement of the overall goal of an organisation, and different capabilities can be traded-off with each other for this purpose.

This is based on the W3, W7 and 1) above, and assumes the high-level structure of needs with trade-off relationships.

Within these four, there is a clear difference between characteristic 1 and the others in that the capabilities in question are not directly associated with the capability needs of the stakeholders. While there is no need to deny the validity of characteristic 1, such “capabilities” will be referred to as performances or functions for a clarity. Also further in this thesis, *capability* will only refer to the ability and distinguished from task itself, its objective (desired outcome, or effect), and wherewithal for it (capability solution). The definition of capability in this thesis will be “The ability to complete a task under specified conditions and level of performance” (Adopted from (OUSD(R&E) USDoD, 2020)).

2.4.2 Capability Development Activity

In the military domain, Capability-Based Planning (CBP) has been an internationally pursued approach to defence capability development since the first decade of 21st century. In (Davis, 2002), it was defined as “planning, under uncertainty, to provide capabilities suitable for a wide range of modern-day challenges and circumstances while working within an economic framework that necessitates choice”. Implementation-independent guidance is provided in (Joint Systems Analysis Group Technical Panel 3, n.d.; Taliaferro et al., 2019; Walker, 2005).

CBP links procurement decisions to strategic goals (Joint Systems Analysis Group Technical Panel 3, n.d.), which means that strategic goals, which are defined top-down, flow down and define the stakeholders needs for each solution and capability component. In this sense, in the context of CBP, stakeholder needs pertain to capabilities rather than particular solutions. In response to these needs, suppliers develop solutions and components thereof (i.e., capability solutions), which are to be acquired by acquirers. The nature of this activity was described in (M. D. Guenov et al., 2006) from the suppliers’ viewpoint, addressing the challenge in customer (acquirer) needs elicitation. In order to enable considering such capability needs, it was deemed essential to investigate the characteristics

and activities of CBP to reflect them on individual system design. An illustrative diagram of CBP activities is given in Figure 4.

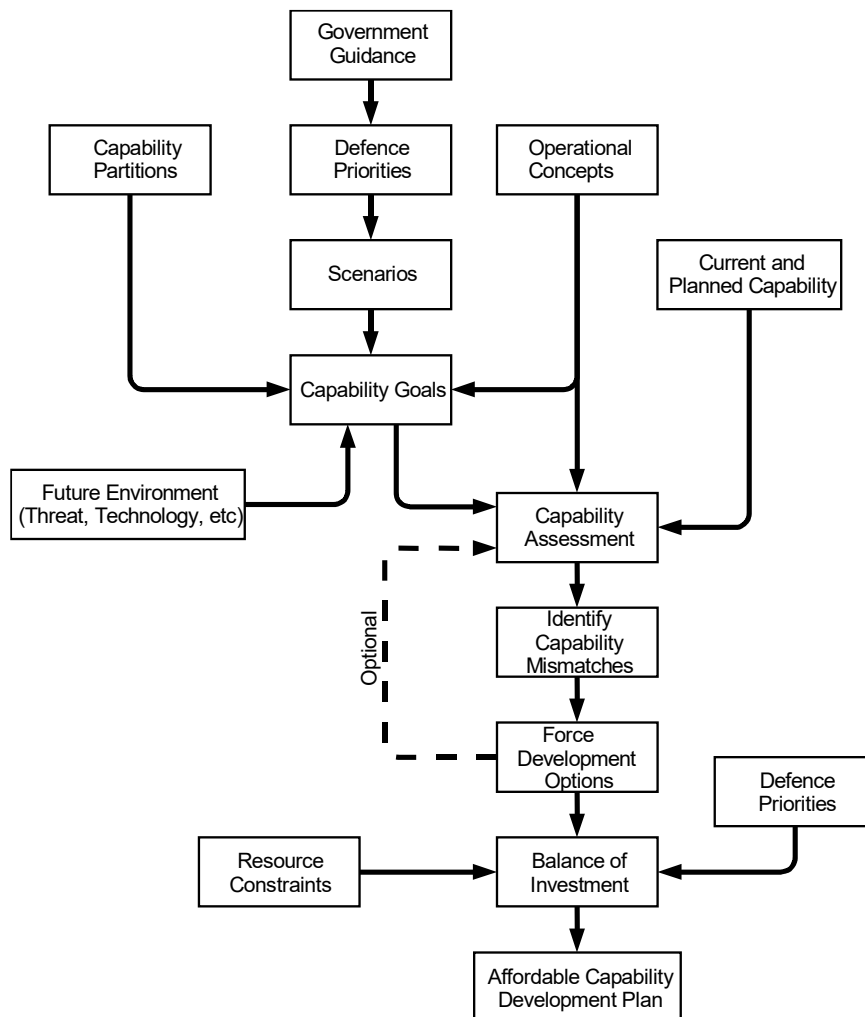


Figure 4 Generic Process Chart of Capability-Based Planning (Joint Systems Analysis Group Technical Panel 3, n.d.)

While this research does not aim to develop a method which is compliant to a particular implementation of CBP (e.g., Joint Capabilities Integration and Development System, or JCIDS, of the US), general principles of CBP need to be understood and what can be expected from CBP need to be clarified to use for the input of this research.

The following are important characteristics of CBP identified from the literature review:

Focus on goals and end-states. CBP starts with what needs to be done and explores broad solution alternatives, rather than narrowing down to particular ones early in the process. The most typical such example is simply replacing retiring equipment by developing a new one of the same type. Outcome-focused, solution-independent goal definition potentially facilitate development of innovative solutions and cost-effective solutions, and they are not limited to materiel solutions. In this sense, a famous proverb in marketing, “People don’t want to buy a quarter-inch drill. They want a quarter-inch hole.”, agrees with the concept of CBP.

Recognition of interdependence and trade-offs of capability components. A wide range of solution alternatives, constituted by different capability components, are explored, which include materiel and non-materiel solutions.

Top-down approach. Needed capabilities are derived from top-level strategy and prioritised in accordance with strategic guidance. A taxonomy of capability partition such as the Joint Capability Areas (JCA) of the U.S. is an example.

Use of situational scenarios, providing context against which to measure the level of capability. For a given solution and a given task, the realised levels of capability are dependent on the contextual factors. For example, a task to “defend friendly airspace” may be successfully accomplished when the attacking adversary systems are obsolete, otherwise it may not. While this example is at a tactical level and influential to low-level capabilities, high-level factors generally have greater and wider influence. For example, if two armed conflicts occur simultaneously rather than one, it will affect strategic-level as well as tactical-level. In CBP, a set of assumptions regarding these factors is compiled as a situational scenario, which describes a plausible situation. In (Taliaferro et al., 2019, Chapter 4; Webb et al., 2010), factors which ought to be included in scenarios are listed as follows:

- The political military context, e.g., how the situation came about, who is allied with whom, the degree of strategic warning, forward stationing of forces, etc.
- The objectives and strategies of all actors involved (i.e., success criteria)

- The forces, such as size, character, nominal capabilities
- Force effectiveness, accounting for cohesion, morale, etc., of all involved forces
- The natural environment
- Other assumptions.

Additionally, a guidance on the use of scenarios in defence planning is provided in (Bakken, 2003, Annex 1). It is envisioned that scenarios, if defined appropriately, collectively capture the range of uncertainty regarding the situation anticipated by the planner. They can be either specific (having a fixed relationship to particular nation and geography) or generic (having generalised characteristics that are representative of a class of adversaries or a geographical region). However, a failure to recognise the breadth of the range of uncertainty may result in unreliable results, as overspecialised, inflexible solution alternatives may be overrated in such cases.

Continuous execution based on gap analyses. The CBP process is essentially twofold: First, the capability of As-Is solutions for tasks are assessed against situational scenarios to identify capability mismatches (gap or excess). Subsequently, To-Be solution alternatives are generated to mitigate the mismatches and assessed in terms of the capabilities and the costs associated with the implementation. The entire process takes place continuously (or periodically) to maintain the capability in a dynamic context.

“Joint” perspective. The overall activity, including capability assessment and solution formulation, are carried out organisation-wide (i.e. jointly across services). Through the assessment of total organisation-wide capability, the elimination of redundancy across services, and the optimal resource allocation, organisation-wide coherence and efficiency are expected to be achieved. While this aspect is a crucial consideration for an effective CBP and viewed as one of its major advantages over traditional procedures, the extent of realisation of the potential benefits of the joint perspective is dependent on the implementation of CBP in an actual organisation.

2.4.3 Capability Systems and Systems-of-Systems Engineering

The engineering approach to capabilities is conceptualised as Capability Systems Engineering (Kemp & Daw, 2014). As the distinctive characteristics of capability systems from traditional product systems, three points were identified:

1. They focus on delivering higher-level outcomes, or effects, rather than lower-level outputs, or performance.
2. They are enduring and evolutionary. Upgrading existing capability is usual, rather than building a new one.
3. They are constituted by different components spanning people, processes, equipment, and organisations.

These characteristics of capability systems have some overlaps with those of systems-of-systems (SoS), which is a more common concept. As an international standard, ISO/IEC/IEEE 21839 (ISO/IEC/IEEE, 2019) described concepts relevant to SoS. According to it, the definition of system of systems is “set of systems or system elements that interact to provide a unique capability that none of the constituent systems can accomplish on its own”. Hardware, software, humans, procedures, and facilities can be a part of the systems. Also, citing (Maier, 1998), five key characteristics of SoS were listed. Their description was made in (J. S. Dahmann, 2015; Henshaw et al., 2013) as well. They can be summarised as:

1. **Operational independence of constituent systems.** Each constituent system can operate independently, without the need to interact with other systems, to fulfil some number of purposes on its own.
2. **Managerial independence of constituent systems.** Each constituent system is likely to be managed by organizations with a level of independence, with potentially different goals and objectives for the constituent systems. Constituent systems may be owned, operated, or evolved independently by different organisations.
3. **Geographical distribution of constituent systems.** In such case, primary concern of SoS engineering is the information exchange between constituent systems.

4. **Emergent behaviour.** As in conventional systems, there are behaviours exhibited as consequences of interactions of elements (constituent systems) of the whole (system-of-systems), rather than by each individual element itself.
5. **Evolutionary development processes.** SoS is not created once and for all, but evolves as constituent systems, or functions thereof, are added, removed, or modified. These developments may take place asynchronously based on the independent development processes of the constituent systems.

Within these five, operational and/or managerial independence (1, 2) are identified as the principal distinguishing characteristics of SoS from other systems.

In the defence-specific context, (J. S. Dahmann, 2015) describes three overlapping domains of systems-of-systems: Missions SoS, Platform SoS, and Information Technology SoS (Figure 5). In the research presented in this thesis, missions SoS are of particular concern; they are defined as “sets of systems working together to provide a broader capability to support a military mission”. It is common for missions SoS to involve operationally and / or managerially independent systems with separate evolution plans, interacting with each other to generate synergy effects (which is a form of emergent behaviour).

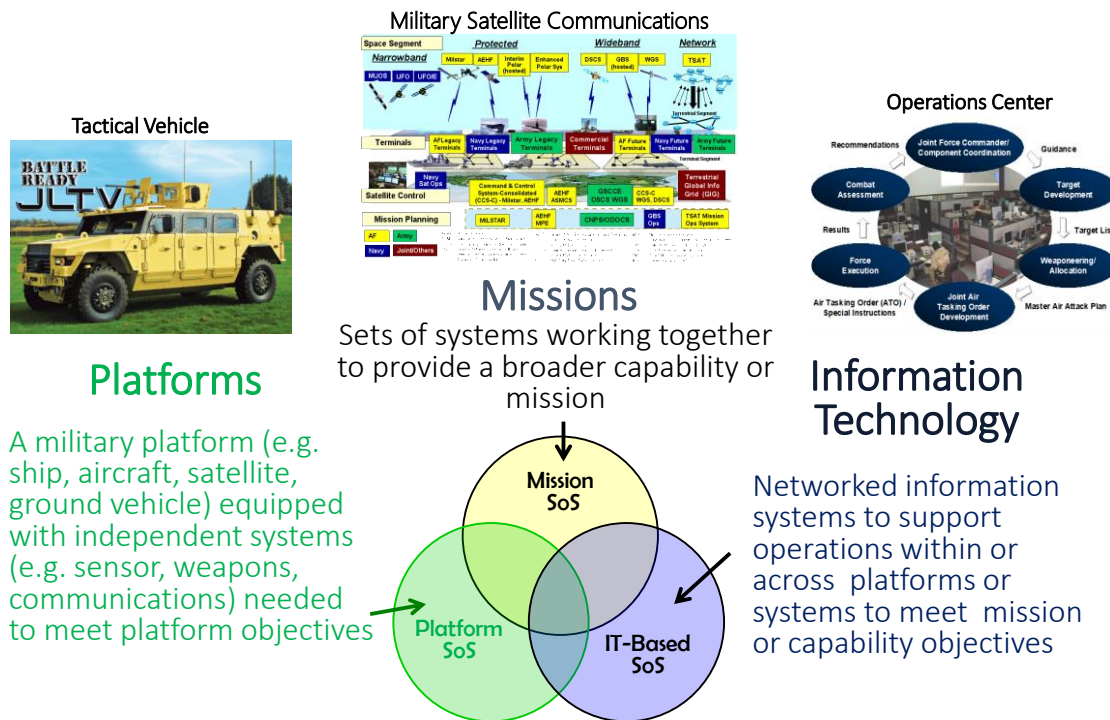
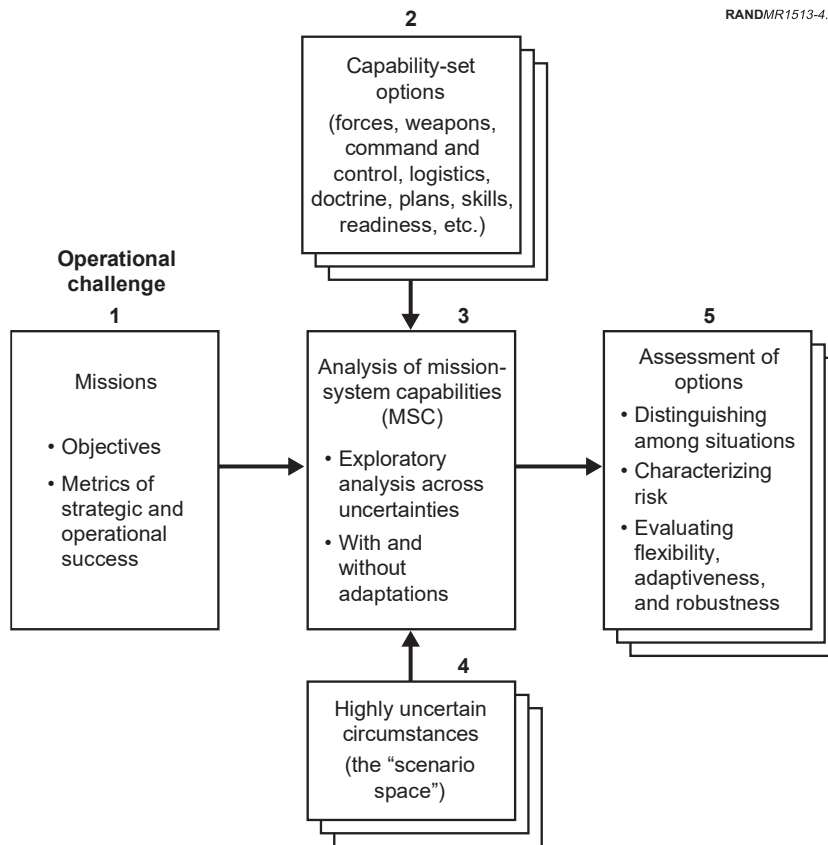


Figure 5 Systems of Systems Domains in Defence (J. S. Dahmann, 2015)

For missions SoS, specific engineering approaches were proposed. It will be discussed in detail in the next section.

2.4.4 Analytical Methodologies for Mission Execution

In (Davis, 2002, Chapter 4), it was pointed out that CBP needs to be focused on “building-block” or “modules” of tasks, which can be referred to as “missions”. They need to be largely, if not completely, independent from other tasks so that each capability can be defined, developed, and tested within the domain of a mission. It was also proposed that a mission can be viewed as a “system” composed of subordinate critical components of subtasks, whose achievements collectively represent the achievement of the entire mission. Using “a family of models and games”, exploratory analyses are carried out. Situational scenarios are inputted to the analysis to account for the mission uncertainties. Overall, this process, called “Mission-System Analysis” (Figure 6), demonstrated the validity of using mutually independent missions as the focal point of CBP. At the mission level, the relationship among solution, context, and realised capability level is analysed by operational analysis.



NOTE: For a given operational challenge (1), consider a set of options (2) for meeting the challenge; apply mission-system analysis methods (3) across a wide range of circumstances (4) to generate a comparative assessment of the options with appropriate metrics (5).

Figure 6 Mission-System Analysis (Davis, 2002)

In recent decades, analytical techniques at mission level have developed by applying methodologies of systems engineering. In the defence community of the United States, Mission Engineering (ME) has been employed in recent years. It is based on concepts from systems engineering and provides a link between mission-level operational analysis and capability development.

ME is intended to inform the broader activity of Mission Integration Management (MIM) as its technical sub-element. The official definition of MIM is “the synchronization, management, and coordination of concepts, activities, technologies, requirements, programs, and budget plans to guide key decisions focused on the end-to-end mission”. ME is defined as “the deliberate planning, analyzing, organizing, and integrating of current and emerging operational and system capabilities to achieve desired warfighting mission effects” (OUSD(R&E)

USDoD, 2020). It is envisioned that MIM activities shall be carried out for missions such as Close Air Support, Air Defence, Interdiction, Intelligence, Surveillance, and Reconnaissance. These missions are complex, and their solutions are likely to involve a set of systems integrated with each other to form a capability system.

ME has two characteristics distinct from traditional approaches. First, it applies systems engineering approach to both the mission itself and the solution. They are decomposed and assignment relationships are considered. Decomposition structure of the mission is expressed as a Mission Thread (MT), which is “a sequence of end-to-end activities and events presented as a series of steps to achieve a mission” (OUSD(R&E) USDoD, 2020). Along the MT, assignment of systems or other capability components to each task are considered. This forms a Mission Engineering Thread (MET), which is “Mission threads that include the technical details of the capabilities and systems required to execute the mission” (OUSD(R&E) USDoD, 2020). A typical example of MT is “kill chain” in a strike mission, which involve the tasks of find, fix, target, track, engage, and assess. Capability elements such as sensors and weapons are assigned to each task. The emphasis on MT and MET in ME highlights that mission success is dependent not only on what systems are used and how they are integrated, but also on how to employ them. In the concept of capability engineering, this aspect can be understood as a part of “concept” in TEPID OIL (Chapter 1).

The second characteristic of ME is the reliance on models throughout the process. “Model development, integration, use, analysis, and curation” are considered as core elements of ME activities (OUSD(R&E) USDoD, 2020). In line with the digital engineering, models are envisioned to provide consistency and traceability to the engineering activities of missions and systems. They represent both systems themselves and their interaction with the outer world in missions.

For the models of missions, the inputs are alternatives of MET with different features, while the outputs are mission metrics including Measures of Effectiveness (MoEs) and Measures of Performance (MoPs). Each mission metric represents different aspect of accomplishment of a mission or a part thereof.

Here, a distinction between MoE and MoP needs to be clarified. According to (International Council on Systems Engineering, 2023), the definition of MoE is “Measures that define the acquirer’s key indicators of achieving the mission needs for performance, suitability, and affordability across the life cycle”. On the other hand, the definition of MoP is “Measures to assess whether the system meets design or performance requirements and has the capability to achieve operational objectives”. As can be read, MoEs need to indicate the overall mission achievement, while MoPs do not necessarily. The distinctive characteristic of MoEs is that they are defined from the acquirer’s perspective, without any particular solution in mind (Roedler et al., 2005). On the other hand, MoPs are low-level measures related to the details of solutions. Therefore, in the capability-based context, assessment of solutions is primarily made based on MoEs, rather than MoPs.

From the viewpoint of individual system development, the emphasis on high-level outcome (i.e., MoEs) requires understanding on the quantitative relationships between the individual system design and the eventual effect created in a mission execution by the entire capability system. Emergent behaviour created by the interactions between capability components should also be a part of it. As with Mission-System Analysis and ME, models are typically used for this purpose.

2.4.5 Evolutionary Development

As presented in Section 2.4.3, evolutionary development is a common characteristic of capability systems and SoS. As a practical representative model from the viewpoint of implementers, wave model for SoS systems engineering (Figure 7) was proposed (J. Dahmann et al., 2011). According to it, a system-of-systems undergoes multiple, overlapping iterations of evolution. This evolution is driven by development of its constituent systems. Analyses are carried out continuously to direct each iteration, informed by continuous input from external environment. The architecture of SoS, which governs integration of these constituent systems, also evolves. Consequently, the SoS evolves through multiple waves of iteration.

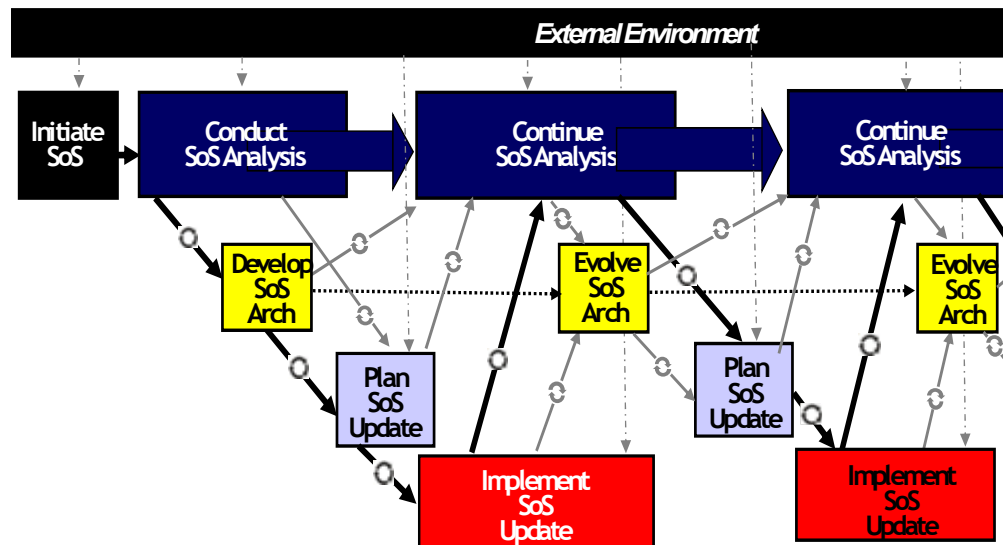


Figure 7 SoS Systems Engineering Wave Model (J. Dahmann et al., 2011)

While capability systems may contain additional capability components (e.g., people, processes, organisations) as their constituents, the components themselves may possess the enduring characteristic of SoS. Therefore, the wave model is also likely to be applicable to them.

2.5 Operation Modelling

As described in previous sections, models are commonly used to analyse systems in operation and establish a link between their design and the operation outcomes. Approximated models are constructed to represent a portion of military operation and executed to collect data. The type of the model to be used may vary according to the subject and purpose of the analysis; it can range from simple mathematical model, consisting of differential equations, to complex agent-based computational model, where behaviours of individual entities are described in a bottom-up manner. As a model becomes more aggregated, the resolution becomes lower in general because of the more approximation introduced. This trade-off relationship is due to the modelling effort and computational power to improve both aggregation and resolution, which are often prohibitively (computationally) expensive. In any cases, models must be able to capture the difference within their input and output with appropriate resolution to serve their purposes.

In the defence community, models are commonly classified into campaign, mission, engagement, and engineering levels, from high to low degrees of aggregation (Hill & Miller, 2017; Piplani, 1994, sec. 4.4). Since capability systems are intended for performing tasks no lower than mission level, the aggregation level needs to be sufficiently high. If the purpose of model pertains to the engineering decisions, its resolution needs to be sufficiently high (Figure 8). Therefore, such models need to employ analytical methods as discussed in Section 2.2.

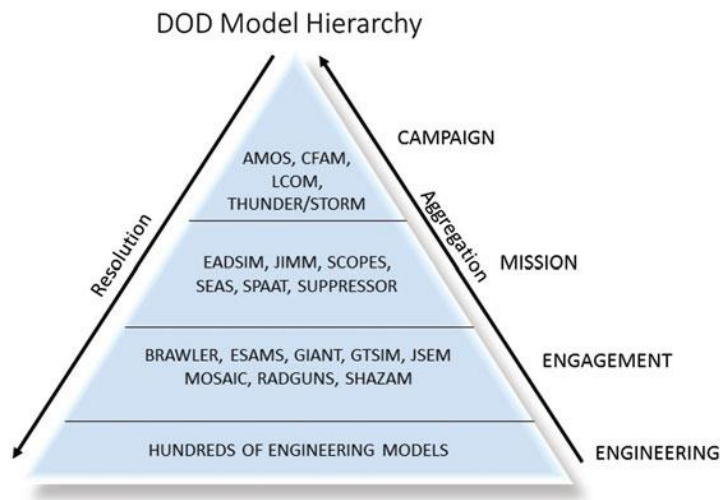


Figure 8 Hierarchy of Models Used within U.S. Department of Defense (Hill & Miller, 2017)

Output variability is another characteristic of models; models are not necessarily deterministic, where the same input always produces the same output. This reflects the uncertain nature of the subject (which is the modelling of military operations). Uncertainties might be related to agent's (human or AI) decisions, behaviour of systems, or environment (e.g. weather). Sometimes, stochasticity is intentionally built in the model to replicate these uncertainties. For instance, consequence of guided weapon use (i.e., hit or miss) can be simply simulated by generating a random number between 0 and 1, then comparing it with the weapon's "probability of kill (PK)" (e.g., 0.7), instead of calculating the full dynamics and behaviour of the weapon. While it may degrade the fidelity of the model, such approximation, as discussed above, is beneficial for reducing the computational load of the simulation, particularly for high aggregation level model

with a large number of entities. This approach contrasts with many low-level engineering problems, where deterministic models (e.g., Computational Fluid Dynamics or Finite Element Method) are constructed to replicate the sole “true” phenomenon and deviation from it can be considered as errors. Expecting such output variabilities within a model, investigation of the input-output relationships require data from multiple runs and a statistical processing of the data.

2.6 Critical Discussion

As discussed in Section 2.4, the researcher has investigated the characteristics of abstract concepts such as *capability* and *capability development* and their relation to systems design. Then state-of-the-art research on system design and evolution planning was conducted to identify potential research gaps. It was found that the methodology proposed by Biltgen (Biltgen, 2007; Biltgen & Mavris, 2007), as well as the recently developed EBD technique (Allison & Kolonay, 2018), are comprehensive and appropriate for identifying design alternatives of systems and systems-of-systems which are effective in single mission. Also, their focus is on high-level mission outcome, which is indicated by MoEs rather than MoPs. This is in line with the concept of capability, which is discussed in this chapter. However, the approaches are limited to single mission level and lack the explicit consideration of needs of capability acquirers. Such consideration is usually essential for systems development, especially for multi-role systems, which contribute to realisation of multiple capabilities. This also requires accounting for the capability planning activity, where the priorities of different capabilities are determined in a top-down manner.

It was found that dynamic design exploration methods, such as EEA, also require expansion to make them applicable to capability development. Indeed, they provide a framework for value-based evaluation of evolving system in a dynamic context by introducing context space, needs space, value space, performance model, and value model. However, the appropriate measures for the performance and the value are not obvious in themselves and therefore need to be defined, considering the practice of capability planning. This highlights the need for a new framework bridging two areas; the first one is mission analysis, pertaining to

objective analytical methods. The other is definition and expression of stakeholder needs, which are determined by their subjective value perception and influence their decisions.

Therefore, the initial objective (Chapter 1) of this research is defined here more precisely as:

Objective 1: Devise an analytical framework which aids systematic decisions regarding the evolution of capability components over a long planning horizon in a dynamic environment.

Also, when using the framework in practice, a specific method for the definition and expression of needs is needed. In value-based analysis, quantitative expression of needs is made through weighing importance of different capabilities. However, considering the capability planning process practiced in actual organisations, these quantitative data are not necessarily available to the analysts. They may be defined only vaguely or qualitatively, or they may not be explicitly expressed, or they may not be disclosed due to security reasons. Furthermore, for what-if analysis, analysts need to hypothesise the future needs for use in each scenario. It is likely that analyses need to be done in absence of information or direction from the real acquirer. Therefore, a method to interactively translate qualitative information of the importance of different capability needs into a quantitative, formalised form was found to be necessary.

This justifies the formulation of a second research objective:

Objective 2: Develop an interactive method for expressing the needs of the capability acquirer in relation to value.

Additionally, the literature review revealed a need for a method for analysing the outcome of missions, which are represented by a given set of MoEs. As discussed in Section 2.5, there exist modelling techniques grounded on empirical data-driven approaches. To collect and process data, statistical methods based on the technique of design and analysis of experiments are commonly employed. They help to understand the trade-off relationships between the factors related to system design and mission context. These can include interactions between

multiple factors, which represent emergent behaviours and thus are suitable for systems and SoS. Examples of such approaches are presented in (Kujawski, 2014; Telford, 2007).

An alternative to these approaches is the use of SEM (Structural Equation Modelling). In SEM the user can specify causal relationships as a hypothesis and verify it against empirical data. Additionally, the hypothetical causal model can contain latent constructs, which are quantities not measurable directly. Regarding system design, (M. Guenov, 2008) proposed using SEM for customer needs definition, where the needs are not clearly expressed and thus established using elicitation (i.e., market survey). A method more closely related to this research was proposed in (Huang et al., 2021), where quantitative relationships among different individual military capabilities and their aggregated capability are estimated from empirical data which contain indicator values of these capabilities. In the terminology of this research, indicators of individual capabilities can be assumed to be MoPs, while those of aggregated capabilities can be assumed to be MoEs. On the other hand, their research lacks consideration into how capability realisation is influenced by differences in factors regarding design and mission context. In their SEM model, these factors were not included. Therefore, it is concluded that if the design and contextual factors have influence on MoPs and MoEs, then that SEM technique can potentially be used for analysing trade-off relationships affecting the mission outcome.

Therefore, a third objective was formulated as follows:

Objective 3: Develop a method for trade-off analysis regarding Sol's design, which accounts for mission context.

2.7 Research Scope

As stated in the objective and discussed in Section 1.2, an attempt was made to formulate a generic framework which provides characterisation of the acquirer's needs, design features of Sol, and contextual factors of Sol. The framework also established the end-to-end quantitative relationships spanning these characterised factors, mission outcomes, realised capability levels and evaluation

metric for Sol alternatives with a chain of mission models and a value model. Specific methods to specify the models to be used in the framework were also proposed in response to Objectives 2 and 3. The solution to Objective 2 pertains to the value model, while the solution to Objective 3 pertains to the mission models.

This framework was intended to enable quantitative investigation of the “benefit side” of evolving systems. As mentioned in the previous section, the other “cost side” is virtually inseparable from the benefit side for practical decision makings, as the important index for decision makers is usually cost-to-benefit. However, since the required cost is highly dependent on the nature of the system whose alternatives are to be explored, its exploration needs to be done with a solution-specific cost model. Modelling methods for this are being developed for different systems in separate research which enable investigating system evolvability (e.g., (van Heerden, 2018; Van Heerden et al., 2018)). Therefore, cost modelling was excluded from the scope of this research, which was intended to be generic.

In line with EEA, this framework enables considering future uncertainties through planning multiple scenarios, which are assumed to collectively capture the range of possibilities. Since scenarios themselves are deterministic, the analysis is on a “what-if” basis within the scope of this research. This contrasts with formal probabilistic approaches, where every uncertainty is assigned with probability. Theoretically, probabilistic approaches may enable identification of the “best” options, or design optimisation. However, it is generally difficult to evaluate the probabilities regarding future, and unless such evaluation is reasonably done, the results shall turn out to be groundless. On the other hand, the scenario-based approach is still capable of providing information and facilitates preparation for the uncertain future. Since the aim of the proposed framework was to provide information to aid decision making, rather than making decisions by itself or optimising design, assignment of probabilities was kept beyond the scope of this research.

One of the assumptions behind the proposed framework is the ability to quantitatively measure the capability (i.e., as a level between 0 and 1). In each

mission, fulfilment is expected to be indicated by a set of MoEs corresponding to different aspects of the mission and their aggregation is expected to represent the overall capability level. However, the specific aggregation function which maps the MoEs to the capability level is likely to be dependent on the mission and the acquirer preference and it was deemed unrealistic to provide a generic method to specify it. Therefore, the development of such a method was excluded from the scope of this research.

Lastly, the research presented in this thesis was focused on the development of an analytical framework. The managerial aspect to implement it to the practical problem in a practicing organisation was therefore not covered. Since this framework involves information generated in different practices and processes it for decision making, it is crucial to ensure the coherence throughout the process to assure the traceability and integrity of the analysis.

2.8 Summary and Conclusions

The literature review presented in this chapter aimed to identify the areas for further investigation, necessary for achieving the overall research aim defined in Chapter 1. In particular,

- The method for evolution planning, the existing framework of EEA, which provides a scenario planning-based approach for modelling future uncertainties surrounding systems, was found to be generic enough to allow expanding its applicability to capability components.
- The VCDM and their underpinning principle of VFT, which uses a pre-defined value metric to represent decision maker's preference, was found to be applicable to capability-based assessment. However, identified was a need for adaptation. Specifically, characterisation of exogenous factors surrounding the Sol is necessary for applying EEA. Also, modelling of capability needs, quantification of capability levels, and definition of value metric are necessary for applying VCDM.

Despite that the above discussed concepts are interrelated, it was found that they were considered exclusively within narrower fragmented areas such as capability planning, operational analysis, and systems

engineering. It is therefore envisaged that a framework brings them together. This conjecture led to the refined definition in Chapter 1.

- Value model and performance model are input-output models in VCDM whose form needs to be known to carry out analyses. The value model needs to reflect the decision maker's needs (capability acquirer's needs in this case) and thus an interactive method to specify it was deemed necessary. Also, within the performance model, trade-off relationships among the factors of Sol's design and Sol's context exist and a method to analyse them was deemed necessary. Accordingly, research objectives 2 and 3 were formulated as complementing Objective 1.

Presented in the following three chapters are the methods and tools developed in order to achieve the above specified objectives, starting with the framework definition.

3 Framework Definition

The analytical framework formulated in response to Research Objective 1 is described in this chapter. The aim of the framework is to produce information useful for a systematic decision making, regarding the acquisition of evolving capability components. From the literature review described in the previous chapter, quantitative evaluation of alternatives was deemed suitable for this purpose. Existing approaches and methods such as Value-Focused Thinking (VFT) (Keeney, 2008), Value-Centric Design Methodology (VCDM) (Brown et al., 2009), and Epoch-Era Analysis (EEA) (Ross & Rhodes, 2008) were identified as the candidate enablers. On the other hand, as these methodologies have not yet been applied to capability components and capability systems, the need for adaptation was recognised. Specifically, an appropriate value definition for capability components needed to be formulated. Also, the associated influential factors needed to be identified and quantitative relationships among them needed to be specified. All of these and other necessary terms described below had to form a consistent ontology, underpinning the framework.

The chapter is structured as follows. The ontology is formulated in Section 3.1. The quantitative relationships of the framework are devised in Section 3.2. Assumptions made for this purpose are also discussed in this section. Finally, conclusions are drawn in Section 3.3.

3.1 Capability Planning Ontology

According to the *Oxford Advanced Learner's Dictionary*, ontology refers to “a list of concepts and categories in a subject area that shows the relationships between them” (Lea & Bradbery, 2020). Since this research attempts to utilise concepts from different areas of practice, a consistent ontology needs to be formulated as the foundation of the capability planning framework proposed in this research.

First, as stated in Section 2.4.1, the following definition was adopted from the defence domain.

Capability: The ability to complete a task under specified conditions and level of performance. (Adopted from (OUSD(R&E) USDoD, 2020))

As seen in this definition, a capability can only be defined under an existence of a particular task. A task is defined as follows:

Task: A clearly defined action or activity. (Adopted from (United States Department of Defense, 2013), as cited in (OUSD(R&E) USDoD, 2020))

The entity which possesses a capability to perform a task, or a solution for the task, can be viewed as a system. Termed capability system, it can be constituted by people, processes, equipment, and organisations.

Capability System: A system which performs a task.

Capability, task, and capability systems form a triad, where the tasks are “what” to be done and capability systems are “how” to do it. In the original definition in the reference (OUSD(R&E) USDoD, 2020), the assignment relationship of a task to an individual or an organisation (“who”) is also mentioned. This is done from the commanding perspective of defence organisations. However, in this framework, whose purpose is development of systems, this aspect is not considered and thus omitted from the framework definition.

The relationships among these three concepts are defined as:

- A capability system has a capability.
- A capability system performs a task.

Regarding tasks, composition-decomposition relationships can be considered. By examining “What subtasks must be performed in order to perform this task?”, “children” tasks can be derived from a “parent” task. Therefore,

- A task can be composed of other tasks.

This characteristic implies the existence of a hierarchy of tasks.

Capability systems have the following characteristic:

- A capability system is constituted by one or multiple **capability components**.

Adopting the definition in (Henshaw et al., 2011), capability component is defined as follows:

Capability Component: A system required to build and realize a capability (Henshaw et al., 2011).

Capability components include people, processes, equipment, and organisations, which are commonly referred to as TEPID OIL and DOTMLPF in defence organisations of UK and the US, respectively (Taliaferro et al., 2019).

Also, the following characteristic applies:

- A capability component can be a part of one or more capability systems.

Task fulfilment can be affected by entities outside the capability system, which are collectively termed as its **operational environment**, defined as follows:

Operational Environment (of a Capability System): Entities outside the system boundary of a capability system.

In the definitions in this research, capabilities are always associated with task fulfilments, which involve interaction with the operational environment. Also, capabilities have a latent feature (Yue & Henshaw, 2009). That is, it is not possible to directly measure capabilities based on the characteristics of capability systems. Instead, a manifest **effect** created through the task fulfilment under specific conditions can be investigated and measured as the proxy.

Effect: A set of changes in the world which is created as the consequence of an activity.

This effect can have an impact on the capability system itself as well as on its operational environment. For example, losses of friendly assets, as well as neutralisation of intruding air assets, can be parts of the effect of an air defence task.

- An effect is created as a capability system performs a task.

- Capability system's operational environment influences the effect.

As stated in the definition, the capability of a capability system to perform a task implies a certain level of performance, which can be referred to as **capability level** of a capability system.

Capability Level: The level of performance achieved when a capability system performs a task, as assessed by the acquirer.

The determination of a capability level involves both objective measurement of the effect and subjective assessment by the acquirer. The measurements used here should be Measures of Effectiveness (MoEs) rather than Measures of Performance (MoPs), as they are mission metrics defined from the acquirer's perspective and in a solution-independent manner. In the previous example of air defence, two MoEs, which are enemy losses and friendly losses, can be used. On the other hand, how to weigh and aggregate each measurement depends on the acquirer's preference. In one situation, they may prioritise maximising the enemy losses while accepting friendly losses to some degree. In another situation, they may wish to minimise friendly losses, while accepting fewer enemy losses. Therefore,

- A capability level is determined as an effect is assessed against need.
- A capability level indicates the achieved capability.
- A capability system is assessed according to the capability level it achieves.

Assuming that the acquisition is "capability-based" (i.e., the acquirer is indifferent between solutions which can realise the same levels of capability), capabilities can be viewed as the expression of the needs rather than particular solutions (capability systems).

- The acquirer's needs are expressed in terms of capabilities.
- The acquirer prioritises tasks in accordance with (his/her) needs.

Also, as the acquirer needs multiple capabilities corresponding to different tasks, priorities regarding tasks is expected to exist as a part of the acquirer's needs:

Task Priority: The prioritisation of one task over another when multiple tasks exist.

Finally, to express the time-dependent nature of the subject of the analysis, the concept of time period was included in the ontology. It gives a particular set of conditions exogenous to the capability system. Alternatives of a capability system are considered for each time period. The following premises are made:

- A need remains constant during a specified time period (timeframe).
- One or multiple alternatives of capability system belongs to a time period.
- An operational environment of the capability system belongs to a time period.

Reflecting the relationships described in this section, the overall ontology is expressed as the form of the UML class diagram in Figure 9.

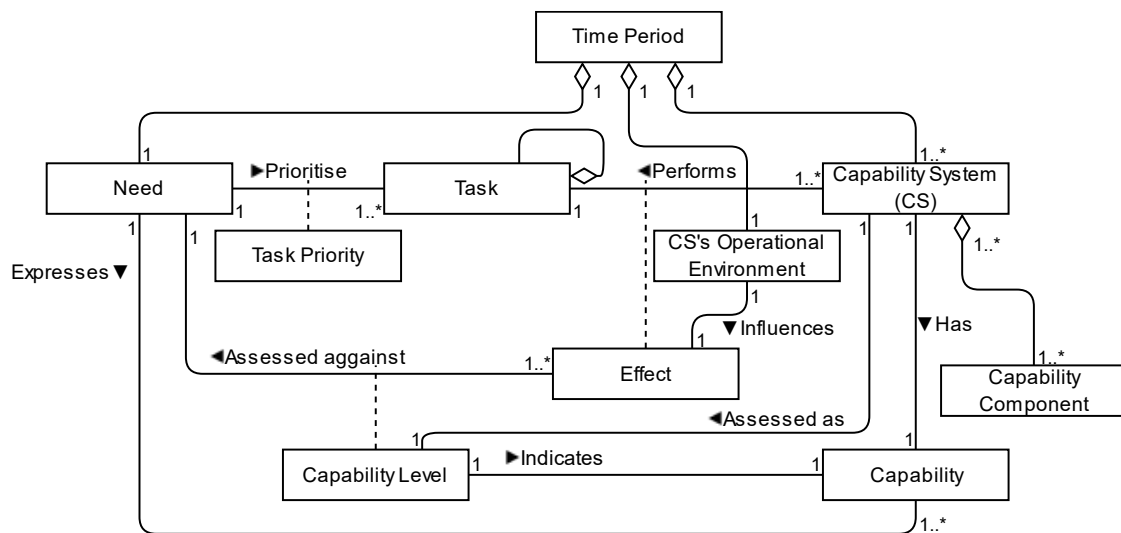


Figure 9 Ontology Model as a UML Class Diagram

The above definition applies to tasks of any scale. That is, these relationships are valid regardless of the tier of the task within the task hierarchy.

In some cases, an entire organisation (e.g., armed forces) is built around its fundamental task (e.g., “national defence”). Such an organisation can be viewed as a solution in itself to perform the task. These fundamental task, termed **organisational task**, is defined as follows:

Organisational Task: A task to which an organisation is assigned.
--

It is the “root” task in a task hierarchy, from which all other tasks (e.g., air defence) are derived as its subtasks by task decomposition. Also correspondingly, **organisational capability** and **organisational capability level** can be considered.

3.2 Framework Formalisation

Based on the ontology described in the previous section, a purposeful framework is formulated. Quantification of concepts is introduced, as well as mathematical relations linking these concepts. These are formulated to enable quantitative analysis for capability evolution planning.

3.2.1 System Boundaries Definition

The proposed framework aids the assessment of different alternatives of particular set of capability components, whose design is assumed to be under the control throughout the timeframe of concern. Adopting a term from the ISO standard of system life cycle processes, these capability components are collectively referred to as the **system-of-interest (Sol)**, which is defined as:

System-of-Interest: System whose life cycle is under consideration (ISO/IEC/IEEE, 2023)
--

The Sol in this framework is involved in performing one or more tasks as constituting part of corresponding capability systems. It can contain non-materiel components, but one can still consider their different “designs”. For example, the same fighter aircraft with personnel with two different skill levels can be counted as two different systems (human-machine systems), and the skill level can be differentiated by a decision to put more / less resources (e.g., investment in personnel training).

On the other hand, entities exogenous to the Sol are considered to be beyond the control, and their variations are treated as uncertainties imposed on the Sol.

From the viewpoint of the Sol, these entities are collectively referred to as the **(Sol's) context** defined as follows:

Context: The environment in which the system operates (Ricci, 2014, p. 78)

The boundary separating a Sol from its context is defined according to the purpose of the analysis to be carried out. If the analysis is regarding the organisation-wide capability planning, all capability components within the organisation shall be the Sol. If it is regarding the design exploration of particular systems, the boundary of these systems should be the boundary of Sol. It should be emphasised that in such case, capability components which cooperate with the Sol within capability systems are considered to be beyond the control and classified as a part of the context. This reflects the managerial independence between the constituent systems, which is a commonly observed characteristic of SoS (or capability systems).

In summary, difference between an Sol and a capability system is that the former is defined for the purpose of an analysis, while the latter is determined for each task. Set-theoretic relationship between the Sol, capability systems, and the outer world is illustrated in Figure 10. The boundaries of capability systems vary across missions, and their operational environments are shown as the yellow areas. The blue areas, together with the yellow areas, represent the context of Sols, which was defined earlier.

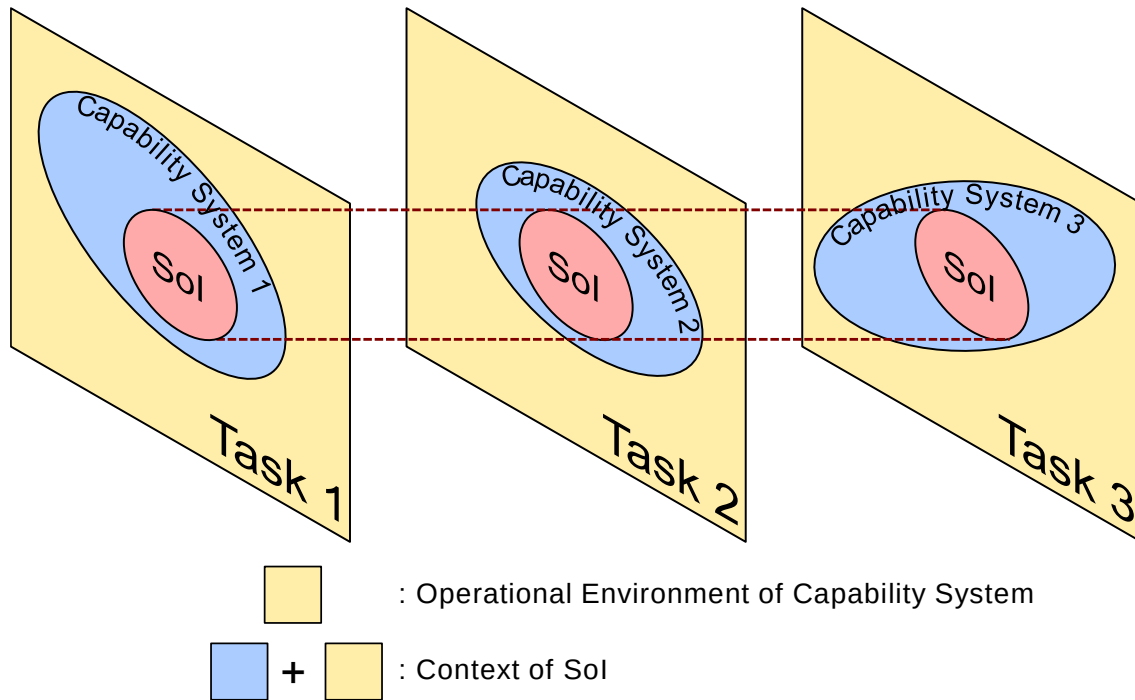


Figure 10 Boundaries of Sol and Capability Systems

It should be noted that in the original ISO standard, an additional classification of “enabling systems”, is present. It is defined as “system that supports a system-of-interest during its life cycle stages but does not necessarily contribute directly to its function during operation”, which addresses stakeholder concerns by forming a solution alongside Sol (ISO/IEC/IEEE, 2023). This definition appears to be similar to that of capability system in the proposed ontology. This is in a sense that some systems exogenous to Sol may involve solutions, and that the scope of solutions varies according to the situation. This relationship is illustrated in Figure 11.

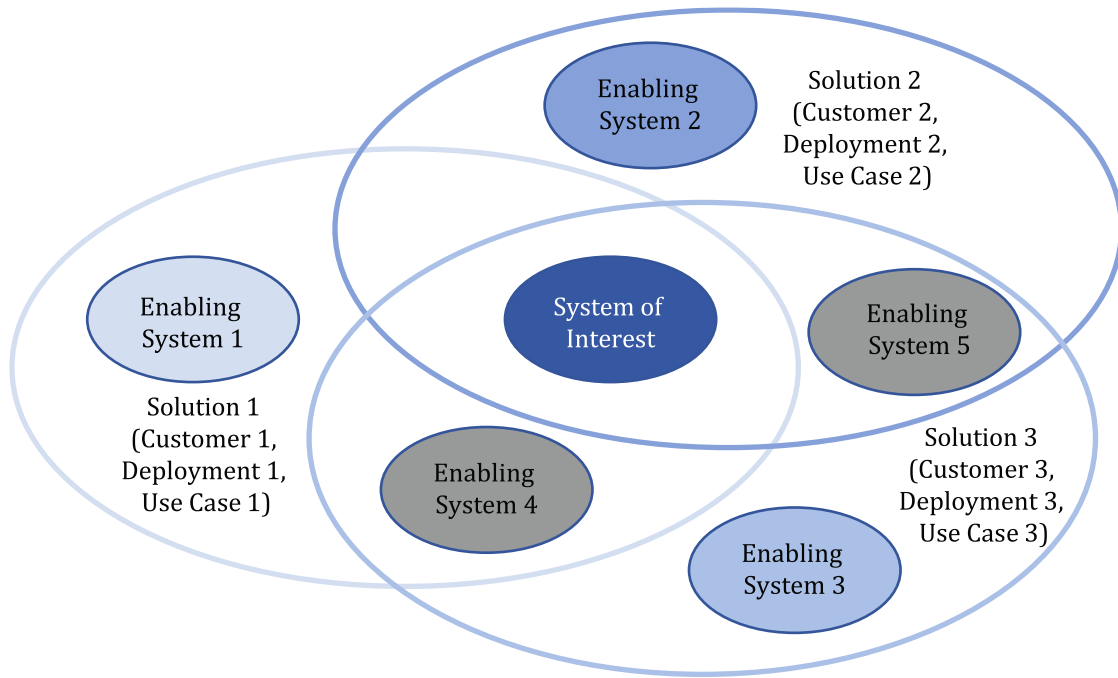


Figure 11 System Solution Contexts (ISO/IEC/IEEE, 2023)

However, these views are not entirely the same. The ISO/IEC/IEEE view (Figure 11) seems to imply that operational performances are largely realised by the Sol alone, and thus enabling systems are treated as “supporting” system. The proposed ontology, in contrast, assumes that performances are exhibited through an integration of Sol and other systems. Instead, ISO/IEC/IEEE proposes separate classifications of “interfacing systems” and “interoperating systems”.

In any case, this ontology is focused on the operational phase of the life cycle (which is termed “In-Service” in CADMID cycle (UK Ministry of Defence, 2005) of UK defence acquisition), and consideration of systems which are involved in other phases (e.g., “Manufacture” phase in CADMID) is not envisioned. For example, while production capacity of entire aircraft involves in the “Manufacture” phase, it is not likely to be a part of a capability system, as it does not affect the operation of existing examples of the type. On the other hand, production capacity of certain consumable parts within the aircraft can be a part of “logistics” component of a capability, as it can affect the “In-Service” phase.

3.2.2 Capability Levels

From the perspective of the acquirer, who is responsible for the development of the entire organisation, the following assumption can be made:

Assumption: Acquirer (customer) needs can be identified with the capability to perform their organisational tasks.

Given that a capability level can be quantified, a higher organisational capability level can be assumed to be the acquirer's needs in the first place. As it will be discussed later, a linkage between this fundamental need and design features of individual capability components is established in the proposed framework.

Capability levels were defined to be expressed with numbers between 0 and 1, where:

0: Lowest (least desirable) acceptable level

1: Highest (most desirable) level.

By definition, a capability level below 0 means that the realised capability for such a task is unacceptable to the acquirer.

As described in the ontology, this quantity is determined by assessing the effect created through task fulfilments. As this effect is expressed as a set of MoEs, the following assumption can be made:

Assumption: Complex mission achievement can be measured by aggregating measures of effectiveness.

This aggregation needs to reflect the intention of the acquirer. Specifically, which MoEs are to be used, and how they are aggregated need to be specified. If a mapping from MoEs to capability level can be explicitly expressed, this process can be done systematically. Generally, some elicitation from stakeholders is needed for identification of this mapping, although specific methodology for this was not a subject of this research.

Next, the relationship between the composition-decomposition relationships of tasks and capability levels is considered. For this, the following assumption was made:

Assumption: A set of representative subtasks can be appropriately derived from a supertask so that the subtasks' fulfilment collectively indicates the fulfilment of the supertask.

For example, from a supertask of "defending an island", subtasks to counter maritime and airborne threats may be easily derived. However, various coordination efforts between these two tasks may be required for successful accomplishment of the supertask. If they are not deemed negligible, they must be included as additional subtasks. Universal Joint Task List (UJTL) of the United States Armed Forces (JCS USDoD, n.d.) is an example of a complete taxonomy of tasks of an organisation.

Also, the applicability of this assumption is generally limited to tasks which can be assumed to be independent from each other. In such a case, the fulfilment of one subtask is not affected by that of another subtask. By practitioners in military CBP domain (for example, (Davis, 2002, p. 27)), the feasibility and practicality of defining tasks so that they can be treated as *largely*, if not completely, independent from each other is supported. However, in a strict sense, there can be interrelationships between task which might not be negligible. In such a case, the overarching supertask needs to be considered as a whole.

Additionally, the following assumption is made:

Assumption: The fulfilment of a supertask can be evaluated based on the fulfilment of its subtasks and their importance over each other.

In theory, in order to derive the capability level of the supertask from the capability levels of its subtasks, the impacts of subtask fulfilments on supertask fulfilment need to be known. However, obtaining this information by analytical methods (e.g., modelling and simulation) can be difficult. This is especially the case when the task in question is at the high-level of the hierarchical structure, with a broad scope and complex capability solutions involving many systems and capability

components. Instead, with the above assumption, this process can be simplified by substituting such complex (and often unreliable) analytics with importance of subtasks, which can be subjectively determined by the acquirer. The acquirer is likely to have accumulated knowledge regarding the task importance through experiences, (e.g., strategic-level wargames), therefore this assumption was considered to be reasonable and practical.

With the relationships among the capability levels of supertask and subtasks known, the entire hierarchy of tasks can be traversed. That is, organisational capability level can be expressed as the aggregation of capability levels of all subordinate tasks. Higher value of this organisational capability level can be considered as the fundamental needs of the capability acquirer, as mentioned in Section 3.1. It can serve as the indicator of the needs satisfaction, which can be the “objective function” for the individual system design. Perhaps this can be better understood, with the help of an example: Consider the organisational task, “national defence”. The Sol is a fighter aircraft. Finding an alternative of fighter aircraft which can realise the best national defence capability is the goal for the design activity. However, it is hardly possible to directly link individual design features of fighter aircraft to national defence. As systems usually contribute to different parts of the task in different ways and to different extents, “divide and conquer” approach using task breakdown is likely to be necessary.

Generally, as tasks are broken down further, solutions to fulfil them become smaller, and analysis into the relationship between solution and capability level becomes simpler. Hence, task breakdown helps analysing the contribution of Sol, which is discussed later.

However, there should be a limit on the level of decomposition, considering the need to facilitate exploration of a wide range of solutions by retaining some freedom regarding their form. This is because a too elementary (leaf) task presupposes a particular form of solution. For example, if one decomposes air defence task further into detection task and engagement task, barrage balloons, which do not possess nor require detection capability, may not be considered as a (potential contributor to a) solution.

Additionally, separately considering too elementary tasks leads to a failure to account for interactions between systems and to evaluate the effect of capability systems integration. In the example above, it is highly unlikely that good detection system (e.g., radar) and good engagement system (e.g., missiles) immediately mean the success of air defence. Effects of other factors, such as interfaces and communication, need to be captured as well for a correct assessment. However, it is very difficult to do so if tasks are excessively fragmented.

Further in this thesis, tasks at an appropriate resolution will be referred to as **missions**, which form the lowest level of the breakdown structure of the tasks in this framework.

Since the “parent” task is performed by the aggregation of solutions performing its “children” tasks, the capability level of the “parent” task is given by a weighted average of those of the “children” tasks. Here, the weights can be understood as the relative importance of one “child” task over another in performing the “parent” task (**importance weight**). In the previous example, if the military has solutions for both “countersea” and “counterair” tasks, one can assume that they have a solution for the overarching “island defence” task. If the countersea capability level is 0.8 and the counterair capability level is 0.4, the total island defence capability level should be somewhere between 0.8 and 0.4, determined by relative importance of the countersea task over the counter air (or vice versa). These relationships are applicable to the mission capabilities and the organisational capabilities. When the capability level of mission i is V_i , its importance weight is W_i , and the number of missions which compose the entire organisational tasks is N , the organisational capability level V_o can be expressed as:

$$V_o = \sum_{i=1}^N W_i V_i \quad (3-1)$$

Where $0 \leq W_i \leq 1$ and $\sum_{i=1}^N W_i = 1$. This relationship is illustrated in Figure 12. The organisational capability level V_o is visualised as the total area of grey rectangles.

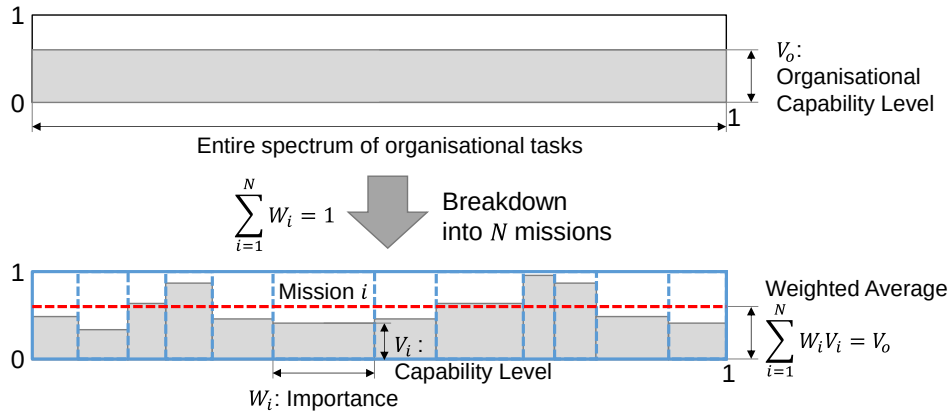


Figure 12 Breakdown of Organisational Tasks into Missions

3.2.3 Contribution of Sol to Organisational Capability

In the previous section, an assumption was introduced so that acquirer needs can be identified with the capability to perform their organisational tasks. On the other hand, what needs to be known from the viewpoint of individual systems is how the differences in their respective design impact the fulfilment level of organisational tasks. This can be referred to as the contribution an Sol makes to the organisational capability, leading to the following assumption:

Assumption: The acquirer (customer) can compare solution alternatives based on their contribution to the fulfilment of the organisational tasks.

An Sol does not necessarily contribute to all missions within the entire spectrum of organisational tasks. Relevant missions, whose capability levels can be affected by the Sol, constitute a set M . For each mission in M , different alternatives of Sol yield different capability levels. Contribution of an alternative X to the organisational capability can be quantified by extracting relevant terms from equation (3-1) as follows:

$$C(X) = \sum_{i \in M} W_i V_i(X) \quad (3-2)$$

As M can be viewed as the undertakable mission spectrum of the Sol, alternatives can be compared within this limited spectrum. Importance weights W_i of missions are upscaled to w_i (**importance weight for Sol**) by multiplying a

constant $1/\sum_{i \in M} W_i$, so that they sum to 1. The relative importance of one mission over another, which is given by a ratio of W_i to W_j , remains the same with w_i to w_j . This spectrum is useful for comparing alternatives of Sol, as the complete task breakdown structure is not necessarily available for analysis in practice. This may be due to the vastness of the entire spectrum of organisational tasks.

Within this Sol mission spectrum, the relative contribution of an alternative X to the organisational capability, $c(X)$, can be defined as:

$$c(X) = \sum_{i=1}^n w_i v_i(X) \quad (3-3)$$

This is because $c(X) = C(X)/\sum_{i \in M} W_i$ and $c(X)/c(Y) = C(X)/C(Y)$ always hold, for any X and Y . This allows using $c(X)$ to compare contributions made by different Sol alternatives.

An example is given in Figure 13. The contribution C and relative contribution c are visualised as total areas of rectangles, each of which representing a mission. In this case, $N = 12$, $n = 3$ and $M = \{4, 6, 9\}$. In the Sol mission spectrum, these missions are re-numbered into $\{1, 2, 3\}$: this means that $v_1 = V_4$, $v_2 = V_6$, $v_3 = V_9$.

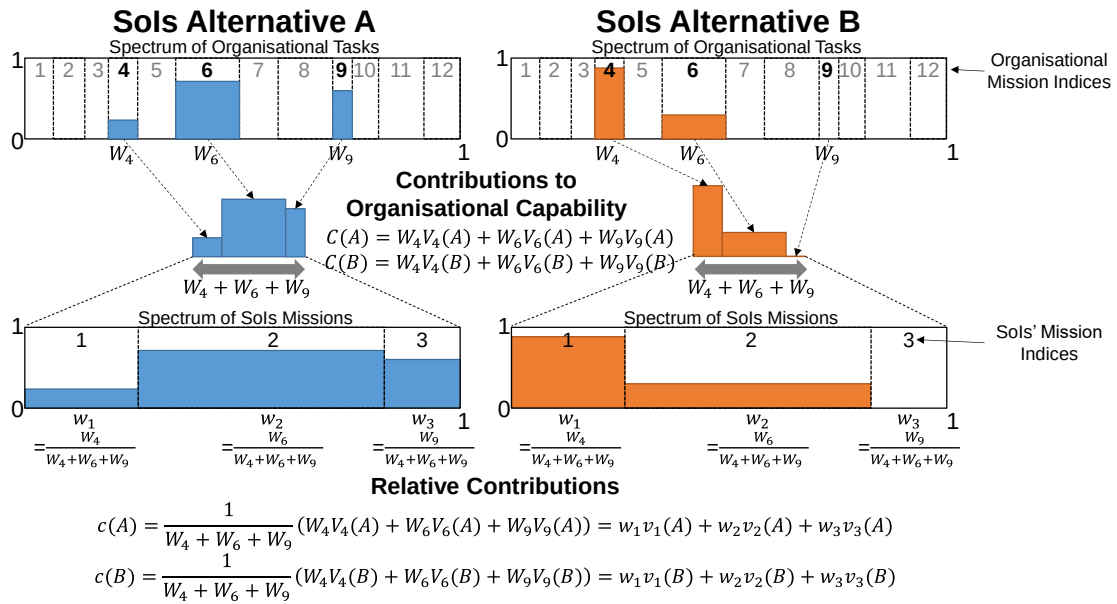


Figure 13 Calculating Relative Contribution to Organisational Capability

The proposed framework, on its own, does not specify elimination criteria for solutions with unacceptable capability levels. Even if achieved levels of some of the tasks fall below 0, high levels in other tasks can still compensate for the “underachievers”. This is because levels of any pair of tasks within the spectrum can be traded off. In realistic situations, whether such state can indeed be acceptable shall depend on the nature of the tasks. If the tasks are non-essential “bonus” tasks, they may be retained to explore the trade-off possibilities. In contrast, if the tasks in question are essential, it is likely that their levels need to be above 0, and therefore underachieving solution alternatives are likely to be excluded from the consideration. In light of the principle of the proposed framework, which is entirely based on value trade-offs and eliminated thresholds and requirements, this may seem to be an exceptional treatment. Despite this, as required by stakeholders and analysts, such elimination criteria can be easily implemented when analyses are carried out.

3.2.4 Value Definition Based on Contribution

In line with the concept of VFT and VCDM described in Section 2.1, a 1-D metric of **value** needs to be pre-defined first such that the preferences of the acquirer organisation are reflected. Later, each alternative is evaluated in accordance with the value definition. In the previous section, it was shown that the contribution to the organisational capability, quantified as $c(X)$, expresses the benefit offered by the Sol. Therefore, $c(X)$ was suitable as the value function of the Sol. It is worth noting that the linkage between the benefit and money was not established, as the cost aspect was not considered within the scope of this research.

$$V(X) = c(X) = \sum_{i=1}^n w_i v_i(X) \quad (3-4)$$

$v_i(X)$ in equation (3-4) can be viewed as value attributes, or objectives, in the terminology of decision analysis. This form of $V(X)$ implies that:

1. Sol which can realise higher capability levels (higher $v_i(X)$) are valued more highly.

2. Sol which can contribute to important missions (those with higher w_i) are valued more highly.
3. Sol which can contribute to more missions (more non-zero $v_i(X)$) are valued more highly, as some alternatives can contribute to missions which other alternatives cannot.
4. Trade-off between mission capabilities in terms of value is allowed, at a ratio of one importance weight over another ($w_i(X)/w_j(X)$).

Additionally, by considering parent tasks at the levels in between the top-level organisational task level and the mission level in the task breakdown structure, a value tree of the Sol can be formed. Specifically, missions undertakable for Sol, together with their “ancestors” tasks, are extracted from the organisational task breakdown structure. Then, the importance weights are recalculated so that the proportional relationships of the weights between tasks which share the common “parent” task are preserved. In the tree, the sum of the weights of capabilities at the same level of the hierarchy is always 1. This implies that the value of the Sol can be decomposed in a hierarchical and “mutually independent, collectively exhaustive” manner, corresponding to the task breakdown (Figure 14). A specific method to determine importance weights in the Sol value tree is proposed in Chapter 4.

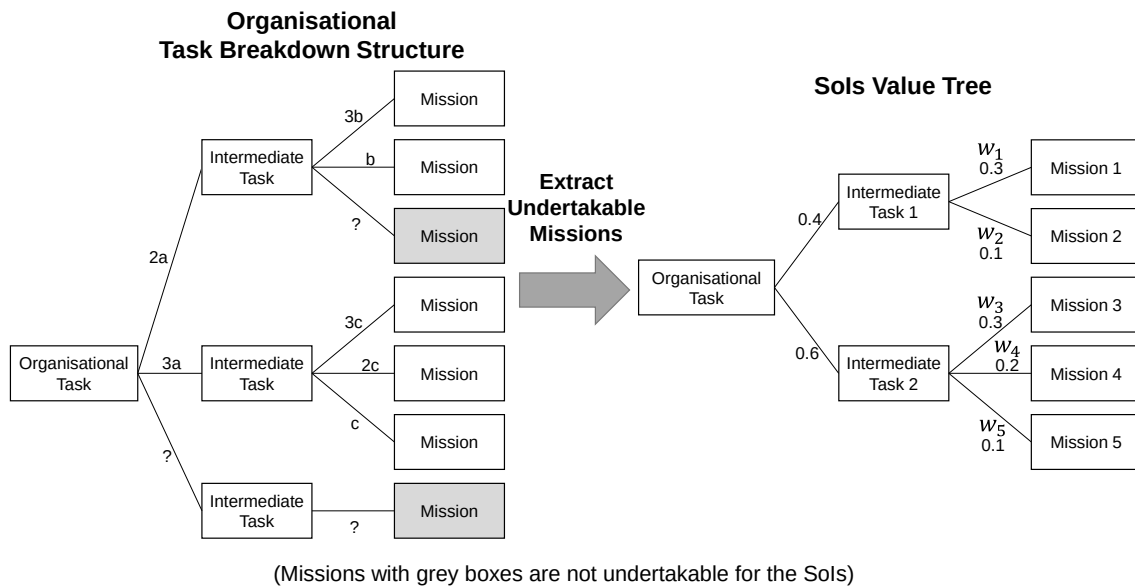


Figure 14 Formation of Sol Value Tree

Using the resulting value function, alternatives can be evaluated in relation to the acquirer needs. The function can be viewed as a mapping of the realised capability levels to value in accordance with the acquirer needs, which are characterised by the set of missions and their importance weights.

3.2.5 Value-Based Evaluation among Temporal Changes

In the previous section, an approach for value-based evaluation of capability-based systems was presented. To account for temporal changes including design evolutions, the proposed approach was expanded by incorporating the methodology of EEA (Ross & Rhodes, 2008).

The design of a Sol may evolve over time during its lifecycle. In this framework, different designs are assumed to be characterised by a finite set of **design variables**. Each design variable represents a decision which differentiates one design from another, and can be either discrete or continuous. Design variables span the **design space**. An instance of Sol design is represented by a **design vector, D** .

The factors exogenous to the Sol, which are beyond the control of designers, are also subject to changes and uncertainties. These factors can be classified into context and needs. Like design, the context of a Sol is assumed to be characterised by a finite set of **context variables**, each of which differentiates one context from another. These variables span the **context space**. An instance of context is represented by a **context vector C** . In this framework, context variables which affect the realised mission capability level are of primary concern. Factors regarding adversary systems and systems cooperating with Sol are typical examples of such context variables. The characterisation of a capability System, Sol, context and operational environment is expressed in the form of UML class diagram in Figure 15.

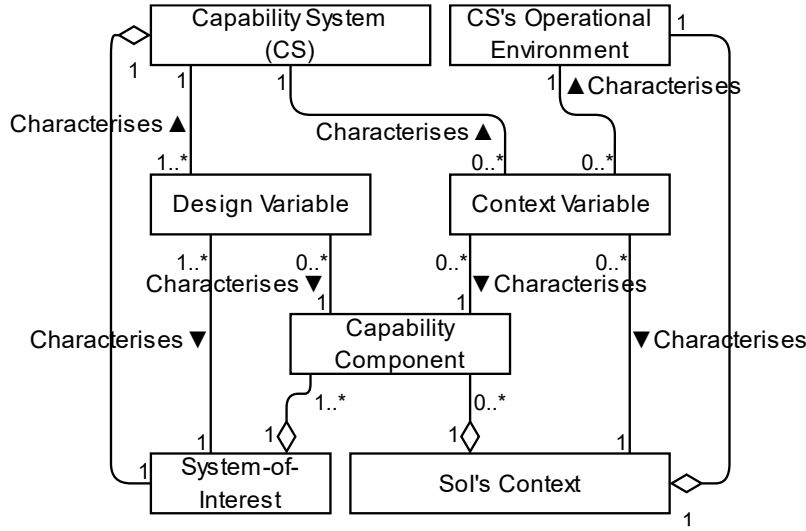


Figure 15 Characterisation of Capability System, Sol, Context and Operational Environment as a UML Class Diagram

A mission can be viewed as a function which takes a pair of Sol design and context vectors as the input, and outputs the realised capability level. Mathematically, the function $f_{mission\ i}$ is expressed as follows, using $\mathbf{D}$, $\mathbf{C}$ and v_i :

$$v_i = f_{mission\ i}(\mathbf{D}, \mathbf{C}) \quad (3-5)$$

This $f_{mission\ i}$ can be viewed as a composite function (procedure) of an objective part and a subjective part. The former outputs the MoEs indicating the objectively measurable effect created through mission execution and can be expressed as a function $f_{effectiveness\ i}$. The latter takes these MoEs as the input, aggregates them to output eventual v_i . This can be expressed as a function $f_{success\ i}$, as it pertains to the mission successfulness as perceived by the stakeholder. For example, when there are two MoEs which are enemy loss and friendly loss, one armed force may prioritise minimising the latter to maximising the former, while another may do the opposite.

In the concept of design loop (Ricci et al., 2014), the former (objective) part belongs to the performance model, while the latter belongs to the value model, because of the involvement of subjective preference. In a mathematical formulation, the following equations can be defined using the vector of mission Measures of Effectiveness (MoEs) $\mathbf{E}$.

$$E = f_{effectiveness\ i}(D, C) \quad (3-6)$$

$$v_i = f_{success\ i}(E) \quad (3-7)$$

$$v_i = f_{success\ i}(f_{effectiveness\ i}(D, C)) \quad (3-8)$$

While appropriately defining $f_{success\ i}$ reflecting the acquirer preference could be a problem worth pursuing, this was not included in this research. Instead, $f_{success\ i}$ was assumed to be pre-defined beforehand and available for use in the framework. On the other hand, a specific method to analyse $f_{effectiveness\ i}$ is discussed in later Chapter 5.

Here, it is assumed that the form of the functions is invariant over time. If there is a significant change in the mission itself and applying the same function becomes inappropriate, introduction of a new mission is required. For example, ballistic missile defence can be called as a subset of air defence with higher target performance, which could be characterised with slightly different C . However, many unique elements, such as space-based sensors, make a design space considerably different from that of ordinary air defence mission. In such a case, considering separate mission model will make sense.

Factors related to the value perception of the acquirer organisation are classified as a part of the needs. It is characterised by the value function, specifically the set of missions and their importance weights. In the form of mathematical expression, what changes over time is the function f_{value} , which maps each mission capability level v_i onto the value metric. Their relationship is expressed as:

$$V = f_{value}(v) \quad (3-9)$$

Assuming the additive value function introduced in the previous section, f_{value} can be parametrised with a vector w of mission importance weights w_i . Vector w can be viewed as a **needs vector** defined in the **needs space**, therefore w_i can be referred to as **needs variables**. With all possible missions the Sol may be involved in performing in the future, the needs space can be defined as a

hypercube. With this definition, addition (w_i becomes zero to nonzero) or removal (w_i becomes nonzero to zero) of missions can be modelled, as well as changes of importance during the lifecycle. Equation (3-4) can be simply rewritten as:

$$V = \mathbf{w} \cdot \mathbf{v} \quad (3-10)$$

Changes occurring within the Sol boundary (i.e., changes of design) are expressed as transition from one ***D*** to another, while changes occurring outside the boundary (i.e., changes of context or needs) are expressed as one pair of (***C***, ***w***) to another pair.

As in EEA, an ***epoch*** is defined as a discrete time period with fixed context and needs (Ricci, 2014, p. 81). It is characterised with a context vector and needs vector (***C***, ***w***), which together specify an external condition to the Sol.

By sequencing epochs, a ***scenario*** which describes one possible future awaiting the Sol can be characterised. Likewise, an evolution trajectory of the Sol can be characterised by a sequence of its designs. Value of the Sol can be calculated for each epoch, and a value trajectory of the Sol is calculated by sequencing values along the timeline. The overall relationships among vectors, functions and values across epoch is described in Figure 16.

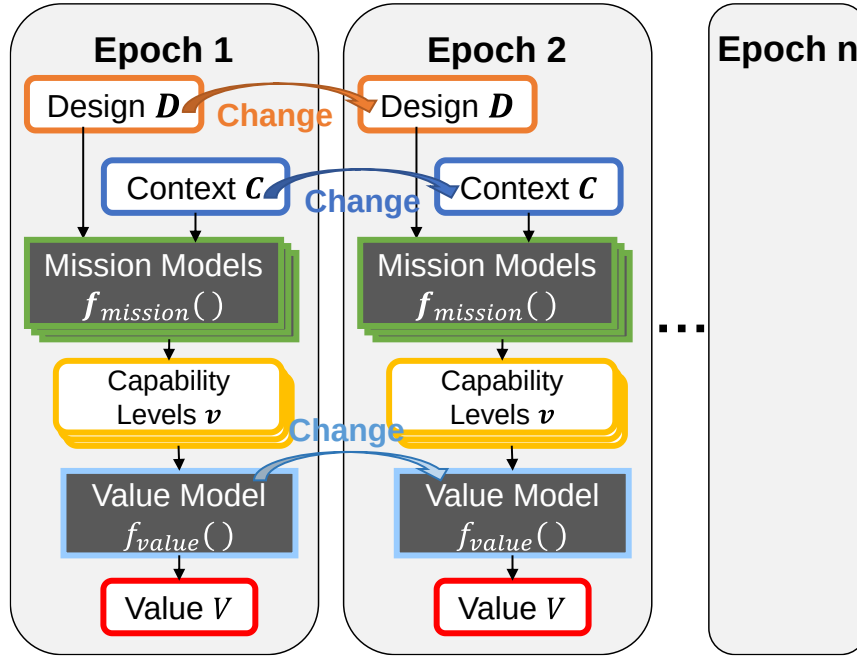


Figure 16 Temporal Representation of the Proposed Framework

As described above, the approach taken by EEA to model and analyse future uncertainties is scenario-based. The proposed framework inherits this approach from EEA, which was employed as the basis. The underpinning assumption for these scenario planning methods is as follows:

Assumption: A finite number of representative deterministic scenarios collectively capture the range of future uncertainties.

In contrast to some other uncertainty modelling approaches such as probabilistic methods, this scenario planning is characterised as “less formal but more practical” (O. de Weck et al., 2007). Since each scenario itself is deterministic, it cannot account for the likelihood of each future unless probabilities are assigned to each scenario. However, the scenario specification using context and needs vectors gives a flexibility in modelling possible futures. Context variables can be factors related to interoperating systems, adversary systems, or operational environment including conditions of missions (e.g., distance from the airbase to the area of operation). Needs variables, which are importance weights of missions, describe priorities of the acquirer at various task levels, from high strategic level to low tactical level. Since the proposed framework was intended

for the use in early conception stage of the system development, the ability to account for a wide range of future possibilities was considered to be suitable for the purpose. The characterisation of a scenario is expressed in the form of UML class diagram in Figure 17.

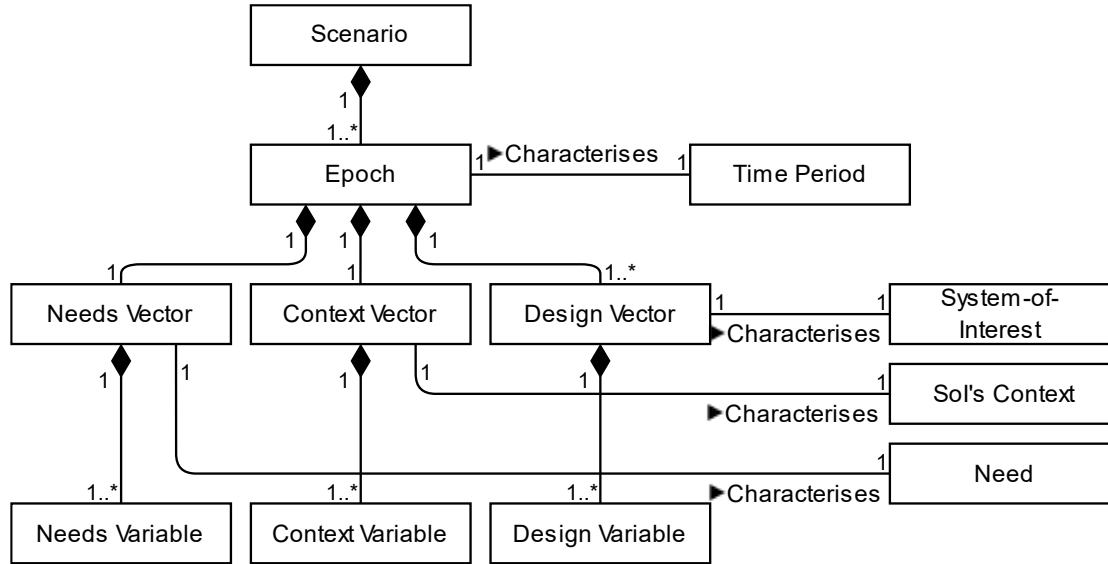


Figure 17 Scenario Characterisation as a UML Class Diagram

The proposed framework is entirely based on the principle of VFT, which first defines a value metric reflecting the acquirer's preference, then assign value to each alternative and use it for evaluation. With the value defined as a 1-D cardinal number, trade-off relationships across tasks can be specified with importance weights of tasks.

The overall inputs of this framework are a set of alternatives of design evolution trajectory and a set of scenarios. The outputs are value trajectories calculated for each combination of a design evolution trajectory and a scenario. The information produced by this process is expected to enable a value-based "what-if" analysis.

To use the proposed framework in actual analysis, w_i need to be determined to reflect the acquirer needs, and the $f_{mission i}$ need to be identified. In the following chapters, methods for these are described.

3.3 Summary and Conclusions

A framework for an exploration of evolving capability components was devised. It is founded on a consistent ontology, which was formulated by incorporating concepts from engineering system design, capability planning, and operational analysis. Assuming that the benefit to the capability acquirer is the contribution to the organisational capability, a value metric was formulated. It is an aggregation of capability levels achieved in relevant missions defined as an additive value model. In line with the principle of VFT and VCDM, comparative evaluation of different design alternatives is enabled with such a metric.

Mathematical relationships bridging this value metric and factors affecting it were established. These factors are acquirer's needs, Sol's design, and Sol's context. Needs variables are the mission importance weights in the value model, representing the acquirer's priorities across tasks. Design and context variables comprise the inputs of mission models, which output the capability levels of mission and affect the value through them.

Value-based analysis of Sol lifecycle accounting for future changes was enabled by adopting the methodology of EEA. Using the parametric expressions of needs and context, possible futures exogenous to the Sol are planned as scenarios. These are characterised as the sequences of state transitions of needs vectors and context vectors. Alternatives of evolving Sol designs, which are expressed as state transitions of design vectors, can be evaluated against these scenarios. Trajectories of Sol value are outputted as the end product of the application of the methods in framework.

Also, the proposed framework enables value-based what-if analysis of evolving Sols from the acquirer's perspective. From the practical viewpoint, it enables:

- Consideration of multi-role Sols, which contribute to the organisation's task fulfilment in multiple task (mission) areas.
- Trade-off between task performances in accordance with the acquirer's needs.
- Accounting for shifts of needs in terms of high- and low-level task priorities.

- Consideration of the impact of various contextual factors on mission accomplishments and Sol value.
- Modelling of a wide range of future possibilities regarding needs and Sol context.

With these features, the framework is expected to be capable of producing information which aids decision making in the evolution planning of capability components.

4 Capability Needs Prioritisation

4.1 Introduction

Proposed in this chapter is a method in accordance with Objective 2 of this research:

Objective 2: Develop an interactive method for expressing the needs of the capability acquirer in relation to value.

In the design loop (Ricci et al., 2014) discussed in Chapter 2, value assessment of a System-of-Interest (SoI) is not necessarily directly associated with the performance measures it achieves in tasks. Following this philosophy, a value model is proposed to explain the trade-off relationships between performance measures in terms of value. In the capability-based context, the model needs to reflect the stakeholder's capability needs, which are determined as a part of the capability planning activity.

As formalised in Section 3.2.4, an additive value function was assumed as the form of the value model. The model expresses the value as a weighted sum of capability levels. To identify it, importance weights associated with each task need to be determined reflecting the priorities of capabilities, as perceived by the stakeholders. Since such perception can be qualitative and intangible, a need for a specific clarification and quantification method was recognised.

In response to this need, a method which utilise pairwise comparisons with ratio scale is proposed. It is intended to translate the qualitative capability needs of the stakeholders, in terms of their perceived value, into quantitative trade-off ratios.

The rest of this chapter is organised as follows: In Section 4.2, details of the proposed method are described. In the subsequent Section 4.3, results from the demonstration (use case) are presented. Conclusions are drawn in the final Section, 4.4.

4.2 Pairwise Comparisons within Value Hierarchy

As the framework developed in this research (described in Chapter 3) is founded on the Value-Centric Design Methodology (VCDM), design exploration of Sol is treated as a decision making based on a value metric. Inherently, some value-based tools for decision analysis are applicable to the framework. The capability prioritisation, described in the previous section, can be associated with a specific area of methods for criteria weighting in Multi-Criteria Decision Making (MCDM); here, *criterion* is synonymous with objective in the terminology of decision analysis. It should be noted that the problem is not related to the whole process of MCDM, which usually include other activities such as evaluation of alternatives against criteria.

Criteria weighting methods determine quantitative importance weights from limited available information regarding the priorities among criteria, as perceived by the decision maker. In this research, an approach was taken to develop a capability prioritisation method by adopting a criteria weighting method. The following points regarding capability planning were considered:

- The capability acquirer is assumed to be able to express relative priorities among capabilities, by comparing one capability over another.
- The relative priorities are not necessarily expressed quantitatively.
- Comparison takes place at each level of the task hierarchy, possibly from different perspective (e.g., strategic or tactical) by different group of people.
- Comparison results change over time.

A review of criteria weighting methods can be found in (Odu, 2019). Since capabilities are prioritised subjectively by the acquirer, subjective weighting methods are suitable for this purpose. Inherently, as the methods deal with subjective judgement, there is no perfect method to create a perfect model. Also, applicability of each method depends on the availability of information elicited from the decision maker. For example, the point allocation method assumes that the decision maker can precisely express the quantified importance, which is not likely the case in this research. Swing weighting method requires alternatives to

be known a priori to determine weights, making it incompatible with Value-Focused Thinking (VFT) described in Section 2.1. In contrast, ranking method only require ordinal ranks of criteria. However, it cannot account for the intensity of importance (i.e., how much one criterion is more important than another).

Assuming that the capability acquirer can specify relative intensity of importance of tasks, if not precisely, a method based on pairwise comparisons with a ratio scale was found to be suitable for the purpose. It is a criteria weighting method employed in the Analytic Hierarchy Process (AHP) (T. L. Saaty & Vargas, 2012), which is a comprehensive MCDM method (In the AHP, pairwise comparison is also used for prioritising alternatives in terms of the criteria). The method systematically quantifies importance weights from qualitative ratings indicating the importance of one task over another. Specifically, it requires the user (the acquirer) to provide qualitative answers to a series of questions such as: “How important is the fulfilment of task Y compared to that of task Z, in order to perform their parent task X?”. The answers are to be chosen from verbal statements defined as the Fundamental Scale (Table 1). Intensities of importance are normally quantified as natural numbers or their reciprocals corresponding to the given statements. An exception is where the consistency rule (which is described later) is imposed as a constraint, making the intensities rational numbers in such a case.

Table 1 Fundamental Scale (T. L. Saaty & Vargas, 2012)

Intensity of importance	Definition	Explanation
1	Equal importance	Two activities contribute equally to the objective
2	Weak	
3	Moderate Importance	Experience and judgment slightly favor one activity over another
4	Moderate plus	

5	Strong importance	Experience and judgment strongly favor one activity over another
6	Strong plus	
7	Very strong or demonstrated importance	An activity is favoured very strongly over another; its dominance demonstrated in practice
8	Very, very strong	
9	Extreme importance	The evidence favoring one activity over another is of the highest possible order of affirmation
Reciprocals of above	If activity i has one of the above nonzero numbers assigned to it when compared with activity j , then j has the reciprocal value when compared with i	A reasonable assumption
Rationals	Ratios arising from the scale	If consistency were to be forced by obtaining n numerical values to span the matrix

By applying an algorithm to the quantified intensities corresponding the statements, quantitative weights are given systematically. The specific algorithm is as follows. Calculations proceed from the higher tier to the lower within the value hierarchy. Tasks and missions which share the same “parent” task in the value hierarchy form are grouped. First, in the group formed by the n tasks within the tier 1, a pairwise comparison matrix A is populated with a_{ij} . Each a_{ij} indicates the elicited intensity of relative importance of the i th task over the j th ($1 \leq i \leq n, 1 \leq j \leq n$). The resulting matrix A is a positive reciprocal matrix, where $a_{ij} = 1/a_{ji}$ ($i \neq j$) and $a_{ij} = 1$ ($i = j$). This implies that specification of intensities of $n(n-1)/2$ pairwise comparisons within the group identifies A .

Principal eigenvalue of A can be calculated as $\lambda_{\max}$, and its corresponding normalised eigenvector w becomes the priority vector, whose elements w_i indicating the “local” importance within the group.

$$Aw = \lambda_{\max}w \quad (4-1)$$

$$\sum_{i=1}^n w_i = 1 \quad (4-2)$$

The same process is applied to the groups of tasks at lower tiers. “Global” weights are calculated stepwise from the upper tiers to the lower, by multiplying the obtained “local” weights with the “global” weight of the “parent” task. With the “global” weights, a complete value hierarchy characterising the value model is obtained. Weights of high-level tasks affect the importance weights of lower-level capabilities. This enables modelling changing acquirer capability needs at various levels, affecting the eventual value of Sol.

Within a group with $n \geq 3$, the specification of rating may be inconsistent. This refers to the deviation from the transitivity rule across pairs. For example, if task 1 is of “moderate importance” in relation to task 2, and task 2 is of “moderate importance” in relation to task 3, task 1 should be of “extremely importance” in relation to task 3, as the intensities $a_{12} = 3$ and $a_{23} = 3$ makes $a_{13} = a_{12}a_{23} = 9$ if transitivity rule holds. However, the result of elicitation may not follow and task 3 may be of “very strong importance” in relation to task 1 ($a_{13} = 7$). Since the acquirer elicitation relies on verbal qualitative ratings, such discrepancy may occur.

As a measure for consistency, the use of the Consistency Index (CI , eq.(4-3)) is suggested in the AHP. If $CI \geq 0.1$, it is suggested to “study the problem and revise the judgments” (R. W. Saaty, 1987).

$$CI = \frac{\lambda_{\max} - n}{n - 1} \quad (4-3)$$

Theoretically, this method is applicable to groups of any size (i.e., any n). However, the need for carrying out $n(n - 1)/2$ pairwise comparisons may

become prohibitively expensive (laborious) in practice. A possible way to reduce n is to appropriately introduce supertasks containing a group of tasks.

By using pairwise comparison results as the parameters, rather than importance weights themselves, capability needs could be parametrised with a smaller number of parameters. Also, the proposed approach agrees with the relative nature of the acquirer's perception regarding importance, where one can hardly determine the absolute importance of a given task within the entire task spectrum. This enables modelling changes regarding needs in a concise and consistent manner. For example, importance of task A in comparison with that of task B may change from "equal" to "moderate" without affecting other tasks. This feature is important for scenario planning, where different scenarios are created for what-if analysis by making changes in some of the comparison results.

4.3 Demonstration Test

To demonstrate and verify the developed method, a test was carried out. As a test case, 3-tier hierarchical task structure consisting of 2 intermediate-level tasks (IT1, IT2) and 5 missions (M1, M2, M3, M4, M5) was considered. Here, M1, M2, and M3 are subtasks of IT1, while M4 and M5 are subtasks of IT2. The overall hierarchy of tasks and missions, together with the pairwise comparisons required to determine the importance weights, is illustrated in Figure 18. 5 pairwise comparisons, as indicated in the figure with numbers, are required to prioritise the entire capability needs.

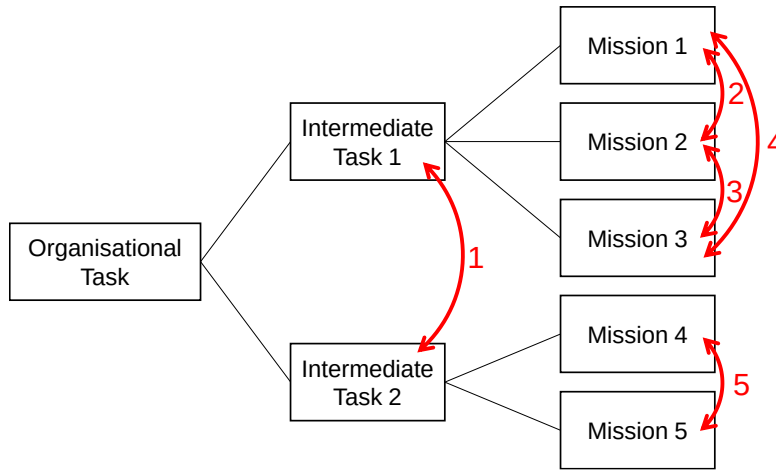


Figure 18 Task Hierarchy and Pairwise Comparisons

These comparisons correspond to giving answers to 5 questions, correspondingly numbered as follows:

1. How important is IT1, compared to IT2?
2. How important is M1, compared to M2, for the accomplishment of IT1?
3. How important is M2, compared to M3, for the accomplishment of IT1?
4. How important is M3, compared to M1, for the accomplishment of IT1?
5. How important is M4, compared to M5, for the accomplishment of IT2?

As the answers to the aforementioned questions, following statements are assumed:

1. IT1 is of moderate importance in comparison to IT2.
2. M1 is of moderate importance in comparison to M2.
3. M2 is of strong importance in comparison to M3.
4. M1 is of extreme importance in comparison to M3.
5. M4 is of moderate importance in comparison to M5.

First, importance of IT1 and IT2 were determined. Following the definition of the fundamental scale (Table 1), a_{12} can be determined to be 3, while a_{21} is $1/3$, giving pairwise comparison matrix $A = \begin{pmatrix} 1 & 3 \\ 1/3 & 1 \end{pmatrix}$. By numerically solving Equation (4-1), its principal eigenvalue $\lambda_{\max}$ was calculated as 2, while the

corresponding eigenvector was identified to be $\begin{pmatrix} w_{IT1} \\ w_{IT2} \end{pmatrix} = \begin{pmatrix} 0.75 \\ 0.25 \end{pmatrix}$, as $\sum_{i=1}^2 w_{ITi} = 1$ must hold.

Subsequently, M1, M2, and M3 were prioritised. This time, $A = \begin{pmatrix} 1 & 3 & 9 \\ 1/3 & 1 & 5 \\ 1/9 & 1/5 & 1 \end{pmatrix}$ and the transitivity rule does not hold. Numerical calculation gives principal eigenvalue $\lambda_{\max} = 3.0291$, and $CI = \frac{3.0291-3}{3-1} = 0.0146$. Since CI is less than 0.1 in this case, the result can still be considered as sufficiently consistent. $\sum_{i=1}^3 w_i = w_{IT1} = 0.75$ and therefore $\begin{pmatrix} w_1 \\ w_2 \\ w_3 \end{pmatrix} = \begin{pmatrix} 0.5037 \\ 0.1991 \\ 0.0472 \end{pmatrix}$. Finally, regarding M4 and M5, $A = \begin{pmatrix} 1 & 3 \\ 1/3 & 1 \end{pmatrix}$ and $\begin{pmatrix} w_4 \\ w_5 \end{pmatrix} = \begin{pmatrix} 0.1875 \\ 0.0625 \end{pmatrix}$, as $\sum_{i=4}^5 w_i = w_{IT2} = 0.25$. Using icicle chart, the result was visualised as Figure 19.

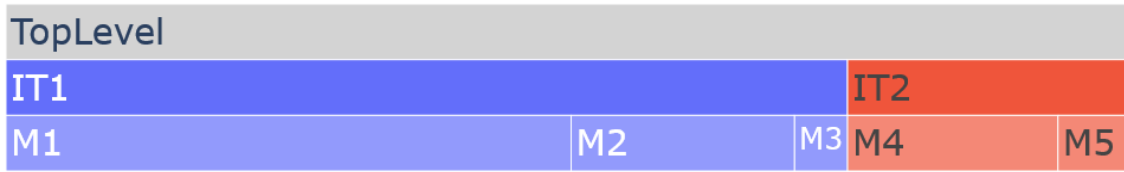


Figure 19 Icicle Chart of Value Hierarchy

A case where pairwise comparison results changed can also be considered. If the relative IT1 over IT2 becomes “strong” rather than “moderate”, the first statement will be as follows:

- IT1 is of strong importance in comparison to IT2.

While the other statements remain unchanged.

In such case, the pairwise comparison matrix for IT1 and IT2 becomes $A = \begin{pmatrix} 1 & 5 \\ 1/5 & 1 \end{pmatrix}$, as $a_{12} = 5$ and $a_{21} = 1/5$ by applying the fundamental scale. The principal eigenvalue $\lambda_{\max}$ was again 2, while the corresponding eigenvector was identified to be $\begin{pmatrix} w_{IT1} \\ w_{IT2} \end{pmatrix} = \begin{pmatrix} 0.8333 \\ 0.1667 \end{pmatrix}$. As w_{IT1} and w_{IT2} affects w_1 to w_5 , they are

recalculated as $\begin{pmatrix} w_1 \\ w_2 \\ w_3 \end{pmatrix} = \begin{pmatrix} 0.5037 \\ 0.1990 \\ 0.0472 \end{pmatrix}$ and $\begin{pmatrix} w_4 \\ w_5 \end{pmatrix} = \begin{pmatrix} 0.1875 \\ 0.0625 \end{pmatrix}$. As seen from the result, the change affected the intermediate-level task while the relative trade-off ratios at the lower mission level remained unaffected. The revised icicle chart is shown in Figure 20.



Figure 20 Icicle Chart of Value Hierarchy (Changed)

The change of mission importance weights can also be visualised in a form of stacked bar chart (Figure 21), which is useful for tracking changes.

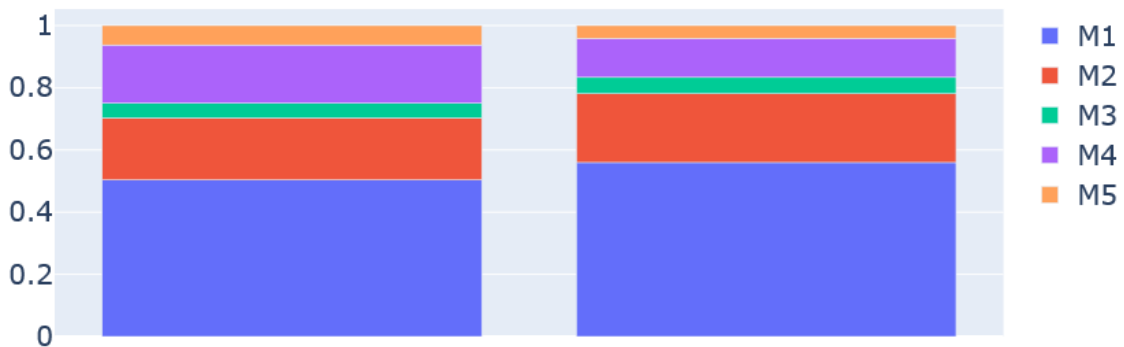


Figure 21 Stacked Bar Chart of Mission Importance Weights

From the results of the demonstration test, it can be concluded that the proposed method is capable of systematically translating the pairwise comparison ratings, which were expressed qualitatively, into quantitative importance weights representing the trade-off relationships between tasks and capability levels. This corresponds to the identification of value model reflecting the value perception of the capability acquirer.

It needs to be noted that, in practice, qualitative rating scale need to be “calibrated” so that the intention of stakeholders is appropriately reflected on the eventual quantitative value trade-offs. Additionally, when pairwise comparison

results are provided by different group of people (e.g., strategic-level planner and tactical-level planner), consensus need to be made between them to align what is implied by the ratings; otherwise the result shall be inconsistent. Ultimately, these calibration and consensus may require iterations, i.e., attempt to evaluate Sol alternatives, analyse how different comparison ratings affect their values, and adjust the comparison ratings if needed. Using the implemented framework, which is described later in Section 6.3, this iteration is expected to be easily and interactively carried out.

4.4 Conclusions

Responding to Objective 2 of this research, a specific method to systematically prioritise capability needs was formulated and tested. From existing criteria weighting methods, pairwise comparison method with the ratio scale, as used in AHP, were adopted for the purpose. It determines quantitative importance weights based on qualitative ratings describing the relative intensity of importance of one task over another.

In the demonstration test, the proposed method proved to be able to systematically determine the importance weights of tasks on different tiers in a task hierarchy.

Within the overall framework proposed in Chapter 3, these weights correspond to the quantitative trade-off ratios of capability levels regarding different tasks. The method was found to be able to produce a value model which constitutes a part of the framework.

5 Mission Effectiveness Estimation

5.1 Introduction

In this chapter, a method formulated with regard to Objective 3 of this research is described:

Objective 3: Develop a method for trade-off analysis regarding Sol's design, which accounts for mission context.

First, the problem is re-formulated in line with the framework defined in Chapter 3 using related mathematical notations. As discussed in Section 3.2.5, the relationship between a design vector $\mathbf{D}$, a context vector $\mathbf{C}$, and a measures of effectiveness (MoEs) vector $\mathbf{E}$ in mission i can be expressed as follows using a hypothetical function $f_{effectiveness\ i}$:

$$\mathbf{E} = f_{effectiveness\ i}(\mathbf{D}, \mathbf{C}) \quad (5-1)$$

Where:

$$v_i = f_{mission\ i}(\mathbf{D}, \mathbf{C}) = f_{success\ i}(f_{effectiveness\ i}(\mathbf{D}, \mathbf{C})) \quad (5-2)$$

$f_{effectiveness\ i}$, defined as such, can be viewed as the representation of the entirely objective relationships between the input $(\mathbf{D}, \mathbf{C})$ and the outcome $(\mathbf{E})$ regarding the mission execution. This is because it was defined to be decoupled from $f_{success\ i}$, which pertains to the stakeholder's subjective preference in determining v_i . In the terminology of the design loop concept (Ricci et al., 2014), $f_{effectiveness\ i}$ is a part of performance model while $f_{success\ i}$ is a part of value model. To enable value-driven model-based exploration of design alternatives, both these models need to be identified and the loop needs to be "closed". While $f_{success\ i}$ was assumed to be given *a priori* in this research, the need for identifying $f_{effectiveness\ i}$ remains. Therefore, the trade-offs regarding the Sol's design and context, which were stated in the Objective 3 of this research, need to be analysed against MoEs. This calls for the ability to estimate the output from the input and vice versa regarding $f_{effectiveness\ i}$. To enable reasonable (reasonably accurate) and quick

estimation, simple executable model representation of $f_{effectiveness\ i}$ is needed. If the following assumption holds:

Assumption: MoEs achieved by a given solution can be estimated from empirical data.

What needs to be developed is a statistical method which enables replicating the model accounting for the variability within the dataset. Operation modelling activity, which is discussed in Section 2.5, is a potential data source for this.

In response to this need, a method based on Structural Equation Modelling (SEM) (Loehlin, 1998) was developed in this research. Because of SEM's ability to specify causal relationships, it potentially can yield parsimonious and interpretable model relating design, context, performance, and effectiveness in the mission.

The rest of this chapter is organised as follows: In Section 5.2, general description of SEM is given and its relationship with the proposed method is discussed. In the subsequent section, 5.3, results from the test case (use case) are presented. Conclusions are presented in the final section, 5.4.

5.2 Structural Equation Modelling

SEM is a generic multivariate statistical method to establish hypotheses on underlying relationships among variables. Hypotheses are to be specified by the user of the method and denoted in the form of a path diagram describing causation and correlation of variables, then tested against observation data.

Explored in this research was the possibility of using SEM for analysing input-output relationships of $f_{effectiveness\ i}$. The input is a set of factors related to capability systems (design variables constituting D) and their context (context variables constituting C) whose variation causes impact on the output. The output, which are affected by the varied input, is a set of MoEs (E), as well as some measures of performance (MoPs) (P). MoPs are not related to the mission accomplishment as a whole but are related to the parts thereof. By definition, MoPs do not directly affect $f_{success\ i}$ and mission capability level, hence they do

not appear in the rest of the framework, which pertains to the higher level of the hierarchy of capabilities. Nevertheless, their data can be utilised alongside MoEs to support modelling $f_{effectiveness\ i}$.

If observations are made through experiments (e.g., simulations), the input factors are varied under the control of the analysts. Factors can be continuous, as well as discrete dummy variables, whose levels represent alternatives of attributes of SoI and its context. For example, a variable representing “sensor type” can take a value either 0 or 1, each of which representing “infrared” or “radar”. Allocation of input values to simulation runs may call for consideration on Design of Experiments (DoE).

Since the purpose of this analysis is to understand the impact of changes occurring within the solution and the environment, outcome of interest is the increment or decrement of MoEs caused by these changes. While this approach is similar to sensitivity analysis, it is also intended to handle qualitative changes as well as quantitative ones. Qualitative differences of alternative solutions, such as type of a sensor used, inclusion or exclusion of a system, personnel skill level and so forth, are codified using dummy variables. By doing so, different heterogenous solution alternatives can be modelled and analysed.

5.2.1 Model Specification without Latent Variables

The most straightforward candidate of the hypothetical path diagrams is the one with all possible one-way paths from each input variable X_i to each output variable Y_i (Figure 22, note that constant term is omitted). It is equivalent to multiple linear regression, as shown in equation (5-3). ε_i is the error term associated with an output variable Y_i .

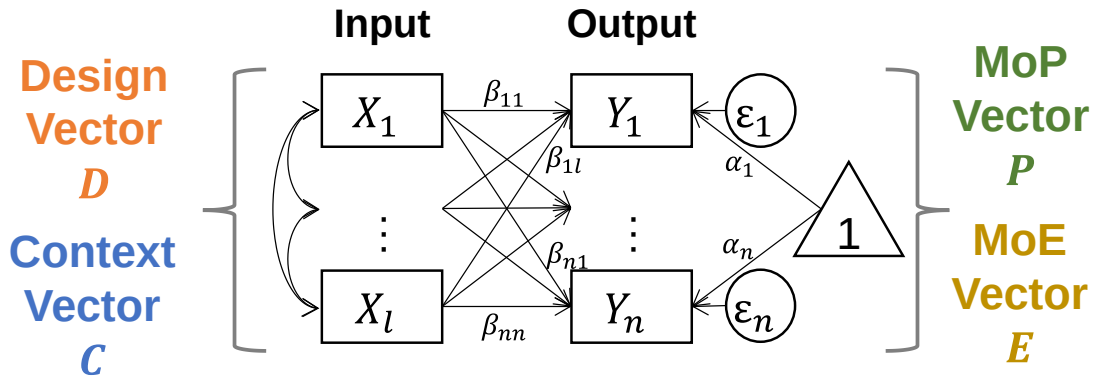


Figure 22 Hypothetical Path Diagram with All Direct Paths

$$Y = \alpha + BX + \varepsilon = \begin{pmatrix} Y_1 \\ \vdots \\ Y_n \end{pmatrix} = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} + \begin{pmatrix} \beta_{11} & \cdots & \beta_{1l} \\ \vdots & \ddots & \vdots \\ \beta_{n1} & \cdots & \beta_{nl} \end{pmatrix} \begin{pmatrix} X_1 \\ \vdots \\ X_l \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_n \end{pmatrix} \quad (5-3)$$

This path diagram contains the following unknowns: n intercepts in α , ln regression coefficients in B , l input means $M(X_i)$, l input variances $V(X_i)$, and n error variances $V(\varepsilon_i)$. Relations among the variables are given per elements of mean vector μ and variance-covariance matrix Σ defined as equation (5-4) and (5-5).

$$\mu = \begin{pmatrix} M(X_1) \\ \vdots \\ M(X_l) \\ M(Y_1) \\ \vdots \\ M(Y_n) \end{pmatrix} \quad (5-4)$$

$$\Sigma = \begin{pmatrix} V(X_1) & \cdots & Cov(X_1, X_l) & Cov(X_1, Y_1) & \cdots & Cov(X_1, Y_n) \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ Cov(X_l, X_1) & \cdots & V(X_l) & Cov(X_l, Y_1) & \cdots & Cov(X_l, Y_n) \\ Cov(Y_1, X_1) & \cdots & Cov(Y_1, X_l) & V(Y_1) & \cdots & Cov(Y_1, Y_n) \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ Cov(Y_n, X_1) & \cdots & Cov(Y_n, X_l) & Cov(Y_n, Y_1) & \cdots & V(Y_n) \end{pmatrix} \quad (5-5)$$

By following the path tracing rule, their equations are derived as (5-6) to (5-9). Assumptions were made that the error term mean $M(\varepsilon_i)$ are zero.

$$M(Y_i) = \alpha_i + \sum_{k=1}^l \beta_{ik} M(X_k) \quad (5-6)$$

$$V(Y_i) = \sum_{k=1}^l \beta_{ik}^2 V(X_k) \quad (5-7)$$

$$Cov(X_i, Y_j) = \beta_{ji} V(X_i) \quad (5-8)$$

$$Cov(Y_i, Y_j) = \sum_{k=1}^l \beta_{ik} \beta_{jk} V(X_k) \quad (5-9)$$

As there are $l + n$ unique elements in $\boldsymbol{\mu}$ and $(l + n)(l + n + 1)/2$ unique elements in $\boldsymbol{\Sigma}$, the degree of freedom of this model is $\{l(l - 1) + n(n - 1)\}/2$. The objective of model fit is to estimate these unknown values which make model-implied mean vector $\hat{\boldsymbol{\mu}}$ close to the sample mean vector $\mathbf{m}$, simultaneously make model-implied variance-covariance matrix $\hat{\boldsymbol{\Sigma}}$ close to the sample variance-covariance matrix $\mathbf{S}$. $\mathbf{m}$ and $\mathbf{S}$ can be calculated from the observations.

The diagram can be simplified by removing some paths where the causal relationship between the input variable and the output variable is deemed negligible. This is equivalent to fixing some regression coefficients (β_{ij}) to zero.

5.2.2 Model Specification with Latent Variables

As well as observed variables, latent variables can be included in the path diagrams. Unlike observed variables which are manifest, latent variables represent hypothetical constructs which are observed only indirectly. Their very existence is hypothetical and therefore can be considered as part of the hypothesis which the path diagram is representing. Reflective and formative constructs are what typical latent variables represent; reflective one is a construct which is thought to be causing effects on observed variables which are “effect indicators”, while formative one is a construct which is thought to be a consequence of effects caused by observed variables which are “causal indicators”. In addition to these, MIMIC (Multiple Indicator, Multiple Cause)

construct, which has both causal indicators and effect indicators (e.g., (Hancock, 2012)) can be considered. In Figure 23, basic path diagrams with these three types of latent construct are compared. A latent construct is expressed as η_i , while the error associated with it is expressed as ζ_i .

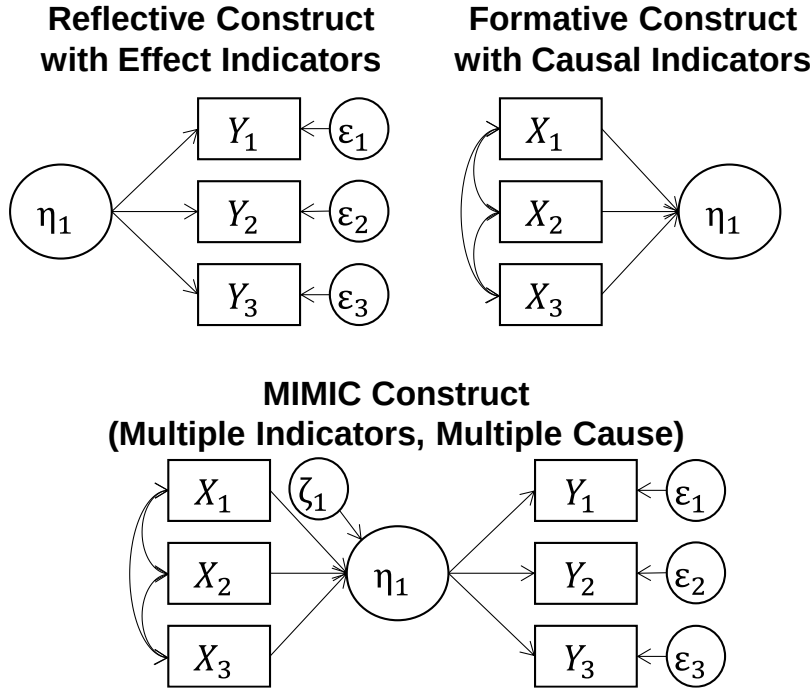


Figure 23 Three Types of Latent Variable Construct

From the application viewpoint, a possible use of latent constructs in mission models is to express some form of “building blocks” of mission accomplishment which cannot be directly measured. They are supposed to be influenced by some of the input factors (i.e., design variables in $\mathbf{D}$ and context variables in $\mathbf{C}$), which are varied under the analyst’s control and modelled as manifest variables. At the same time, the latent constructs have some influence on MoEs, which are the output of mission models. Therefore, in such case, MIMIC constructs are likely to be present in a path diagram. An example of such a path diagram is illustrated in Figure 24 (Note that intercepts α_i are omitted from the figure). In the form of equations, it is expressed as (5-10) and (5-11).

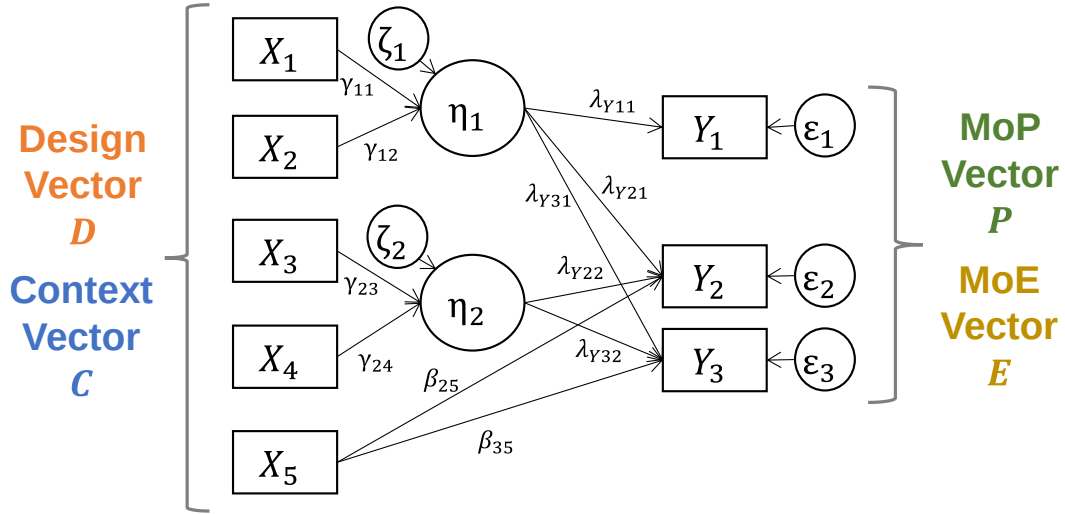


Figure 24 Example of Proposed Hypothetical Model

$$\begin{aligned}
 Y &= \alpha + BX + \Lambda_Y \eta + \varepsilon \\
 = \begin{pmatrix} Y_1 \\ \vdots \\ Y_n \end{pmatrix} &= \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} + \begin{pmatrix} \beta_{11} & \cdots & \beta_{1l} \\ \vdots & \ddots & \vdots \\ \beta_{n1} & \cdots & \beta_{nl} \end{pmatrix} \begin{pmatrix} X_1 \\ \vdots \\ X_l \end{pmatrix} + \begin{pmatrix} \lambda_{Y11} & \cdots & \lambda_{Y1m} \\ \vdots & \ddots & \vdots \\ \lambda_{Yn1} & \cdots & \lambda_{Ynm} \end{pmatrix} \begin{pmatrix} \eta_1 \\ \vdots \\ \eta_m \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_n \end{pmatrix} \quad (5-10)
 \end{aligned}$$

$$\eta = \Gamma X + \zeta = \begin{pmatrix} \eta_1 \\ \vdots \\ \eta_m \end{pmatrix} = \begin{pmatrix} \gamma_{11} & \cdots & \gamma_{1l} \\ \vdots & \ddots & \vdots \\ \gamma_{m1} & \cdots & \gamma_{ml} \end{pmatrix} \begin{pmatrix} X_1 \\ \vdots \\ X_l \end{pmatrix} + \begin{pmatrix} \zeta_1 \\ \vdots \\ \zeta_m \end{pmatrix} \quad (5-11)$$

In these equations, unknowns are n intercepts in α , regression coefficients in B , Γ and Λ_Y , l input means $M(X_i)$, l input variances $V(X_i)$, m disturbance variances $V(\zeta_i)$, and n error variances $V(\varepsilon_i)$. Definitions of mean vector μ and variance-covariance matrix Σ remain the same as equation (5-4) and (5-5). Equations of their elements are derived by following the path tracing rule.

The concept of Mission Thread in Mission Engineering suggests that performances in individual tasks within a mission (hereinafter called *subtasks*) can be viewed as “building blocks” of mission accomplishment. If these performances are not directly measurable, they need to be modelled as latent constructs. The allocation of systems (i.e., capability components constituting a capability system) onto subtasks within a mission, conceptualised as Mission Engineering Thread, also implies that design variables related to a particular subtask are influential to the accomplishment of only the subtask itself, and not to other subtasks.

For example, in an air defence mission, which is supposed to have “sense”, “decide”, and “influence” subtasks, components allocated to “sense” subtask are ground-based Early Warning Radar (EWR) and onboard radar of fighter aircraft. Therefore the “sensing” performance of the capability system is the aggregation of the individual performances of these components. Accordingly, the design variables related to these components are influential to the “sensing” performance, which is a latent construct. Likewise, “influence” performance may be hypothesised as the aggregated performance of surface-to-air missiles and air-to-air missiles, each of which are characterised by the numbers of munitions and their probabilities of kill. In turn, the MoEs, which are the output of the mission model, may be simply expressed as the aggregation of “sensing”, “deciding”, and “influencing” performances.

Additionally, use of latent constructs potentially allows to hypothesise and establish factors such as “morale of personnel”, or “user-friendliness of equipment”, which are influential to MoEs and assumed to be measured with a set of MoPs, while how to appropriately quantify them is yet unclear. As performance of non-materiel components also counts in capability systems, they are additional important considerations to analyse performance of capability systems, in comparison with engineering system-of-systems. Potential challenges to enable this shall be establishment of methods to collect reliable dataset, which may involve human interaction, occasionally in “live” operational environment.

This approach has a similarity with a work presented by Huang et al. (Huang et al., 2021) in that the underlying structure within mission capability is investigated and some building blocks are introduced as latent constructs. However, their work does not account for exogenous input factors (design and context variables) and latent constructs are modelled as reflective constructs. Seemingly, the focus of their research was on hypothesising the existence of building blocks, rather than investigation of how these building blocks and the total capability are related to the design and context.

The inclusion of latent constructs potentially provides richer information regarding the relationships within a mission model, as it reveals trade-off between accomplishment of subtasks, as well as trade-off between input factors within one subtask. However, while this adoption of Mission Thread and Mission Engineering Thread for mission analysis may give a simple and straightforward hypothesis of the quantitative mission model, it should be noted that such hypotheses always need to be validated against data, and need to be rejected if the data contradicts with them (i.e., a good model fit is not achieved). If no valid latent construct was established as a part of hypothesis, the fall-back model shall be the ones without latent variables, which were discussed in Section 5.2.1. Also, throughout this method, linearity of the mission model is assumed. If the range of variation of each input factors are small enough, this will not cause a problem. However, if it is not the case, it will be unlikely to obtain a good model fit with these linear models.

Within the overarching evolution planning activity of a Sol, the purpose of identification of $f_{effectiveness\ i}$ is to enable estimation of E for a given D , under a given C . While SEM can be a solution for it in principle, the accuracy of the identified model and the efficiency of the identification are dependent on the data on which the identification is based. In terms of accuracy, data for E should be collected from a wider area within the design space and context space, covering as much of the range of each design and context variable. Here, the range of design variables is determined by the scope of exploration of design alternatives, while the range of context variables is determined by the breadth of the scenarios. On the other hand, in terms of efficiency, the number of runs of experiments should be as small as possible. To optimally strike a balance between the accuracy and efficiency, a rigorous approach based on DoE shall be required according to the mission and the form of path diagram. However, this research avenue was not pursued further.

5.2.3 Model Estimation

The overall procedure for estimating a model of $f_{effectiveness\ i}$ is the following:

1. Define E based on $f_{success\ i}$, which is inputted from relevant stakeholders.

2. Define D , C , and P .
3. Determine possible levels of D and C .
4. Enumerate combinations of D and C to be simulated.
5. Model system alternatives and their context.
6. Execute experiments to collect data.
7. Formulate a hypothetical path diagram.
8. Test the hypothesis by fitting data.
9. If a good model fit was achieved, $f_{effectiveness\ i}$ is established. If not, go back to step 7 for another hypothetical path diagram.

In step 2, possible directions of changes expected within the capability system are listed considering their possible evolution. Likewise, those related to the context are listed as well considering future scenarios. They become the axes of design and context variables and span the design space and the context space.

In step 3, possible variations of the capability system and its context are expressed as different values (in the terminology of design and analysis of experiments, they are referred to as “levels”) along the individual axes.

In step 4, combinations of design and context variables are enumerated to form the input to the mission model. Each combination corresponds to a simulation test case. For the better estimation results of the mission model, it is generally preferable to create many combinations with different levels of the variables.

In step 5, each of the system alternatives, which are characterised by design and context variables, is expressed as a model. The models embody the assumptions regarding the properties of the systems, including those which are not expressed by the variables. For example, it is possible that two design variables, which are “radar power” and “number of missiles”, characterise alternatives of a fighter aircraft system, while other system properties such as kinematic performance of the airframe, also need to be present to carry out experiments. In these experiments, factors outside the design and context variables need to be kept constant (i.e., they are parts of numerous “control variables” in the terminology of design and analysis of experiments). But how to keep them constant may depend

on the purpose of the analysis. In the above example, one approach is to keep the “design philosophy” behind the airframe constant (e.g., achieve the same level of kinematic performance) and optimise the airframe for every alternative in line with it. Another approach is to assume one common airframe in all alternatives. Both approaches are valid since models are created based on consistent assumptions. Naturally, in the former approach, the airframe with high radar power and more missiles will be larger than the one with low radar power and fewer missiles. This allows the analyst to consider the cost increment caused by the increased airframe size. On the other hand, in the latter approach, the fighter’s kinematic performance may be degraded if it is equipped with a better radar and more missiles, which make the aircraft heavier. Consequently, the analyst will be able to consider the differences in mission equipment taking performance disadvantages into account. Overall, the analyst is required to consider how assumptions should be made in accordance with the purpose of the analysis.

5.3 Demonstration Test

To demonstrate and verify the proposed method, a case study was carried out to estimate $f_{effectiveness\ i}$ based on experiment data. The data were generated by an off-the-shelf combat simulation gaming software which can simulate combat missions involving various systems including fighter aircraft, ground-based Surface-to-Air Missile (SAM) Systems, ground-based radar site, and so forth. The software was configured as a constructive simulation; that is, all entities are automatically controlled by the programme without any human intervention. The software also allows to adjust the simulated operators’ skill level for each equipment system. Since it was a proprietary software, the details of the simulation process were not accessible and hence the mission model was treated as a black box. However, because of the ability of SEM to estimate the internal structure of black-box systems, such details were not necessary, and any simulation environment can basically be used in conjunction with it.

A hypothetical Defensive Counter-Air (DCA) mission was created to serve as the estimation test case. It is an air defence mission scenario to protect friendly

targets from adversary air attack, where fighter-bomber aircraft and cruise missiles are employed.

Step 1: Definition of E

To assess the capability level of the mission, the following three MoEs were assumed to be used:

- E_1 : Number of neutralised adversary aircraft
- E_2 : Number of adversary missiles hitting the friendly targets
- E_3 : Number of destroyed friendly aircraft

Here, E_1 is “higher the better”, while E_2 and E_3 are “lower the better” type of MoE.

Steps 2 and 3: Definition of D , C , P and Levels of D , C

In the mission, a pair of fighter aircraft equipped with Air-to-Air Missiles (AAMs) and a SAM battery are supposed to intercept the adversary air attack. To detect the adversary units, a fixed ground-based EWR can be integrated to share the collected target information with fighters and SAM.

Therefore, in this mission, fighters, SAM, EWR, and their operating personnel were seen as the capability system. Performance of these equipment and skill level of the personnel were adjustable to vary. The following three factors were chosen as the design variables for this case study:

- D_1 : Presence of EWR (1: Present, 0: Absent)
- D_2 : Performance of AAMs (1: High, 0: Low)
- D_3 : Fighter pilots' skill level (1: High, 0: Low)

Also, the following factor was chosen to form the context variable:

- C_1 : Performance of the adversary fighter aircraft (1: Present, 0: Absent)

As can be seen, the design and context variables in this particular demonstration case were binary rather than continuous. It should be noted that it was due to a limitation imposed by the simulation software used in step 6, which only allowed choosing performance parameters from pre-set levels. The proposed method

itself is compatible with continuous variables, for example, range of AAMs between 50 and 100 nautical miles.

Step 4: Enumeration of Possible Combinations of D and C

As defined in the previous step, the design and context variables and their levels can make $2^4 = 16$ combinations. In this case study, all these possible combinations were retained for the subsequent simulation.

Steps 5 and 6: Modelling and Simulation Execution

Models of the Sol and its context were generated by inputting the enumerated D and C to the combat simulation software. Subsequently, simulation runs were executed to collect data. As 5 runs were executed for each simulation test case, 80 runs were executed out in total.

Steps 7, 8, and 9: Mission Model Estimation

In these steps, JMP 17, a statistics software, was used to process the simulation output. The MoEs were standardised (i.e., a linear transformation was applied to adjust the mean to 0 and the standard deviation to 1) as the pre-processing, since they are expressed in natural scales and discrepancy of their magnitude may affect the model fit. Also, to unify the signs, E_2 and E_3 were inverted to make them “higher the better”.

A hypothetical path diagram was then specified from the Graphical User Interface (GUI) and the data was fit on it. In this case study, latent variables were not included in the model, as valid latent constructs were not found in the data in this use case. While this made the result equivalent to multiple regression, the applicability of SEM to mission effectiveness estimation was demonstrated.

The resulting path diagram is shown in Figure 25. It shows that trade-off relationships between input factors were appropriately captured. As these factors are coded as 0 / 1 binaries, the weight of a path from a design variable to a MoE represents the increment of standardised MoE made by each individual change in design and context. For example, the weights of paths from D_1 exceed those from D_2 and D_3 for all MoE, indicating that addition of EWR into the capability

system is more beneficial than improving missiles or crew's skill. In practice, these benefits realised by the possible “upgrade plans” can be compared to inform the decision making regarding the evolution of a capability system.

By applying inverse transformation of the linear transformation applied to standardise the MoEs, equation (5-12), with MoEs in natural scale, is obtained.

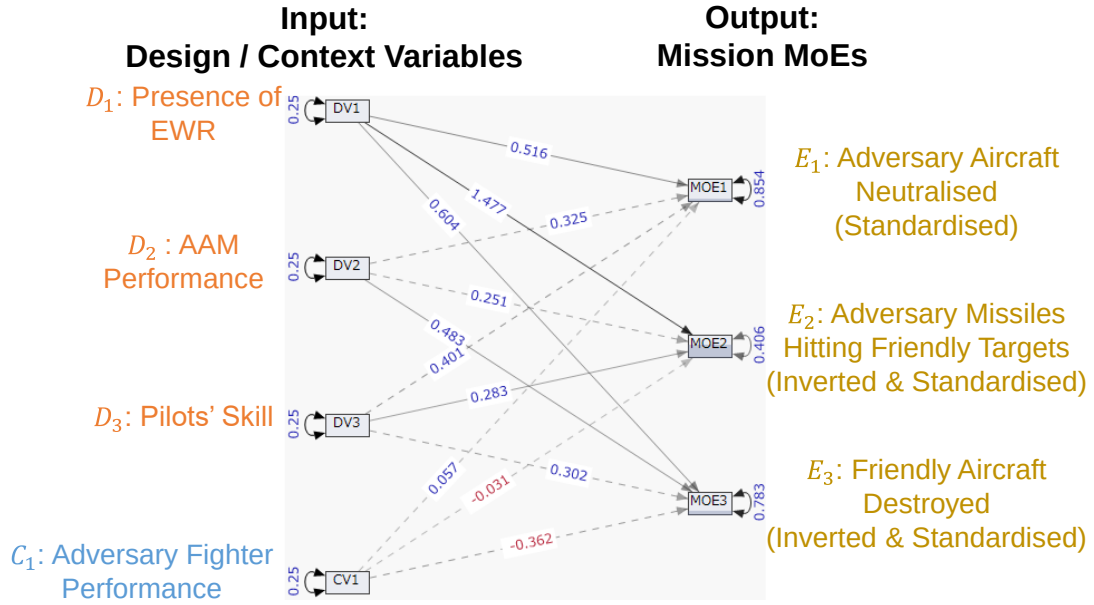


Figure 25 Fitted Model of a DCA Mission

$$\begin{pmatrix} E_1 \\ E_2 \\ E_3 \end{pmatrix} = \begin{pmatrix} 0.675 & 0.425 & 0.525 & 0.075 \\ -2.350 & -0.400 & -0.450 & -0.050 \\ -0.500 & -0.400 & -0.250 & -0.300 \end{pmatrix} \begin{pmatrix} D_1 \\ D_2 \\ D_3 \\ C_1 \end{pmatrix} + \begin{pmatrix} 1.937 \\ 6.100 \\ 1.775 \end{pmatrix} \quad (5-12)$$

The extent of the demonstration effort within the scope of this research was limited by the lack of combat simulation environment. In the future, it is desirable that a simulation environment with more flexibility in the input definition is employed. For example, if individual performance parameters of capability components, such as the radar range and the AAM range, can be adjusted more freely, more detailed data would become available. It is expected that the possibility of the use of latent constructs in the simulation can also be demonstrated with such an environment.

5.4 Conclusions

Responding to Objective 3 of this research, a method was proposed to estimate the trade-off relationships among Sol's design variables and context variables in terms of MoEs' achievement in a mission. The estimation is made possible through the identification of a black-box model of mission execution based on empirical data. The proposed method is based on SEM technique, which establishes a hypothesis of causal relationships through testing it against data. While the user needs to manually specify the hypothesis, this method potentially enables reaching a parsimonious and interpretable model. In the demonstration test, the applicability of SEM to a dataset of mission execution was verified. The method successfully identified the trade-off relationships within a mission.

Additionally, the possibility of introducing latent constructs as a part of causal models was reported, although a demonstration using data was yet to be conducted due to the lack of an appropriate simulation environment and use cases.

With the trade-off relationships identified by the proposed method, estimation of MoEs to be achieved by a given Sol design alternative under a given context is made possible. Since the method is based on empirical data and statistical processing, it ensures objectivity and traceability of decision making.

Industrial evaluation of the framework is presented in the next chapter.

6 Evaluation

6.1 Introduction

As a part of this research, evaluation was carried out for the verification and validation of the developed framework and methods. As suggested in DRM (Blessing & Chakrabarti, 2009), the evaluation of the research outcome was carried out in three phases: namely Support Evaluation, Application Evaluation and Success Evaluation. The first Support Evaluation phase pertains to the verification of the functionality of the capability development support tool developed in this research. Support Evaluations specifically regarding the capability needs prioritisation method and the mission effectiveness estimation method were carried out separately and discussed in chapters 4 and 5, respectively. Support Evaluation effort of the capability development support tool related to the whole framework is discussed in this chapter. For this purpose, a computer programme which can be integrated with the output of the aforementioned individual methods was created. Description of this programme is given in Section 6.2.

The second and third phases of the evaluation, which are Application Evaluation and Success Evaluation, were also carried out. The purpose of the Application Evaluation is to investigate the applicability and usability of the capability development support tool, while the aim of the Success Evaluation is to investigate its overall ability to achieve the intended success of the research. To conduct success evaluation in a comprehensive manner, the capability development support tool needs to be validated through an application to a real industrial problem. However, due to the practical limitations such as time and accessibility to real security-related data, the application and success evaluations were carried out to an appropriate level. That is, employing a realistic, but not real use case, allowing to confirm the applicability and successfulness of the proposed approach with reasonable confidence.

Thus, instead of real-world data, a hypothetical use case which imitated a simplified capability evolution planning problem was developed for the application

evaluation. The use case is described in Section 6.3. Also, a panel of practitioners with relevant industrial experience were consulted. This is important, as they were thought to have sufficient knowledge of the real-world problems and thus be able to judge the potential applicability, usability, and usefulness of the proposed framework, via the evaluation of the capability development support tool. To capture the expert opinion thoroughly and efficiently, the researcher developed a questionnaire. In Section 6.4, the details of this external evaluation session are given. Finally, conclusions are drawn in Section 6.5.

6.2 Software Implementation

The framework and methods described in Chapters 3 to 5 were implemented in a computer programme for the evaluation and validation of the proposed approach. Since the framework was intended to aid decision making, interactive scenario editing and data visualisation features were also included in the programme.

The entire programme was coded with Python 3.9.16 programming language (Python, 2022). For data visualisation, Plotly 5.15.0 (Plotly, n.d.-b) graphing library for Python was used. Also, to create the Graphical User Interface (GUI) realising the interactive features of the framework, Dash 2.10.2 (Plotly, n.d.-a) framework was used. Dash enables creating web-based applications which run on a web server, being accessible and usable from web browsers on the client side. The use of existing libraries (i.e., Plotly and Dash) enabled rapid construction of the programme.

When conducting an analysis in line with the framework, design alternatives, scenarios, and mission models need to be specified. On the premise that design alternatives can be characterised with a set of design vectors, design variables and their possible levels are specified in the code. When the programme is launched, design vectors representing design alternatives are automatically generated by exhaustively making combinations of levels specified for each design variables. Mission models (which can be the output of mission effectiveness estimation method discussed in Chapter 5) are specified in the code, in the form of a set of equations such as equation (5-12).

To allow specifying scenarios, context variables and their possible levels, as well as capability needs hierarchy, are also specified in the code. Specification of context vectors composing scenarios is enabled by allowing to choose appropriate levels from the predefined list of levels, for each epoch and for each scenario. Likewise, each pairwise comparison result of task importance can be specified by choosing appropriate rating from the ratio scale. The specified comparison results are instantly and automatically converted to importance weights, which are needs variables constituting a needs vector. By combining a sequence of context vectors and a sequence of needs vectors, scenarios are created. What-if analyses of different scenarios are enabled in this way, with the interactive scenario editing feature of the programme.

The overall GUI of the programme is illustrated in Figure 26 and Figure 27. The Scenario Planner enables editing each scenario by specifying its parameters (i.e., levels of context variables and pairwise comparison results of tasks) with the GUI. It handles one scenario at a time, and the scenario to be edited can be selected from a dropdown menu located at the top of the window. Situated on the top-left corner of the window of Scenario Planner is the needs specification part. It has tabs, each of which corresponding to an epoch. In each tab, pairwise comparison results of each pair of tasks are inputted from sliders. The importance weights in the value tree are visualised as an icicle chart, which is suitable for hierarchical structures. When any of the pairwise comparison results were modified, all importance weights within the value tree are re-calculated and reflected on the chart instantly. With the built-in function of Plotly, the icicle chart is interactive. Mouse-overing on one of the rectangles shows the value of importance weight. Clicking on it invokes zoom-in, focusing on its children tasks and below. In the “Era” tab, temporal changes of importance weights across epochs are expressed in the form of stacked bar chart.

On the top-right corner of the Scenario Planner window is the context specification part. Like needs specification, it has tabs, each of which with sliders corresponding to context variables. With each slider, the user can specify a level from a pre-defined set.

The bottom part of the Scenario Planner window is occupied by another set of tabs corresponding to each epoch and an era. Each epoch tab contains a data table and a parallel coordinate plot, where Measures of Effectiveness (MoEs), capability levels, and System-of-Interest (SoI) value at the epoch are shown for each design alternative of the SoI. The era tab also contains a data table and a parallel coordinate plot, which show the temporal change trajectory of MoEs, capability levels and SoI value throughout the era, which spans epochs in a scenario.

The Design Browser page also contains a data table and a parallel coordinate plot. It shows aggregated data from all considered scenarios to allow users to investigate how alternatives perform and how they are valued in different future scenarios.

With the built-in function of Plotly, parallel coordinate plots also have an interactive feature, where a user can filter the plots. By dragging lines along vertical axes, ranges of the axes can be specified so that plots which run through these ranges can be highlighted. This enables exploring alternatives by specific MoEs, capability levels, or values of the user's interest.

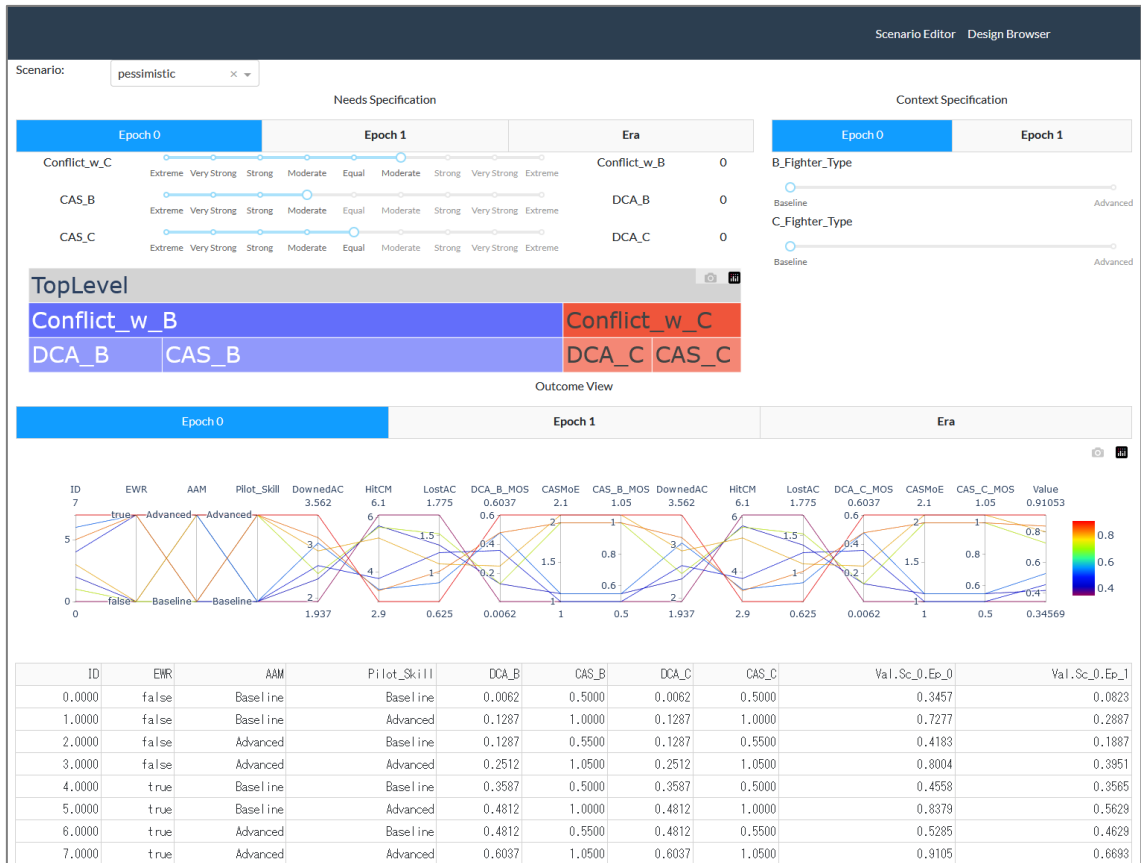


Figure 26 Scenario Planner Page of Demonstration Software

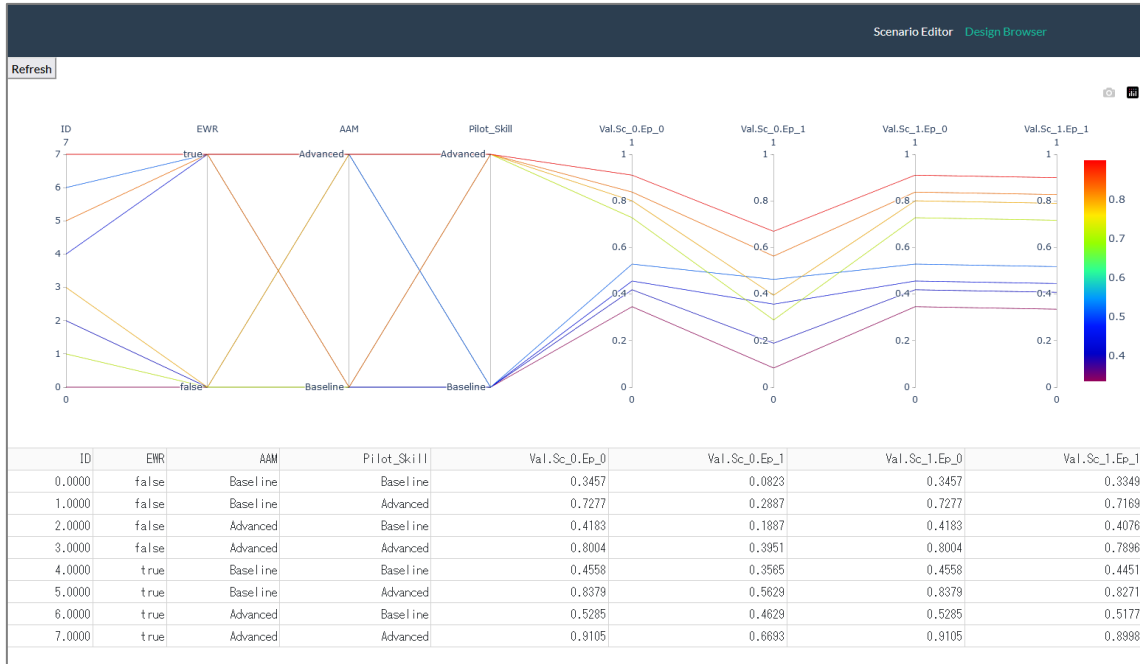


Figure 27 Design Browser Page of Demonstration Software

6.3 Case Study

In order to test and demonstrate the applicability of the developed framework as a whole, a use case was formulated by extending the ones developed for individual methods, which were discussed in Sections 4.4 and 5.4. It was intended to contain all the main issues which this framework was addressing. While the individual methods, which were discussed in Chapters 4 and 5, pertain to single epoch (i.e., a single point in time), the entire framework provides a method to aggregate their outputs to analyse era, expressed as a sequence of epochs. Therefore, the following aspects of regarding evolution of the Sol, as well as scenario planning, were included in the use case:

- (1) Expression of Sol evolution alternatives as temporal changes of Sol designs.
- (2) Expression of temporal changes of high-level capability needs.
- (3) Expression of temporal changes of mission capability needs.
- (4) Expression of temporal changes of contextual factors.
- (5) Calculation of value trajectory spanning the epochs.

Additionally, the following aspects of capability planning needed to be considered as well:

- (6) Involvement of capability components in multiple missions.
- (7) Comparison of value trajectories of Sol evolution alternatives under multiple scenarios.

In this section, the problem setting assumed in the use case is illustrated through a stepwise description of the scenario planning process. For the ease of understanding, it was described in a form of a hypothetical “story”. Since the purpose of the analysis is a decision making based on quantitative method, the story started from pre-identified candidates for future actions. However, this does not imply that these alternatives can come out of nowhere. In realistic situations, they are more likely to be devised following certain discovery efforts involving forecast of future capability needs and context. The alternatives at this stage are possibly devised through brainstorming-like qualitative means, based on the forecast.

Step 1: In the story, nation A was planning the future capability development for their air force in a 10-year time frame, which is expressed as an era composed of 2 epochs. Candidates of future actions were identified as follows:

- Acquire and field Early Warning Radar (EWR).
- Upgrade Air-to-Air Missiles (AAMs) for better performance.
- Train fighter pilots for higher skill level.

Considering the options to be investigated, Sol in this case was defined to include:

- AAMs
- EWR
- Fighter pilots

The following properties of the Sol were assumed to be under control and characterised using the following design variables:

- D_1 : Presence of EWR (1: Present, 0: Absent)

- D_2 : Performance of AAMs (1: High, 0: Low)
- D_3 : Fighter pilots' skill level (1: High, 0: Low)

The current design of the Sol was defined as $(D_1, D_2, D_3) = (0,0,0)$. Assuming that any of these variables could be changed to 1 in the next epoch 1 (for a certain investment), the possible actions to be taken were found range from doing nothing to executing all 3 actions, making $2^3 = 8$ alternatives in total.

Step 2: The capability needs in epoch 0 were identified. It was assumed that the high-level tasks are:

- Conflict with nation B, which has limited air power as of epoch 0.
- Conflict with nation C, which has advanced air power as of epoch 0.

To accomplish these high-level tasks, the air force was assumed to be required to conduct the following missions:

- DCA (Defensive Counter-Air)
- CAS (Close Air Support)

The resulting hierarchy of tasks (including missions) is illustrated as a tree diagram in Figure 28. In line with the capability needs prioritisation method discussed in Chapter 4, the prioritisation of entire capability needs completes by carrying out 3 pairwise comparisons indicated in the figure with numbers.

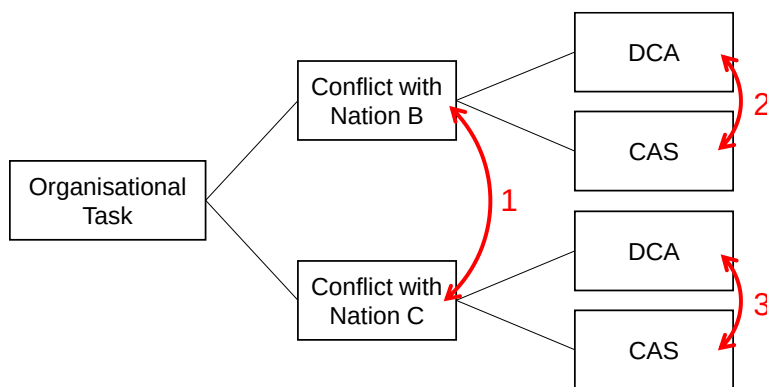


Figure 28 Task Hierarchy and Pairwise Comparisons of a Use Case

These comparisons correspond to giving answers to 3 questions, correspondingly numbered as follows:

1. How important is the conflict with nation B, compared to the conflict with nation C?
2. How important is DCA, compared to CAS, for the conflict with nation B?
3. How important is DCA, compared to CAS, for the conflict with nation C?

Step 3: As each mission listed in the previous step was viewed as an input-output model in the framework, the input and output of these models needed to be defined. Since the design variables were already defined in Step 1, only context variables needed to be defined considering the anticipated future changes. Specifically, the performance of adversary fighter aircraft was assumed to change in the future and therefore included as a context variable. As each of the adversary nations B and C has fighter aircraft, two variables corresponding to them were defined. These factors were assumed to influence both the DCA missions and the CAS missions.

Likewise, MoEs relevant to each mission were identified as the output of the corresponding mission models. As $f_{success\ i}$ are given by stakeholders, these MoEs are determined in accordance with the stakeholders' perception of capability levels. Hence, these MoEs are given a priori. In this demonstration (similarly to the case study carried out in Chapter 5), the following three MoEs were assumed to be identified for the DCA missions (either with Nation B or C):

- E_1 : Number of neutralised adversary aircraft
- E_2 : Number of adversary missiles hitting the friendly targets
- E_3 : Number of destroyed friendly aircraft

Regarding the CAS missions, it was assumed that the capability levels can be expressed using a single MoE:

- E_1 : MoE of CAS missions

Step 4: Having specified the input and output variables of the missions in the previous step, models representing the relationships between them were established as $f_{effectiveness\ i}$. Here, the result from Section 5.3 was adopted unchanged as the model of the DCA missions in this demonstration.

$$\begin{pmatrix} E_1 \\ E_2 \\ E_3 \end{pmatrix} = \begin{pmatrix} 0.675 & 0.425 & 0.525 & 0.075 \\ -2.350 & -0.400 & -0.450 & -0.050 \\ -0.500 & -0.400 & -0.250 & -0.300 \end{pmatrix} \begin{pmatrix} D_1 \\ D_2 \\ D_3 \\ C_1 \end{pmatrix} + \begin{pmatrix} 1.937 \\ 6.100 \\ 1.775 \end{pmatrix}$$

The assumed $f_{success\ i}$ is as follows:

$$v = 0.1E_1 - 0.1E_2 - 0.1E_3 + 0.6$$

Regarding the CAS mission model, the following model of $f_{effectiveness\ i}$ and $f_{success\ i}$ was assumed.

$$E_1 = (0 \quad 1.000 \quad 0.100 \quad -0.600) \begin{pmatrix} D_1 \\ D_2 \\ D_3 \\ C_1 \end{pmatrix} + 0.900$$

$$v = 0.5E_1$$

The $f_{effectiveness\ i}$ and $f_{success\ i}$, which compose mission models, were coded into the computer programme to output capability level v from design and context variables in D and C .

From these expressions, the influence of the input factors, which are the design and context variables, could be observed. For example, D_1 (Presence of EWR) is highly influential to the realised capability level of DCA mission, while it has no influence on that of CAS.

Step 5: Following the specification of relevant variables in the previous steps, possible futures were modelled into scenarios. In this case study, two “pessimistic” and “optimistic” scenarios were considered. While the situations (i.e., needs and context) as of epoch 0 were common in both scenarios, those of epoch 1 differed from each other. For each scenario, the needs and context were codified as variables.

Regarding the needs, tasks and missions within the task hierarchy were prioritised in accordance with the answers to the questions listed in Step 2.

For epoch 0, the answers were assumed to be as follows:

1. Conflict with nation B is moderately more important, compared to that of nation C.
2. For conflict with nation B, CAS is moderately more important than DCA.
3. For conflict with nation C, DCA and CAS are equally important.

In the “Pessimistic” scenario, the capability needs as of epoch 1 were assumed to be as follows:

1. Conflict with nation C will be strongly more important compared to that of nation B.
2. For conflict with nation B, DCA and CAS will be equally important.
3. For conflict with nation C, DCA will be strongly more important than CAS.

In the “Optimistic” scenario, they were assumed to be as follows:

1. Conflict with nation B will be moderately more important compared to that of nation C. (Unchanged)
2. For conflict with nation B, CAS will be strongly important than DCA.
3. For conflict with nation C, DCA will be moderately more important than CAS.

By applying the algorithm presented in Chapter 4, the capability needs as of each epoch were quantitatively expressed using importance weights.

Regarding the context, the following assumptions were made:

- In both scenarios, the performance of the fighter aircraft of Nation B and C are at “Baseline” level as of epoch 0.
- In “Pessimistic” scenario, the performance of the fighter aircraft of Nations B and C becomes “Advanced” level in epoch 1.
- In “Optimistic” scenario, the performance of the fighter aircraft of Nation C becomes “Advanced” level in epoch 1.

These scenarios were inputted into the “scenario planner” page of the computer programme using the sliders in each tab corresponding to an epoch. Screenshots of the programme output with specified scenarios are shown in Figure 29 to Figure 32.



Figure 29 Needs and Context as of epoch 0 in “Pessimistic” Scenario

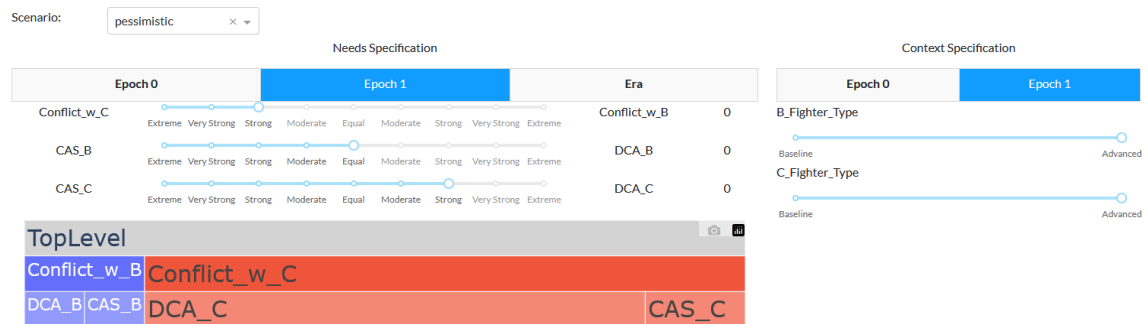


Figure 30 Needs and Context as of epoch 1 in “Pessimistic” Scenario



Figure 31 Needs and Context as of epoch 0 in “Optimistic” Scenario



Figure 32 Needs and Context as of epoch 1 in “Optimistic” Scenario

Step 6: For each scenario and each epoch, design alternatives were evaluated. First, estimated vectors of MoEs, $\mathbf{E}$ were calculated by inputting design variables representing the corresponding design and context variables specified in the scenarios into the $f_{effectiveness\ i}$. Subsequently, by applying $f_{success\ i}$, as coded in the programme, these MoEs were mapped to the mission capability levels, v_i . Then, these levels were aggregated into the value during each epoch (V). By repeating this process for each epoch in a scenario, trajectories of value were calculated for each evolution alternative.

Step 7: The resulting value trajectories were visualised in the form of parallel coordinate plots. This allows to investigate the advantages and disadvantages of each alternative under different scenarios. The output is shown as Figure 33.

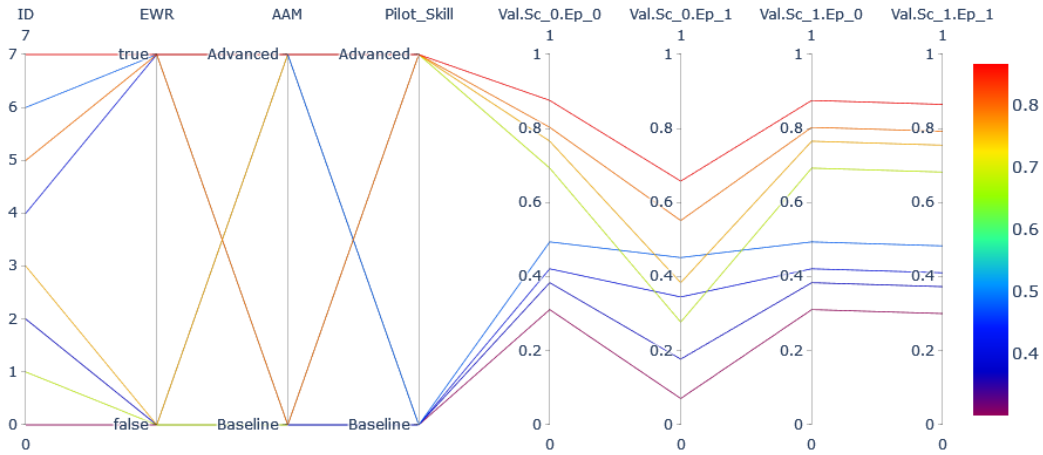


Figure 33 Visualisation of Value Trajectories

From the plot, the following observations were made:

- In both scenarios, values of evolution alternatives diminish as the epoch advances, and needs and context associated with it change.
- As a general trend, regardless of the alternatives, the decrements of the values are larger in “Pessimistic” scenario compared to “Optimistic” scenario. This is because of the higher priority of DCA missions and the

higher performance of the adversary fighter aircraft in “Pessimistic” scenario, which lowers the capability level of the DCA mission.

- As possible measures to counter the diminishing value, improvement of pilot’s skill and addition of EWR could be compared. The former is effective in CAS missions, while the latter is effective in DCA missions.
- As the consequence, addition of EWR was found to be the more effective approach in “Pessimistic” scenario, while pilot skill improvement was found to be more effective in the “Optimistic” scenario.
- Improvement of AAM, which is another evolution path, was found to be effective to a similar degree in both DCA and CAS missions. Thus, not much difference was observed between the scenarios.

Overall, it was demonstrated that the proposed framework was capable of estimating the value trajectories, which is the direct representation of the stakeholder’s perceived desirability of evolution alternatives. It must be noted that the above presented 7-step procedure covered the simple case, i.e., without reworks and iterations. In the above example, a set of alternatives was formulated in the first step and no changes were made in subsequent steps. However, in realistic situations, it is possible that the design space, context space and needs space are re-defined during or after the clarification of the scenarios. If possibilities of new adversary equipment or technology are recognised, potential missions as well as additional design variable may be specified as a result. For example, if a new hypersonic ground attack missile was found to be developed by one of the adversaries, the analyst needs to modify the DCA mission to model its employment in anticipated combat. In response to this, development of new interception missiles may be considered as additional alternatives to counter the new threat effectively. As mentioned above, additional design variables may be required to express such alternatives.

6.4 Industrial Evaluation

To verify the applicability, usability and usefulness of the framework, as well as to validate the framework against its purpose, an external evaluation was carried out. The validation was to cover the following two aspects; the validity of the

assumptions underpinning the framework, and the possibility of the industrial impact to be created by the framework. Since the framework presented in this thesis was developed as a capability development support tool, it is desirable to verify and validate it by demonstrating the ability to solve real industrial problems related to capability development. However, a live industrial evaluation was not possible within the scope of this research because of constraints regarding time and availability of real data from security-related practices. Therefore, as an initial verification and validation effort, a panel of external experts were invited for an evaluation, following a presentation of the methodology and a demonstration of the software with a realistic, but not real use case. The experts were expected to be able to judge the applicability, usability, and usefulness of the approach in the absence of full-scale demonstration. Hence, sufficient knowledge on the real-world capability development problems, which the developed support tool intended to address, was required. Specialists of operational analysis from BAE Systems Plc were considered to be appropriate for this role. A week prior to the evaluation session, preliminary information describing the research was provided to the evaluators. A copy of the document is included in Appendix A of this thesis.

To capture the opinion of the evaluators thoroughly and efficiently, a questionnaire composed of multiple-choice questions and open questions was designed. The detail of the questionnaire is given in Section 6.4.1. A copy of complete questionnaire form is included in Appendix B of this thesis.

An evaluation session took place on 15th June 2023 at the Warton site of BAE Systems. It included 5 participants from the operational analysis section of BAE Systems. Their years of experience ranged from 1 year to 35 years. Prior to the session, the participants received a document providing an introduction to the developed framework, as well as a questionnaire to be answered at the session. On the day of session, the researcher delivered a presentation and demonstration of the framework. Questions and answers and free discussion followed. At the end of the session, the evaluators were asked to complete the questionnaire.

6.4.1 Questionnaire

The questionnaire is composed of two parts, which are Likert-type multiple-choice questions and open questions. The first section of Likert-type questions was intended to capture the level of agreement of respondents regarding the validity of the following 7 assumptions on which the framework is based. Discussions on these assumptions were given in previous chapters. The respondents were required to choose 1 answer from 5 choices, which are “Strongly Disagree”, “Disagree”, “Neither Agree or Disagree”, “Agree”, and “Strongly Agree”.

1. It is reasonable to assume that a finite number of representative deterministic scenarios collectively capture the range of future uncertainties (as an alternative to a probabilistic approaches).
2. It is reasonable to identify customer needs with the capability to perform their organisational tasks.
3. It is reasonable to assume that a set of representative subtasks can be appropriately derived from a supertask so that the subtasks’ fulfilment collectively indicates the fulfilment of the supertask.
4. It is reasonable to evaluate the fulfilment of a supertask based on the fulfilment of its subtasks and their importance over each other.
5. It is reasonable for the customer to compare solution alternatives based on their contribution to the fulfilment of the organisational tasks.
6. It is reasonable to measure a complex mission achievement by aggregating measures of effectiveness.
7. It is reasonable to estimate measures of effectiveness achieved by a given solution from empirical data (e.g., simulation results).

The second section of the Likert-type questions was to capture the evaluators’ view on whether the individual methods, which were discussed in Chapters 3 and 5, serves their intended purposes.

8. A hierarchy of tasks (i.e., from strategic to tactical level) adequately describes customer needs and their changes.
9. Pairwise semi-quantitative comparisons for task prioritisation help to translate intangible customer needs into formal value functions.

10. Estimated causal models of missions help to understand underlying relationships between solutions and their mission effectiveness.

The third section of the Likert-type questions was intended to capture the evaluators' view on the validity of the overall framework in terms of the analytical outcome, as well as the impact on the industry.

11. The proposed framework enables effective identification of solution alternatives with a high value under a given future scenario.
12. The proposed framework is capable of adding value to the industry by providing useful knowledge for product development.

The following open questions were intended to collect recommendations and identify the advantages and disadvantages of the framework and the methods.

1. What possible difficulties and limitations do you anticipate in applying the proposed framework to real-world problems? (e.g. If you strongly disagree with one of the assumptions).
2. Regarding the proposed capability needs prioritisation method, what are the advantages and disadvantages in relation to the current industrial practice?
3. Regarding the proposed mission effectiveness prediction method, what are the advantages and disadvantages in relation to the current industrial practice?
4. Is there anything you found particularly useful within the proposed framework?
5. Is there anything you found particularly not useful within the proposed framework?

6.4.2 Feedback

The answers to the Likert-type questions are summarised in Table 2 to Table 4. The numbers in each column indicates the numbers of respondents who chose a particular level of agreement. Table 2 indicates that the panel mostly agreed on the validity of the underpinning assumptions of the framework.

Table 2 Evaluators' Levels of Agreement Regarding Framework's Underlying Assumptions

Question No.	Strongly Disagree	Disagree	Neither Agree nor Disagree	Agree	Strongly Agree
1				5	
2				4	1
3				3	2
4				5	
5				5	
6			1	4	
7				4	1

As can be seen in Table 3, the panel largely agreed that the proposed methods serve their intended purposes. In addition, evaluators pointed out that the capability needs prioritisation method has a secondary benefit in that it forces stakeholders to consider and somehow determine the priorities, preventing treating all capabilities to be equally important. While this view was expressed from managerial rather than technical viewpoint, it further supports the validity and usefulness of the framework.

Table 3 Evaluators' Levels of Agreement Regarding Individual Methods' Usefulness

Question No.	Strongly Disagree	Disagree	Neither Agree nor Disagree	Agree	Strongly Agree
8			2	2	1
9				5	
10				5	

Also, as indicated in Table 4, the validity of the framework was largely agreed by the panel. The ultimate output of the framework, which is the value trajectories under multiple scenarios, were considered as an appropriate representation of the long-term desirability of evolving Sol alternatives in uncertain futures. Also, it was expected that the information provided by the framework is useful for decision making in product development activities. From managerial viewpoint, it was recommended that care must be taken to prevent misuses of the framework. For example, over-trusting the results and ignoring underlying assumptions when using and communicating the outputs in practice.

Table 4 Evaluators' Levels of Agreement Regarding Usefulness and Validity of Overall Framework

Question No.	Strongly Disagree	Disagree	Neither Agree nor Disagree	Agree	Strongly Agree
11			1	3	1
12			1	4	

As one of the comments from the evaluators, a point was made regarding the task hierarchy. The framework was based on the assumption that at the same level of the task hierarchy, one capability level can always be traded-off against another one. However, it was pointed out that in practice, there can be certain absolutely important missions, which are not suitable for relative trade-offs. In such a case, the proposed approach can be easily augmented with setting of threshold values for capability levels. In this way, alternatives which cannot reach the set threshold should be rejected and eliminated from further considerations. In the proposed framework, implementation of the threshold idea is possible by simply including elimination criteria when alternatives are generated.

Also pointed out by the evaluators was the difficulty of creating models of missions and systems, especially those for future operations and future systems, which do not exist yet at the time of analysis and thus cannot be defined in detail. The participant with 30 years of experience, who chose "Neither Agree or Disagree" for Questions 11 and 12, pointed out in the first open question the particular difficulty to create mission models. Since it was the only difficulty or limitation mentioned in the answer, it could be understood as the chief concern which led to his neutral view towards the overall framework. Admittedly, modelling requires considerable amount of effort, time, and supporting data. The evaluators agreed that models are needed for assuring traceability of analyses and thus essential for the proposed model-based framework. In some cases, tolerating large model uncertainties may become inevitable. Practically, if significant uncertainties are anticipated, creating multiple scenarios to capture the range of such uncertainties would be recommended.

In addition to these, the following suggestions were made by the evaluators:

- Integration with cost modelling.
- A method to consider asynchronicity of changes, such as changes of software upgrades, which has a fast rhythm (e.g., every 6 months), and those of needs, which has a slow rhythm (e.g., periodic defence reviews, which take place every five years).
- Improvement of output visualisation, particularly when a larger number of variables involve.

Overall, no significant deficiencies were noted by the panel and thus the initial indication of the validity of the framework in terms of applicability, usability and usefulness was confirmed.

6.5 Conclusions

The entire analytical framework developed in this research was evaluated in terms of functionality, applicability, usability, and usefulness.

With an aid of a computer programme incorporating GUI and output data visualisation, the functionality of the framework was verified through an application to a hypothetical use case. The programme proved to be able to produce the intended output, which is value trajectories of Sol alternatives in multiple scenarios. Hence, the ability of the framework to provide information regarding the long-term benefit of the alternatives in differently unfolding futures was confirmed. Also, with the interactive scenario editing feature, the programme allowed to simultaneously change the parameters of scenarios and view the changing value trajectories. It successfully demonstrated the ability to enable conducting what-if analyses.

Subsequently, to evaluate the applicability, usability, and usefulness, industrial evaluation was carried out by consulting a panel of practitioners. With the knowledge regarding real-world problems of capability development, these practitioners evaluated the framework in the absence of an application to a real use case. In terms of usefulness, the validity of the framework was examined as well as the validity of its underpinning assumptions. The potential industrial impact was also examined. While some recommendations were made, the

validity of the framework and its underpinning assumptions was largely agreed. This led to the conclusion that an indication of the achievement of the overall aim of this research was confirmed within the scope of this research.

7 Summary and Conclusions

7.1 Research Summary

The overall aim of this research was formulated as follows:

Aim: Enable the evolution planning of capability components from the perspective of capability acquirer.

To achieve this aim, research activities were designed and executed. Following the Design Research Methodology (DRM) (Blessing & Chakrabarti, 2009), the research was organised into four stages: review-based research clarification (RC), review-based descriptive study-I (DS-I), comprehensive prescriptive study (PS), and initial descriptive study-II (DS-II).

In the first stage (RC), the researcher gained understanding on the current state-of-the-art regarding design practices of evolving engineering systems through a literature review. The review revealed that the current evolutionary design methodologies have only limited applicability to systems such as defence equipment. These systems can be viewed as capability components constituting capability systems, which are developed to realise capabilities to perform tasks (e.g., military missions). Individual capability components and capability systems have multiple-to-multiple relationships. The latter apply also to capability components and tasks. Hence a trade-off between tasks exists with regard to the involvement of an individual capability component. This trade-off relationship needs to be derived as a result of an investigation of the needs of the capability acquirer, who is concerned with multiple capabilities, which pertain to many different tasks fulfilled by different capability components they acquire. In the literature review, the lack of a consideration into such needs was found to be the cause of the limitation of existing methodologies. This finding led to the formulation of the aforementioned research aim.

In addition to this, to enable effective and traceable planning activity, a consistent framework which connects related concepts and relationships among them was thought to be worthwhile. Therefore, an initial objective of the research was formulated as follows:

Initial Objective: Devise a framework which aids evolution planning of capability components.

In the second DS-I stage, the literature review was expanded to gain understanding on the nature and practices regarding capability, capability systems, and capability components. Following the review, the initial research objective was refined into 3 more specific objectives, which are discussed in the following Subsections 7.1.1 to 7.1.3.

The next stage (PS) involved the development of a framework for achieving Objective 1, as well as specific analytical methods for Objectives 2 and 3. These methods were concurrently evaluated in terms of functionality by checking the output of the test programmes. Additionally, the framework responding to the Objective 1 was implemented as a computer programme (demonstration software) to be used for the test and evaluation of the research outcome as a whole.

In the final DS-II stage, the evaluation was extended to investigate the validity of the framework, as well as the overall research, against the research aim. The subject of validation included the underpinning assumptions upon which the framework was developed, and the industrial relevance. For these purposes, the developed framework was applied to a use case and demonstrated to industrial experts. These activities are summarised in Subsection 7.1.4.

7.1.1 Objective 1

In this research, an approach was taken that views evolution planning as a decision-making process. Before making a decision, the decision maker (in this case, capability acquirer) needs supporting information obtained from generation and evaluation of many evolution alternatives considering a long time horizon. In order to make sound decisions, such information needs to be generated in a consistent and traceable manner. For these reasons, the framework of this research needs to be able to produce such information systematically. This, in turn, led to the formulation of first refined research objective, as follows:

Objective 1: Devise an analytical framework which aids systematic decisions regarding the evolution of capability components over a long planning horizon in a dynamic environment.

In the literature review of this research, possibilities to adopt existing methodologies were explored for the achievement of this objective. Epoch-Era Analysis (EEA) and Value-Centric Design Methodology (VCDM) were identified methods to form the basis of the proposed framework. EEA is a methodology for the exploration of evolving System-of-Interest (Sol) alternatives in dynamic environment, while VCDM is a method for the evaluation of Sol alternatives using a value metric representing the preference of decision maker.

The research identified a need to adapt EEA and VCDM in order to be applied to evolution of capability components. Furthermore, concepts found in interrelated but separate fields, such as design, analysis, planning, and operations, needed to be bridged together to form the basis of the framework. Also, the literature review revealed the existence of discrepancies across the fields regarding the definition of the concepts, or the assumed world-view when these concepts are referred to. Thus, following the alignment, an integrated ontology was formulated for consistency. To this purpose, existing literature on capability and capability systems, mainly from the defence domain was examined. Subsequently, based on the formulated ontology and making assumptions where necessary, quantitative relationships among concepts were formalised to reflect on the analytical framework.

In EEA, factors influential on Sol's futures are classified into design, context, and needs. They are represented by design space, context space, and needs space, each of which are spanned by a set of design variables, context variables, and needs variables, respectively. The considered time period (e.g., entire lifecycle) is divided into epochs, throughout each of which context and needs remains constant. A scenario contains sequences of context vectors and needs vectors across epochs, representing one possible unfolding of the future regarding the exogenous factors. Performance model and value model provide mapping from design space to value space, which is populated by the value of Sol designs. By

pairing it with one future trajectory of Sol design, the value of Sol can be calculated along epochs, yielding a value trajectory. To enable this, context variables and needs variables need to be defined so that they can capture all anticipated changes in the future.

In line with the VCDM and Value-Focused Thinking (VFT) underpinning it, the framework was centred around the value metric to be assigned to each Sol alternative. In this case, the value could be identified according to the benefit gained from the capability acquirer's point of view. Their ultimate goal was assumed to be the realisation of higher level of capability across the whole spectrum of tasks, or higher level of organisational capability (which is the capability to fulfil the organisational task). Hierarchical breakdown of tasks from the organisational task was considered. Each task in the hierarchy corresponds to a part of the organisational capability, in line with the capability partition specified in the Capability-Based Planning (CBP) activity. This led to an assumption that the value of Sol can be defined based on its contribution to the fulfilment of organisational task. In turn, this contribution is made from aggregating the contribution of the Sol to a limited range of tasks (missions) within the organisation's task spectrum. Therefore, the value of Sol was defined as the aggregation of capability levels in its relevant, undertakable tasks. The lowest-level tasks, in which these capability levels are measured, were especially referred to as missions. Capability levels were quantified using a standardised constructed measure between 0 and 1, each of which representing the lowest acceptable level and the highest ideal level, respectively. Consequently, the resulting value metric takes a form of additive value function with an underlying value tree structure corresponding to the organisation's capability partition.

Each mission importance weight, or multiplier of each mission capability level, represents trade-off ratio of one mission's capability level over another in the value function. These are to be determined by the capability acquirer, reflecting their needs. These importance weights determine how mission capability levels are aggregated into value. In other words, they affect value through value model, which takes mission capability levels as the input. At the same time, the capability

levels can be viewed as outputs of mission models, whose input is factors related to design and context of the Sol.

The eventual output of the framework is value trajectories of alternatives of evolving Sol. One value trajectory represents the temporal change of stakeholder's perceived value of the Sol in one possible future, which is represented by a scenario. Each scenario is deterministic, i.e., it represents one possible unfolding future. Because of this, consideration of future uncertainties in the proposed framework is made on a "What-If" basis.

The proposed framework enabled the existing method of EEA to widen its applicability into particular type of Sol, which are capability components. It is underpinned by a consistent ontology based on a study into the characteristics of capabilities, capability components, and capability systems. Since it is value-centric (i.e., based on the value perceived by the capability acquirer), it allows trade-offs at various levels as low as individual design and context variables and as high as strategic-level tasks. Furthermore, since it is model-based, the traceability from the input to the output is ensured. With the aid of mathematical formalisation and models, it provides a systematic approach for exploring alternatives of Sol evolution. The output of this framework, which is value trajectory, represents the long-term benefit to the capability acquirer, and it is an important consideration for decision making regarding evolution planning of Sol, alongside cost, schedule, risk and so forth.

7.1.2 Objective 2

As described in the previous subsection, value models in the proposed framework contain importance weights of tasks (including missions) as their parameters driving the value. These importance weights need to be quantitatively determined to enable the trade-offs of capability levels between tasks, reflecting the preference of the capability acquirer in terms of the task priorities. Unfortunately, their exact values are not necessarily available for analyses and the analysts need to systematically determine them by systematically translating qualitative information. Therefore, the second research objective was formulated as follows:

Objective 2: Develop an interactive method for expressing the needs of the capability acquirer in relation to value.

A method using pairwise comparisons with ratio scale was found to be applicable to address this objective. Comparisons are carried out between every pair of tasks which share a supertask in the capability partition. These comparisons qualitatively indicate how much one task is more (or less) important than another for the accomplishment of the (parent) supertask. The comparison results are converted to the ratio scale, and subsequently translated into quantitative importance weights by applying an algorithm. The resulting importance weights, together with the hierarchical structure of the organisation's capability partition, complete a value tree indicating where the value comes from.

In response to the Objective 2, the proposed method enabled expressing the acquirer needs, in terms of different capabilities, with a set of simple qualitative statements which are systematically and consistently translated into quantitative importance weights ready for computational analysis. The method centred around ratio scale for describing trade-off is straightforward and expected to be clear to understand, as the ratio of importance weights between two tasks corresponds with the trade-off ratio of their capability levels.

The ability to account for priorities of high-level tasks was realised by the use of value hierarchy and the proposed method. This in turn ensures some flexibility in the scenario planning process of the framework. As the fundamental driver of change is often a shift in high-level (e.g., strategic level) capability needs rather than that in mission-level, the framework and the method are expected to be beneficial. Also, in real organisations, it is likely that high-level (e.g., strategic) and low-level (e.g., tactical) task priorities are determined by different group of people. In this method, results of high-level priorities flow down and affect the lower value, while trade-off relationships within low-level tasks are conserved. This characteristic is expected to be compatible with the realistic decision-making process.

For evaluation purposes, the proposed method was implemented into an interactive computer programme with a GUI (Graphical User Interface). It allowed users to specify pairwise comparison results, simultaneously viewing the visualised value hierarchy, updated in real-time. The algorithm proved to work as intended, producing a consistent value hierarchy which can be used as the basis for value models. Therefore, it can be concluded that Objective 2 was met. This part of programme, together with outcomes of other part of the research, formed an integral programme for more comprehensive analysis. Evaluation outcome of this overall framework is discussed in Section 7.1.4.

7.1.3 Objective 3

Mission models in this framework are parts of the performance model. It takes Sol's design variables and context variables as the input, outputting a set of Measures of Effectiveness (MoEs) representing the mission outcome. MoEs are subsequently mapped to mission capability levels, which are aggregated into Sol value. In order to explore design alternatives, it is necessary to be able to estimate the input-output relationship of mission models, including the impact of differences in the context. Specifically, a rapid and reasonably accurate method for understanding the trade-off relationships between design variables, or between design and context variable, was found to be suitable for the purpose of this framework. As the basis for estimation, empirical data, such as modelling and simulation outcomes, were assumed to be available.

Therefore, the third objective was formulated as follows:

Objective 3: Develop a method for trade-off analysis regarding Sol's design, which accounts for mission context.

The literature review revealed the potential applicability of the Structural Equation Modelling (SEM) technique (Loehlin, 1998) for this objective. It is a statistical method which allows to identify the most likely model of the underlying relationships within a dataset, while mitigating variabilities. While it is dependent on a hypothetical model, given by the analyst, it potentially enables the obtainment of a parsimonious model, which is simple to understand and use. The

resulting model is a linear one, which describes the trade-off relationship between design or context variables in terms of the MoEs' achievement, thus enabling trade-off analysis.

Additionally, latent constructs, which are not directly measured but their existence can be hypothesised, can be included in a hypothetical model. This ability was expected to offer more flexibility in hypothesis modelling, leading to an enriched understanding on missions' underlying relationships. It is envisioned that some aspects of mission capability, such as the ones corresponding to necessary functions to be carried out as a part of the mission, can potentially be captured as latent constructs.

For the evaluation, SEM technique was applied to a dataset of combat mission simulation results. The specification of hypothetical model and data fitting was carried out with a proprietary statistical analysis software. In this research, due to the limitation of appropriate model and dataset, inclusion of latent constructs in the path diagram was not attempted. Despite this limitation, applicability of the concept of SEM to the problem was verified. Therefore, it can be concluded that Objective 3 was met. The resulting model was to be included in the analytical framework to enable exploring design alternatives under different scenarios with different contextual factors. Evaluation results of this overall framework is discussed in Section 7.1.4.

7.1.4 Evaluation

As suggested in DRM, the evaluation of the research outcome was carried out in three phases: Support Evaluation, Application Evaluation and Success Evaluation.

The first Support Evaluation phase, which pertains to the verification of the functionality of the framework and analytical methods, were carried out concurrently and iteratively with their formulation in the PS research stage. Using hypothetical use cases and the software implementation of the framework, this evaluation was carried out by the researcher.

In the second, Application Evaluation phase, the usability and applicability of the developed framework as a whole were investigated. A use case of a realistic (but not real) defence capability development problem was developed for this purpose. It was kept simple, yet some necessary characteristics of capability evolutions were included. These characteristics include multiple-to-multiple relationships of capability systems and capability components, changing context affecting mission outcomes, trade-off between missions and tasks, and changes in high-level capability needs. Multiple scenarios were created, and evolution alternatives were evaluated against each of them to allow what-if analysis, where the strengths and weaknesses of alternatives in differently unfolding futures were investigated. Results were visually presented in the form of plots representing the end-to-end result from individual design variables to temporal value trajectories. The results successfully indicated the dependence of preferred alternatives on future scenarios, allowing exploration of alternatives based on a value metric defined from the capability acquirer's perspective. This corresponded with the overall research aim, and hence its fulfilment was verified.

The third Success Evaluation phase was carried out to investigate the usability and usefulness of the framework, which pertain to the ability to deliver the intended benefit to the intended user. The validity of assumptions underpinning the overall framework, which is prerequisite for its successfulness, was also investigated as a part of this phase. For this purpose, an evaluation session to consult a panel of practitioners with relevant expertise from a major UK defence and aerospace company was organised. Following a presentation and a demonstration using the use case developed for the Application Evaluation, their opinions were collected using a questionnaire and discussions. The feedback indicated that the framework and methods developed in the research were valid as well as the underpinning assumptions: That is, the information produced by the framework contributes to better planning decisions regarding evolving capability components.

7.2 Research Findings

The findings from the research are summarised in this section:

1. The proposed framework incorporates concepts from different fields including engineering system design, capability planning, and operational analysis. However, the literature review found that the terms and concepts used in these areas are not necessarily well-aligned. Specifically, the multitude of definitions of the concepts, as well as that of “world-views” on which the use of a term is based, posed a challenge. Dealing with this challenge necessitated an aligned vocabulary and consistent ontology to be developed in this research and to be used throughout the framework.
 - A solution for realising a capability to perform a task can be viewed as a capability system, which is constituted by different capability components, spanning people, processes, equipment, and organisation. Because of the multiple-to-multiple relationship between capability components and capability systems, design of individual capability components needs to account for multiple capability systems they are involved in, as well as tasks (missions) which these capability systems perform.
2. Based on the principles of VCDM, decision analysis and EEA, a framework for exploration of evolution alternatives of capability components from the capability acquirer’s perspective was formulated. The associated findings include:
 - Assuming that the capability acquirer’s needs can be identified with the capability to perform their organisational tasks, the needs satisfaction was found to be quantifiable as the degree of contribution to the organisational task fulfilment. This quantity can serve as the Sol’s value metric, which represents the preference of the capability acquirer.
 - By considering importance-weighted capability partition of the organisational tasks, this value metric can be correspondingly structured into an additive value function. The importance weights give trade-off ratio between two tasks (including missions), and they need to be determined reflecting the acquirer’s needs.

- The value metric enabled alternatives comparison of multi-role systems, which contribute to the organisation's task fulfilment in multiple task (mission) areas. Value trade-off between tasks enabled balancing of performance between tasks in accordance with the acquirer's needs, preventing over-specialisation in one task.
 - Missions, which are elements of system operation, were found to be appropriate individual components within the value function (termed *objectives* in the terminology of decision analysis). They serve as the building blocks of the organisational task, which collectively represent the spectrum of operations which the organisation anticipates. Also, the range of the Sol's undertakable operation is represented by a set of missions, which is a subset of the organisational task. The acquirer's needs are defined top-down to this mission level, while capability systems are engineered for each mission.
 - The hierarchical structure of the capability acquirer's needs was expressed with a value hierarchy. The hierarchy corresponds to the capability partition, which is defined in the capability planning activity of the organisation. It enabled a rich description of capability needs, which are expressed at various levels from high strategic level to low tactical mission level.
 - Scenario planning method in EEA allowed modelling possible futures exogenous to Sol by parametrising the context and needs. Changes are modelled as state transitions, while uncertainties are collectively represented by a set of scenarios.
3. The semi-quantitative pairwise comparison method, which was developed for expressing the acquirer's needs in a particular situation through determining importance weights of tasks (including missions), served its purpose. Relative ratio scale directly corresponds to the value trade-off, making it clear for the users. It also enabled expressing individual shifts in

needs concisely as changes in relative ratings, making it suitable for scenario planning in EEA.

4. The SEM-based method for analysing mission outcomes successfully estimated the underlying trade-off structure of missions regarding Sol design variables and context variables. Integrated into the overall framework, the simplicity of the resulting models enabled rapid calculation and found to be suitable for interactive analyses. Theoretically, latent constructs can be included into its hypothesis models, and it potentially offer more flexibility, although it has not been demonstrated yet.
5. The value-focused approach enabled accounting for trade-off relationships at various levels, in terms of performance and value. This allows the exploration of wide range of capability system alternatives, either materiel or non-materiel solutions, without narrowing down to particular ones prematurely.
6. The overall framework enabled modelling various changes which a Sol may encounter during its lifecycle. These changes include the acquirer's needs from high (e.g., strategic) level to low (e.g., mission) level, as well as the Sol's context which affects mission accomplishment. Impact of these changes can be traced end-to-end from the cause to the effect in value, in a chain of trade-off models. This traceability is expected to facilitate sound decision-making practices.
7. The proposed framework employed a "what-if" based approach using deterministic scenarios to analyse future uncertainties. This approach is less formal compared to probability-based methods, where every individual uncertainty is assigned a probability. The what-if approach is not capable of identifying the "best" alternatives under uncertainties either. Despite those limitations, it was demonstrated that the proposed framework was capable of considering a wide range of different possibilities expressed with scenarios. It could provide useful information for decision making on a "what-if" basis, while interpretation of the results in terms of the likelihood of each scenario is left to the analysts.

7.3 Research Contributions

The specific contributions to the academic knowledge and the relevant practices made in this research are summarised as follows:

- Formulated was an analytical framework for exploring alternatives of evolving capability components from the perspective of capability acquirer. For this purpose, concepts from different fields (design, analysis, planning, and operations) were put together, to form a consistent, quantitative framework. Through the evaluation, its usefulness as a tool to produce information for traceable decision making was demonstrated and verified. It contains the following new developments:
 - While the general principle of VCDM was adopted, a specific value model for capability components was formulated from the capability acquirer's perspective. In this model, the contribution which an Sol makes to the organisational capability is quantified on the premise that the goal of the capability acquirer is the maximisation of the organisational capability (to perform their task). The concept of capability partition was adopted from capability planning practices to specify the task area where the contribution is made within the spectrum of organisational tasks.
 - The concept of value hierarchy was adopted from decision analysis and linked with the capability partition, on the premise that the high-level tasks are fulfilled through the fulfilment of their subtasks. The levels of task fulfilment serve as the individual objectives in the value hierarchy.
 - A specific method to reflect the acquirer's needs on the importance weights of each task was devised by adopting the semi-quantitative pairwise comparison method with a ratio scale.
 - While the principle of EEA was adopted, factors constituting scenarios (context and needs) were organised so that the changes which the Sol may encounter during its lifecycle are appropriately modelled.

- Mission models, which take Sol design and context as the input, were introduced as parts of the performance model in EEA. This enabled capturing the impact of the differences in exogenous factors, which are characterised by the context part of scenarios, on mission fulfilment.
- A specific method to understand the trade-off relationships within mission models based on empirical data was devised by adopting the SEM methodology.
- The link between individual system design and high-level capability planning was established by the framework with consistent definition of concepts, composition / decomposition structures, and mapping relationships.
 - From the perspective of individual system development practices, one can easily consider the implication of different Sol alternatives on the high-level capabilities. This facilitates more efficient and more thorough concept generation, as well as requirements definition.
 - From the perspective of capability planning practices, investigation of different solution alternatives can be conducted with greater details, such as individual design features.
- The resulting framework enabled model-based and value-based analysis of evolving Sol. To this purpose, an end-to-end model is constructed to directly link individual design and context variables with the value trajectory, which indicates the desirability of the Sol in one possible future. In summary, the extension and integration of a number of existing concepts and techniques, which were formerly fragmented, enabled the formulation of the novel framework. This outcome constitutes the primary academic contribution made by this research.
- The proposed framework accounts for the multitude of operational uncertainties, which capability systems are expected to encounter in a long planning horizon. These uncertainties within a mission are expressed

as context variables, while the uncertainties across multiple missions are expressed as mission importance weights.

- From the specific application viewpoint, the proposed framework enabled alternatives exploration of multi-role systems, which is a common characteristic of capability components. This was achieved by employing VFT and by considering the fundamental acquirer's needs. Appropriate balance of performances among missions needs to be derived from the acquirer's needs. For this reason, a value model which reflects these needs was proposed.

7.4 Future Work

As a result of the research and its evaluation, the following avenues for further work have been identified:

- Due to the scarce availability of data, the evaluation of the research was conducted with a demonstration against hypothetical data with limited complexity. While it sufficiently proved the validity of concepts and the usability of the framework, further evaluation efforts with more comprehensive data are preferable for the investigation in terms of applicability. Specifically for this purpose, data representing more realistic defence capability planning problems is needed. Also, as this framework was developed to be generic, application to capability domains outside defence is also preferable.
- To ensure that the capabilities realised by solution alternatives are sufficiently captured in the framework, a methodology for the preparation of mission models needs to be researched further. The problem spans both mission outcome (i.e., MoEs) prediction and subjective scoring based on the outcome (i.e., Their aggregation into capability level defined between 0 and 1), which were assumed available in this research. The appropriate modelling method is expected to be dependent on the application.
- While the scenario-based method of the current framework can provide information for decision-making on a "What-If" basis, it is expected to be

more beneficial if uncertainty is explicitly modelled in the framework. Therefore, a method for reasonably weighting scenarios with probability, reflecting the future forecast, would be preferable especially when the number of scenarios is prohibitively large for what-if analysis. This will enable transforming value to utility, which has a formal theory regarding decision makings in uncertainty.

- The framework is suggested to be integrated with a cost model, which outputs the cost of a given design. For exploring alternatives regarding long period, the cost model should include not only initial development and production cost, but also redesign and implementation of upgrades. This is very important for practical acquisition decisions, as virtually no decision can be made without accounting for the cost aspect, which contrasts with the benefit aspect studied in this research. Additionally, if a metric for measuring long-term cost-to-benefit of evolving systems can be defined, it may serve as the ultimate objective metric for the evolution exploration.

REFERENCES

- Abbas, A. E., & Cadenbach, A. H. (2018). On the Use of Utility Theory in Engineering Design. *IEEE Systems Journal*, 12(2), 1129–1138. <https://doi.org/10.1109/JSYST.2016.2602562>
- Allison, D. L., & Kolonay, R. (2018). Expanded MDO for effectiveness based design technologies: EXPEDITE program introduction. 2018 *Multidisciplinary Analysis and Optimization Conference*. <https://doi.org/10.2514/6.2018-3419>
- Anteroinen, J. (2012). Integration of existing military capability models into the comprehensive capability meta-model. *SysCon 2012 - 2012 IEEE International Systems Conference, Proceedings*, 156–162. <https://doi.org/10.1109/SYSCON.2012.6189479>
- Anteroinen, J. (2013). Enhancing the development of military capabilities by a systems approach. *Department of Military Technology*, 33, 112. <https://www.doria.fi/handle/10024/91490>
- Bakken, B. E. (2003). *Handbook on long term defence planning*. North Atlantic Treaty Organization, Research and Technology Organization. <https://doi.org/10.14339/RTO-TR-069>
- Biltgen, P. T. (2007). *A Methodology for Capability-Based Technology Evaluation for Systems-of-Systems*. Georgia Institute of Technology. <http://hdl.handle.net/1853/14520>
- Biltgen, P. T., & Mavris, D. N. (2007). Capability-based quantitative technology evaluation for systems-of-systems. 2007 *IEEE International Conference on System of Systems Engineering, SOSE*. <https://doi.org/10.1109/SYSE.2007.4304300>
- Blessing, L. T. M., & Chakrabarti, A. (2009). DRM, a design research methodology. In *DRM, a Design Research Methodology*. Springer London. <https://doi.org/10.1007/978-1-84882-587-1>

- Brown, O. C., Eremenko, P., & Collopy, P. D. (2009). Value-centric design methodologies for fractionated spacecraft: Progress summary from phase 1 of the DARPA System F6 program. *AIAA Space 2009 Conference and Exposition*. <https://doi.org/10.2514/6.2009-6540>
- Collopy, P. D. (2009). Aerospace system value models: A survey and observations. *AIAA Space 2009 Conference and Exposition*. <https://doi.org/10.2514/6.2009-6560>
- Collopy, P. D., & Hollingsworth, P. M. (2012). Value-Driven Design. *Journal of Aircraft*, 48(3), 749–759. <https://doi.org/10.2514/1.C000311>
- Dahmann, J., Rebovich, G., Lane, J., Lowry, R., & Baldwin, K. (2011). An implementers' view of systems engineering for systems of systems. *2011 IEEE International Systems Conference, SysCon 2011 - Proceedings*, 212–217. <https://doi.org/10.1109/SYSCON.2011.5929039>
- Dahmann, J. S. (2015). Systems of Systems Characterization and Types. In *Systems of Systems Engineering for NATO Defence Applications (STO-EN-SCI-276)*.
- Davis, P. K. (2002). *Analytic Architecture for Capabilities-Based Planning, Mission-System Analysis, and Transformation*. https://www.rand.org/pubs/monograph_reports/MR1513.html
- de Weck, O., Eckert, C. M., & Clarkson, P. J. (2007). A Classification of Uncertainty for Early Product and System Design. *DS 42: Proceedings of ICED 2007, the 16th International Conference on Engineering Design*, 159–160. <https://www.designsociety.org/publication/25407/A+Classification+of+Uncertainty+for+Early+Product+and+System+Design>
- de Weck, O. L., Ross, A. M., & Rhodes, D. H. (2012). Investigating Relationships and Semantic Sets amongst System Lifecycle Properties (Ilities). *Third International Engineering Systems Symposium CESUN*, 1–12. <http://hdl.handle.net/1721.1/102927>

- Fitzsimmons, M. (2007). Whither Capabilities-based Planning? *Joint Force Quarterly*, 1(44), 101–105. <https://apps.dtic.mil/sti/citations/ADA517897>
- Guenov, M. (2008). Covariance structural models of the relationship between the design and customer domains. *Journal of Engineering Design*, 19(1), 75–95. <https://doi.org/10.1080/09544820701213378>
- Guenov, M. D., Barker, S. G., Hunter, C., Horsfield, I., & Coleridge-Smith, N. (2006). An integrated approach to customer elicitation for the aerospace sector. *Systems Engineering*, 9(1), 62–72. <https://doi.org/10.1002/SYS.20045>
- Hancock, G. (2012). Experimental, Quasi-Experimental, and Nonexperimental Design and Analysis With Latent Variables. *The SAGE Handbook of Quantitative Methodology for the Social Sciences*, 318–335. <https://doi.org/10.4135/9781412986311.N17>
- Harper, D. J. (2021). Effectiveness-based design as an important part of the conceptual digital twin: Observations from afrl's expedite program. *AIAA Scitech 2021 Forum*, 1–16. <https://doi.org/10.2514/6.2021-1353>
- Hastings, D., & McManus, H. (2004). *Space System Architecture: Final Report of SSPARC: the Space Systems, Policy, and Architecture Research Consortium (Thrust I and II)*. <https://dspace.mit.edu/handle/1721.1/7518>
- Hazelrigg, G. A. (1998). A Framework for Decision-Based Engineering Design. *Journal of Mechanical Design*, 120(4), 653–658. <https://doi.org/10.1115/1.2829328>
- Henshaw, M., Lister, P., Harding, A. D., Kemp, D., Daw, A. J., Farncombe, A., & Touchin, M. (2011). Capability Engineering – An Analysis of Perspectives. *INCOSE International Symposium*, 21(1), 712–727. <https://doi.org/10.1002/J.2334-5837.2011.TB01238.X>
- Henshaw, M., Siemieniuch, C., Sinclair, M., Barot, V., Henson, S., Ncube, C., Lim, S. L., Dogan, H., Jamshidi, M., & DeLaurentis, D. (2013). *The Systems of Systems Engineering Strategic Research Agenda* (Issue 2).

https://www.researchgate.net/publication/316688269_The_Systems_of_Systems_Engineering_Strategic_Research_Agenda

Hill, R. R., & Miller, J. O. (2017). A history of United States military simulation. *Proceedings - Winter Simulation Conference*, 346–364. <https://doi.org/10.1109/WSC.2017.8247799>

Huang, Q., Zhang, Y., Zhang, B., Ouyang, S., Qi, X., Liu, X., & Zhai, D. (2021). Emerging SEM equipment system combat capability assessment method. *Procedia Computer Science*, 183, 545–550. <https://doi.org/10.1016/J.PROCS.2021.02.095>

Hupman, A., Abbas, A. E., Tibor, E., Kannan, H., Bloebaum, C., & Mesmer, B. (2015). Calculating value gaps induced by independent requirements, deterministic modeling, and fixed targets. *56th AIAA/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference*. <https://doi.org/10.2514/6.2015-1363>

International Council on Systems Engineering. (2023). *Systems Engineering Handbook: A Guide for System Life Cycle Processes and Activities* (D. D. Walden, T. M. Shortell, G. J. Roedler, B. A. Delicado, O. Mornas, Y.-S. Yip, & D. Endler, Eds.; Fifth Edition). John Wiley & Sons, Inc.

ISO/IEC/IEEE. (2017). *ISO/IEC/IEEE 24765-2017 - Systems and software engineering--Vocabulary*. <https://doi.org/10.1109/IEEESTD.2017.8016712>

ISO/IEC/IEEE. (2019). *ISO/IEC/IEEE 21839-2019 - Systems and software engineering -- System of systems (SoS) considerations in life cycle stages of a system*. <https://doi.org/10.1109/IEEESTD.2019.8767116>

ISO/IEC/IEEE. (2023). *ISO/IEC/IEEE 15288-2023 - Systems and software engineering - System life cycle processes*. <https://doi.org/10.1109/IEEESTD.2023.10123367>

JCS USDoD. (n.d.). *Universal Joint Task List*. Retrieved 28 December 2023, from <https://www.jcs.mil/Doctrine/Joint-Training/UJTL/>

- Joint Systems Analysis Group Technical Panel 3. (n.d.). *Guide to Capability-based Planning*. Retrieved 28 December 2023, from <https://www.hsdl.org/c/abstract/?docid=461818>
- Juzgado, P. D. B., & Bonnema, G. M. (2008). On the Origin of Evolvable Systems; Evolvability or Extinction. *7th International Symposium on Tools and Methods for Concurrent Engineering, TMCE 2008*. <https://research.utwente.nl/en/publications/on-the-origin-of-evolvable-systems-evolvability-or-extinction>
- Keeney, R. L. (2008). Applying Value-Focused Thinking. *Military Operations Research*, 13(2), 7–17. <http://www.jstor.org/stable/43920754>
- Kemp, D., & Daw, A. (2014). *Capability Systems Engineering Guide*.
- Kujawski, E. (2014). Interaction Effects in the Design of Computer Simulation Experiments for Architecting Systems-of-Systems. *Systems Engineering*, 17(4), 426–441. <https://doi.org/10.1111/SYS.21280>
- Lea, D., & Bradbery, J. (2020). *Oxford Advanced Learner's Dictionary* (Tenth). Oxford University Press.
- Loehlin, J. C. (1998). *Latent Variable Models: An introduction to factor, path, and structural analysis* (3rd ed.). Lawrence Erlbaum Associates Publishers.
- Maier, M. W. (1998). *Architecting Principles for Systems-of-Systems*. [https://doi.org/10.1002/\(SICI\)1520-6858\(1998\)1:4](https://doi.org/10.1002/(SICI)1520-6858(1998)1:4)
- Odu, G. O. (2019). Weighting methods for multi-criteria decision making technique. *Journal of Applied Sciences and Environmental Management*, 23(8), 1449–1457. <https://doi.org/10.4314/JASEM.V23I8.7>
- OUS(D(R&E) USDoD. (2020). *Mission Engineering Guide*. https://ac.cto.mil/wp-content/uploads/2020/12/MEG-v40_20201130_shm.pdf
- Papageorgiou, E. (2018). Development and application of a value driven design assessment framework to an unmanned air system design. *2018 Advances*

- in Science and Engineering Technology International Conferences, ASET 2018*, 1–6. <https://doi.org/10.1109/ICASET.2018.8376805>
- Papageorgiou, E., Eres, M. H., & Scanlan, J. (2016). Value modelling for multi-stakeholder and multi-objective optimisation in engineering design. *Journal of Engineering Design*, 27(10), 697–724. <https://doi.org/10.1080/09544828.2016.1214693>
- Parnell, G. S., & Tani, S. N. (2013). Foundations of Decision Analysis. In *Handbook of Decision Analysis* (pp. 46–67). John Wiley and Sons. <https://doi.org/10.1002/9781118515853.CH3>
- Piplani, L. K. (1994). *Systems Acquisition Manager's Guide for the Use of Models and Simulations*. <https://apps.dtic.mil/sti/citations/ADA285573>
- Plotly. (n.d.-a). *Dash Documentation & User Guide*. Retrieved 28 December 2023, from <https://dash.plotly.com/>
- Plotly. (n.d.-b). *Plotly Open Source Graphing Libraries*. Retrieved 28 December 2023, from <https://plotly.com/graphing-libraries/>
- Python. (2022). *Python Release Python 3.9.16*. <https://www.python.org/downloads/release/python-3916/>
- Ricci, N. (2014). *Dynamic system perspective for design : ility-driving elements as responses to uncertainty*. <https://dspace.mit.edu/handle/1721.1/90789>
- Ricci, N., Schaffner, M. A., Ross, A. M., Rhodes, D. H., & Fitzgerald, M. E. (2014). Exploring Stakeholder Value Models via Interactive Visualization. *Procedia Computer Science*, 28, 294–303. <https://doi.org/10.1016/J.PROCS.2014.03.037>
- Roberts, C. J., Richards, M. G., Ross, A. M., Rhodes, D. H., & Hastings, D. E. (2009). Scenario planning in dynamic multi-attribute tradespace exploration. *2009 IEEE International Systems Conference Proceedings*, 366–371. <https://doi.org/10.1109/SYSTEMS.2009.4815828>

- Roedler, G. J., Martin, L., & Jones, C. (2005). *Technical Measurement. A Collaborative Project of PSM, INCOSE, and Industry*.
<https://apps.dtic.mil/sti/citations/ADA605916>
- Ross, A. M. 1977-. (2006). *Managing unarticulated value : changeability in multi-attribute tradespace exploration*.
<https://dspace.mit.edu/handle/1721.1/35089>
- Ross, A. M., Gregory O'Neill, M., Hastings, D. E., & Rhodes, D. H. (2010). Aligning perspectives and methods for value-driven design. *AIAA SPACE Conference and Exposition 2010*. <https://doi.org/10.2514/6.2010-8797>
- Ross, A. M., & Hastings, D. E. (2005). The Tradespace Exploration Paradigm. *INCOSE International Symposium*, 15(1), 1706–1718.
<https://doi.org/10.1002/J.2334-5837.2005.TB00783.X>
- Ross, A. M., & Hastings, D. E. (2006). Assessing changeability in aerospace systems architecting and design using dynamic multi-attribute tradespace exploration. *Collection of Technical Papers - Space 2006 Conference*, 1, 551–568. <https://doi.org/10.2514/6.2006-7255>
- Ross, A. M., & Rhodes, D. H. (2008). Using natural value-centric time scales for conceptualizing system timelines through epoch-era analysis. *18th Annual International Symposium of the International Council on Systems Engineering, INCOSE 2008*, 4(1), 2250–2264.
<https://doi.org/10.1002/j.2334-5837.2008.tb00871.x>
- Saaty, R. W. (1987). The analytic hierarchy process—what it is and how it is used. *Mathematical Modelling*, 9(3–5), 161–176. [https://doi.org/10.1016/0270-0255\(87\)90473-8](https://doi.org/10.1016/0270-0255(87)90473-8)
- Saaty, T. L., & Vargas, L. G. (2012). *Models, Methods, Concepts & Applications of the Analytic Hierarchy Process* (Vol. 175). Springer US.
<https://doi.org/10.1007/978-1-4614-3597-6>
- Taliaferro, A. C., Gonzalez, L. M., Tillman, M., Ghosh, P., Clarke, P., & Hinkle, W. (2019). *A Guide to Capability-Based Planning (CBP)* (Defense

- Governance and Management: Improving the Defense Management Capabilities of Foreign Defense Institutions). https://www.ida.org/research-and-publications/publications/all/d/de/defense-governance-and-management_improving-the-defense-management-capabilities-of-foreign
- Telford, J. K. (2007). A brief introduction to design of experiments. *Johns Hopkins APL Technical Digest (Applied Physics Laboratory)*, 27(3), 224–232.
- The United States Air Force. (2019a, April). *F-15 Eagle Fact Sheet*. <https://www.af.mil/About-Us/Fact-Sheets/Display/Article/104501/f-15-eagle/>
- The United States Air Force. (2019b, April). *F-15E Strike Eagle Fact Sheet*. <https://www.af.mil/About-Us/Fact-Sheets/Display/Article/104499/f-15e-strike-eagle/>
- The United States Air Force. (2021, March 11). *Air Force receives first F-15EX*. <https://www.af.mil/News/Article-Display/Article/2534008/air-force-receives-first-f-15ex/>
- The United States Air Force. (2022, August). *F-22 Raptor Fact Sheet*. <https://www.af.mil/About-Us/Fact-Sheets/Display/Article/104506/f-22-raptor/>
- UK Ministry of Defence. (2005). *The Acquisition Handbook* (6th ed.).
- United States Department of Defense. (2013). *Joint Publication 1: Doctrine for the Armed Forces of the United States, March 25, 2013, Incorporating Change 1, July 12, 2017*.
- van Heerden, A. S. J. (2018). *Computational techniques for aircraft evolvability exploration during conceptual design*. <https://dspace.lib.cranfield.ac.uk/handle/1826/18615>
- Van Heerden, A. S. J., Guenov, M. D., Molina-Cristóbal, A., Riaz, A., & Bile, Y. (2018). Combined airframe and subsystems evolvability exploration during conceptual design. *31st Congress of the International Council of the Aeronautical Sciences, ICAS 2018*.

- Walker, S. K. (2005). *Capabilities-Based Planning - How it is Intended to Work and Challenges to its Successful Implementation*.
<https://apps.dtic.mil/sti/citations/ADA434864>
- Webb, N. J., Richter, A., & Bonsper, D. (2010). Linking defense planning and resource decisions: A return to systems thinking. *Defense and Security Analysis*, 26(4), 387–400. <https://doi.org/10.1080/14751798.2010.534647>
- Yue, Y., & Henshaw, M. (2009). An Holistic View of UK Military Capability Development. *Defense & Security Analysis*, 25(1), 53–67.
<https://doi.org/10.1080/14751790902749900>

APPENDICES

Appendix A Preliminary Information for Evaluators

Abstract

Systems such as military equipment are expected to constitute the capability of their customer organisation, alongside other cooperating systems and capability components (e.g., personnel). Intended for accomplishing the tasks to be fulfilled by the organisation, capability is developed through continuous planning and engineering. Individual systems need to be developed in this context, so that their value to the customer is sustained throughout their lifetime.

To enable this, there is a need for individual system design to be better aligned with wider capability development activities. Regarding the exploration of design alternatives of systems-of-interest (Sols), two research areas were identified:

- Design alternatives need to be evaluated using a metric which more directly reflects their contribution to the organisational capability. It needs to reflect the priority of the customer among different capabilities, and/or tasks.
- Considering the long-time horizons of the system lifecycle and the overarching capability planning, there is a need for exploring the evolution of the design of Sols. This needs to be done considering external changes and uncertainties during the planning horizon. Included in these changes and uncertainties are those related to customer needs and (the Sols') contextual factors, as shown in Figure 1.

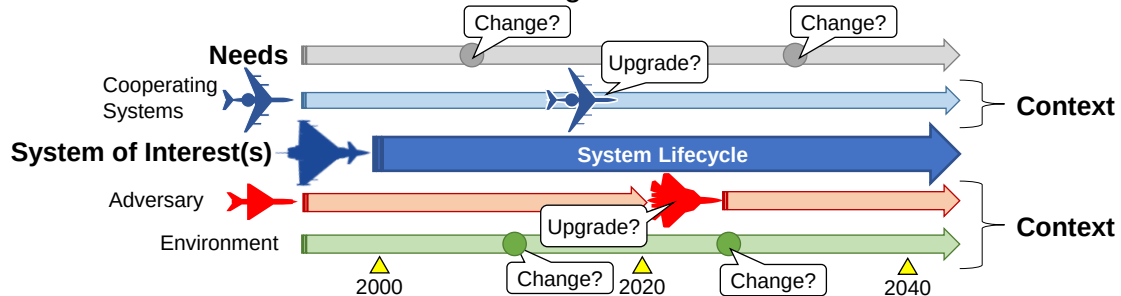


Figure 1 Changes and Uncertainties surrounding Sols

To address the research challenges identified above, a framework was developed for value-based planning of Sols design evolution within the context of capability development. A value-based approach based on a formally defined value function enables evaluating alternatives based on a metric which adequately reflects the customer needs. A value function is devised to represent the degree of contribution of Sols to the organisational capability. The premise is that the extent of this contribution is the value to the customer which the Sols provide. For planning of design evolution, the method of scenario planning is employed. It enables accounting for different possible futures using scenarios with codified change trajectories of customer needs and contextual factors. In order to integrate system design exploration, value-based evaluation, and scenario planning, a consistent domain ontology is formulated. In this ontology, the contribution Sols make is measured as an aggregation of realised capability

levels in their undertakable missions (Figure 2). Missions are derived from breakdown structure of organisational tasks and weighted in accordance with customer needs.

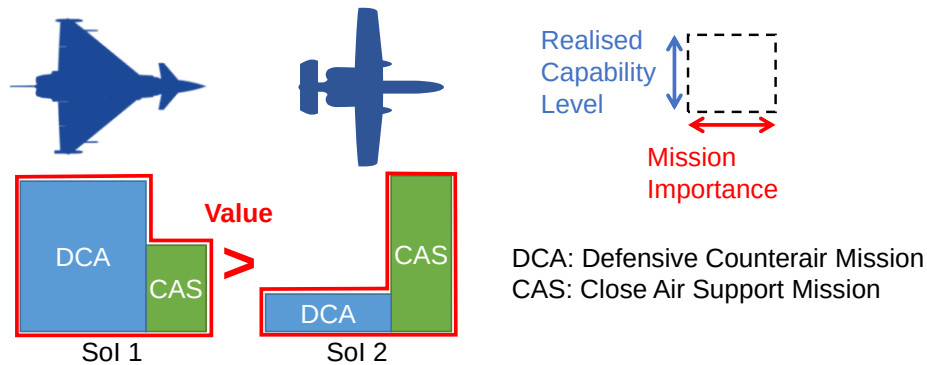


Figure 2 Example of Value of Different Sol Alternatives

Additionally, two specific methods to be used within the framework were developed. The first method identifies the customer needs by traversing the task breakdown structure. Needs at a particular moment of the time horizon (e.g., periodic Defence Reviews) are described by relative importance weights of requisite tasks.

The second method creates mission models based on empirical data, and enables to estimate the measures of effectiveness (MoEs) of a given design, in a given context. These MoEs indicate how well a capability is realised, which affects value. The input-output dataset is assumed to be obtainable through combat simulation. Factors related to other systems (cooperating or adversary) and non-materials (e.g., training level of personnel) can be included in the input as a part of context.

For evaluation purposes, the framework is implemented in an interactive computer program with a graphical interface. It is intended to demonstrate the ability of the framework to conduct what-if analysis against different future scenarios.

The proposed framework is expected to be applicable to practical problems regarding capability development and system design. It enables the forecast of Sols value under future uncertainties. This, in turn, aids gaining knowledge useful for making decisions. Ultimately, this will contribute to the achievement of acquirers' and suppliers' aims for securing long-term benefit.

Outline of the Research

1. Introduction

Organisations, such as armed forces, have **tasks** for which they ought to be ready. Some of these **organisational tasks**, such as anticipated armed conflicts, do not necessarily materialise. Nevertheless, the organisations are required to develop and maintain requisite **capabilities** for organisational task accomplishment. The means for realising these capabilities can be referred to as **solutions**. These are constituted by various capability components spanning

across materiel and non-materiel (commonly referred to as TEPIDOIL¹ or DOTMLPF²) including engineering systems and system-of-systems. In principle, as long as sufficient capability is provided, any combination of these components can be a solution.

“Capability-based” approaches for organisations, including Capability-Based Planning (CBP) (Taliaferro et al., 2019), capability engineering (Henshaw et al., 2011), provide a holistic and outcome-oriented approach for developing solutions across capability components. In such a context, the design of individual systems may be directed toward contribution to capabilities, rather than on-its-own performance for limited tasks. Specifically, if different design alternatives could be comparatively assessed based on such a contribution, then a more effective capability planning and system design would become possible by more directly adhering to the customer needs.

For this purpose, the relationship between the system design and the realised capabilities, which are measured by fulfilment of organisational tasks, needs to be clarified. Organisations certainly need “good” systems, which can contribute much to fulfilling their ultimate organisational tasks (e.g., “National defence”), but the relations between the means and the goal are often too complex. An approach to fill this gap needs to be developed.

Furthermore, consideration of long-term future of systems is also needed. Obviously, ensuring that the system can continuously provide benefit to the customer organisation is crucial. Also, for larger capability constructs (e.g., system-of-systems), evolution is a major consideration which influences the future of the constituent systems. Therefore, if analysis of system lifecycle can be integrated with the larger capability context, it would further contribute to the effectiveness of capability planning, as mentioned above.

The rest of the document is organised as follows. The main concepts behind the research are defined to form an ontology in section 2. The scope of the research is explained in section 3. Outlined in sections 4 and 5 are the individual approaches for models of needs and missions, respectively. In section 6, the approach for verification and validation is discussed.

2. Ontology Development

An ontology refers to “a set of concepts and categories in a subject area or domain that shows their properties and the relations between them”. As this research spans across different relevant areas of practices and studies, a consistent ontology is specially developed.

First, task, solution and capability are defined as follows:

Task: A clearly defined activity specifically assigned to an organisation. (Adopted from (United States Department of Defense, 2013), as cited in (OUSD(R&E) USDoD, 2020))

Solution: An organisation which performs a task.

¹ Training, Equipment, Personnel, Information, Concepts and doctrine, Organisation, Infrastructure, Logistics. UK classification known as the Defence Lines of Development (DLOD).

² Doctrine, Organization, Training, Materiel and equipment, Leadership and education, Personnel, Facilities. US classification.

Capability: The ability to complete a task under specified conditions and level of performance. (Adopted from (OUSD(R&E) USDoD, 2020))

Here, tasks are “what” to be done and solutions are “how” to do it. Solution corresponds to a task and possesses a capability to perform it with a certain level of performance, regardless of whether it is satisfactory or unsatisfactory to the customer. This level will be referred to as **capability level** of a solution.

As special instances of task, organisational tasks are defined as follows:

Organisational Tasks: A set of tasks to which the customer organisation is assigned.

These are “root” tasks from which all tasks are derived. The entire customer organisation is regarded as the solution to perform them. Also correspondingly, **organisational capability** can be considered. Throughout this framework, **needs** of the customer is assumed to be identifiable with this organisational capability. This leads to requiring some form of solution which provides such a capability. If there are two different solutions which provides the same level of capability, the customer is assumed to be indifferent between them.

Additionally, hereinafter in this framework, **systems-of-interest (Sols)** is defined to collectively refer to one or more entities whose alternatives are being analysed. It involves in performing one or more tasks by constituting a part of their solution. It can contain non-materiel components, but one can still consider their different designs. For example, the same fighter aircraft with personnel with two different skill levels can be counted as two different systems (human-machine systems). A clear system boundary separating Sols from external, uncontrollable factors is set and used throughout an analysis with a consistency.

Concepts defined in following sections also form a part of the ontology. For their definition, the reader is directed to the Appendix section.

Organisational Capability Level

Prior to considering contribution to capabilities, a generic scale for capability levels needs to be defined. Assuming that they can be quantified, a standardised number between 0 and 1 can be assigned, where:

- 0: Lowest (least desirable) acceptable level
- 1: Highest (most desirable) level.

This can be directly applied to organisational capability level. However, organisational tasks are usually too broad, and too many systems and capability components are involved in the solution (i.e., entire customer organisation). This complexity makes it not suitable for analysis at an individual system level. As systems usually contribute to different parts of the task in different ways and to different extents, task breakdown is needed to account for this. By examining “What tasks must be performed in order to perform this task?”, “children” tasks can be derived from a “parent” task. There should be a limited number of representative “children” tasks which can collectively capture the range of the “parent”. For example, from a task of “defending an island”, tasks to counter maritime and airborne threats can be derived. Universal Joint Task List (UJTL) of the United States Armed Forces (JCS USDoD, n.d.) is an example of a complete taxonomy of tasks of an organisation.

Generally, as the tasks are broken down further, solutions to perform them becomes smaller, and analysis into the relationship between solution and capability level becomes simpler. Hence, task breakdown helps analysing the contribution of Sols, which is discussed later. But to enable considering a wide range of solutions by retaining some freedom regarding the form of the solution, there is a limit in the level of decomposition. This is because a too elementary (leaf) task presupposes a particular form of solution. For example, if one decomposes air defence task further into detection task and engagement task, barrage balloons, which do not possess nor require detection capability, may not be considered as a (potential contributor to a) solution. Task at an appropriate resolution will be referred to as **missions** in this framework, and they form the lowest level of the breakdown structure of the tasks.

Since the “parent” task is performed by the aggregation of solutions performing its “children” tasks, the capability level of the “parent” task is given by a weighted average of those of “children” tasks. Here, the weights can be understood as the relative importance of one “child” task over another in performing the “parent” task (**importance weight**). In the previous example, if the military has solutions for both “countersea” and “counterair” tasks, one can assume that they have a solution for the overarching island defence task. If the countersea capability level is 0.8 and the counterair capability level is 0.4, the total island defence capability level should be somewhere between 0.8 and 0.4, determined by relative importance of the countersea task over the counter air (or vice versa).

These relationships are applicable to the mission capabilities and the organisational capabilities. When the capability level of mission i is V_i , its importance weight is W_i , and the number of missions which compose the entire organisational tasks is N , the organisational capability level V_o can be expressed as:

$$V_o = \sum_{i=1}^N W_i V_i \quad (1)$$

Where $0 \leq W_i \leq 1$ and $\sum_{i=1}^N W_i = 1$. This relationship is illustrated in Figure 1. The organisational capability level V_o is visualised as the total area of grey rectangles.

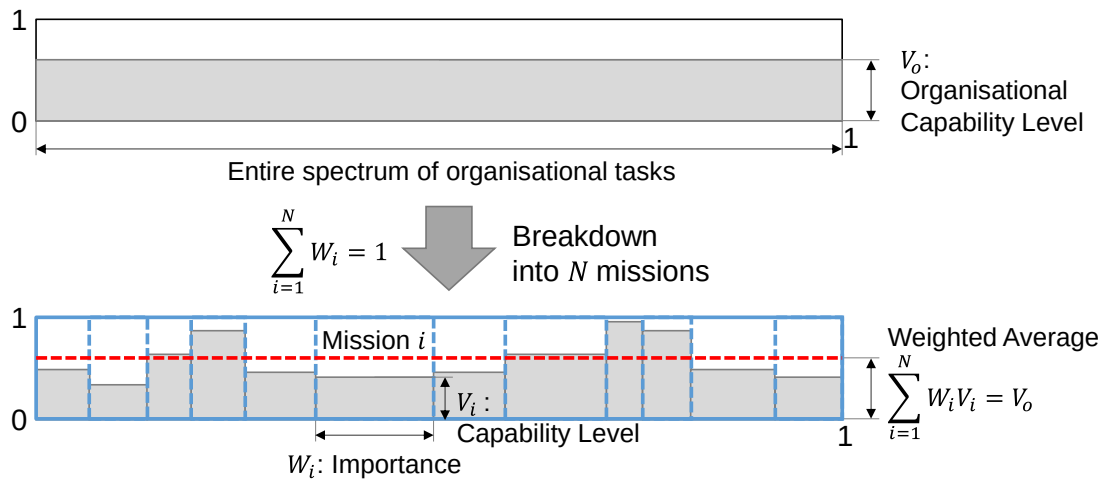


Figure 1 Breakdown of Organisational Tasks into Missions

Contribution of Sol to Organisational Capability

Sols do not necessarily contribute to all missions within the entire spectrum of organisational tasks. Relevant missions, whose capability levels can be affected by the Sols, constitutes a set M . For each mission in M , different alternatives of Sols yield different capability levels. Contribution of an alternative X to the organisational capability can be quantified by extracting relevant terms from equation (1) as follows:

$$C(X) = \sum_{i \in M} W_i V_i(X) \quad (2)$$

As M can be viewed as the undertakable mission spectrum of the Sols, alternatives can be compared within this limited spectrum. Importance weights W_i of missions are upscaled to w_i (**importance weight for Sols**) by multiplying a constant $1/\sum_{i \in M} W_i$, so that they sum to 1. The relative importance of one mission over another, which is given by a ratio of W_i to W_j , remains the same with w_i to w_j . This spectrum is useful for comparing alternatives of Sols, as the complete task breakdown structure is not necessarily available for analysis in practice. This may be due to the vastness of the entire spectrum of organisational tasks.

Within this Sols mission spectrum, the relative contribution of an alternative X to the organisational capability, $c(X)$, can be defined as:

$$c(X) = \sum_{i=1}^n w_i v_i(X) \quad (3)$$

This is because $c(X) = C(X)/\sum_{i \in M} W_i$ and $c(X)/c(Y) = C(X)/C(Y)$ always hold, for any X and Y . This allows using $c(X)$ to compare contributions made by different Sols alternatives.

An example is given in Figure 2. The contribution C and relative contribution c are visualised as total areas of rectangles, each of which representing a mission. In this case, $N = 12$, $n = 3$ and $M = \{4, 6, 9\}$. In the Sols mission spectrum, these missions are re-numbered into $\{1, 2, 3\}$: this means that $v_1 = V_4$, $v_2 = V_6$, $v_3 = V_9$.

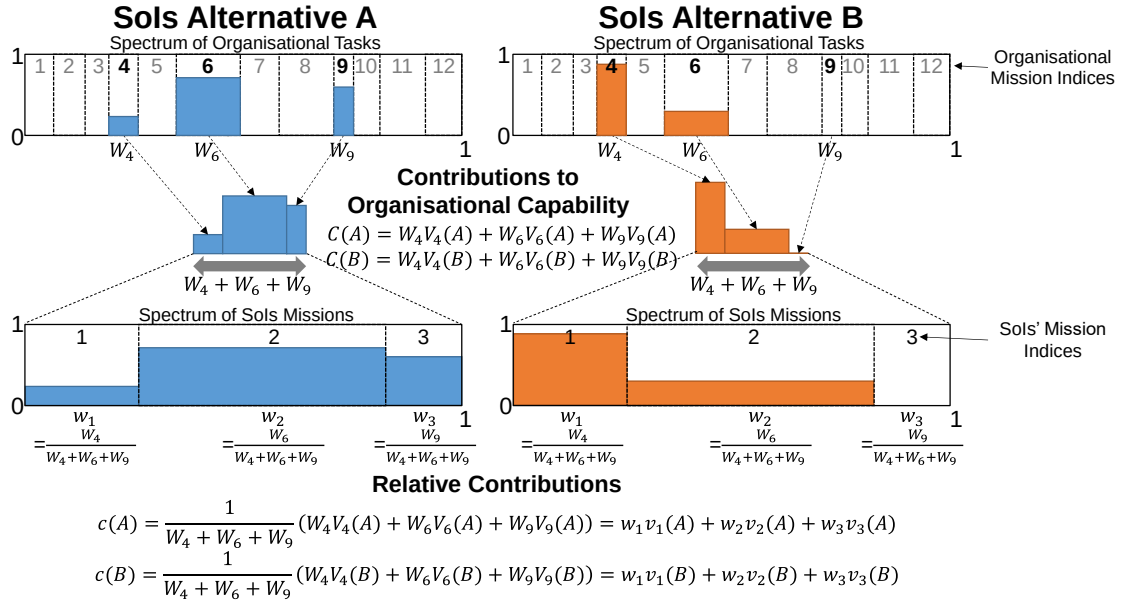


Figure 2 Calculating Relative Contribution to Organisational Capability

Defining Value Based on Contribution

To enable evaluating, comparing, and making decision between alternatives, a 1-D metric which reflects the preferences of the customer organisation is assigned to each alternative. In the terminology of design paradigms such as decision-based design, value-based design (Abbas & Cadenbach, 2018), and value-driven design (Collopy & Hollingsworth, 2012)(Ross et al., 2010), this metric is referred to as **value**. While its specific form can vary, in this framework it is made to simply represent the aggregation of benefits to the customer organisation. Since the contribution to organisational capability can be identified with such benefits, then $c(X)$ can be used as the value function $V(X)$ of an alternative X , where $v_i(X)$ serve as value attributes.

$$V(X) = c(X) = \sum_{i=1}^n w_i v_i(X) \quad (4)$$

The form of $V(X)$ implies that:

5. Sols which can realise higher capability levels (higher $v_i(X)$) are valued more highly.
6. Sols which can contribute to important missions (those with higher w_i) are valued more highly.
7. Sols which can contribute to more missions (more non-zero $v_i(X)$) are valued more highly, as some alternatives can contribute to missions which other alternatives cannot.
8. Trade-off between mission capabilities in terms of value is allowed, at a ratio of one importance weight over another ($v_i(X)/v_j(X)$).

Additionally, by considering parent tasks at the levels in between the top-level organisational task level and the mission level in the task breakdown structure, a value tree of the Sols can be formed. Specifically, missions undertakable for Sols, together with their “ancestors” tasks, are extracted from the organisational task breakdown structure. Then, the weights are recalculated so that the proportional

relationships of the weights between tasks which share the common “parent” task are preserved. In the tree, the sum of the weights of capabilities at the same level of the hierarchy is always 1. This implies that the value of the Sol can be decomposed in a hierarchical and “mutually independent, collectively exhaustive” manner, corresponding to the task breakdown. A specific method to determine weights in the Sols value tree is proposed in Section 4.

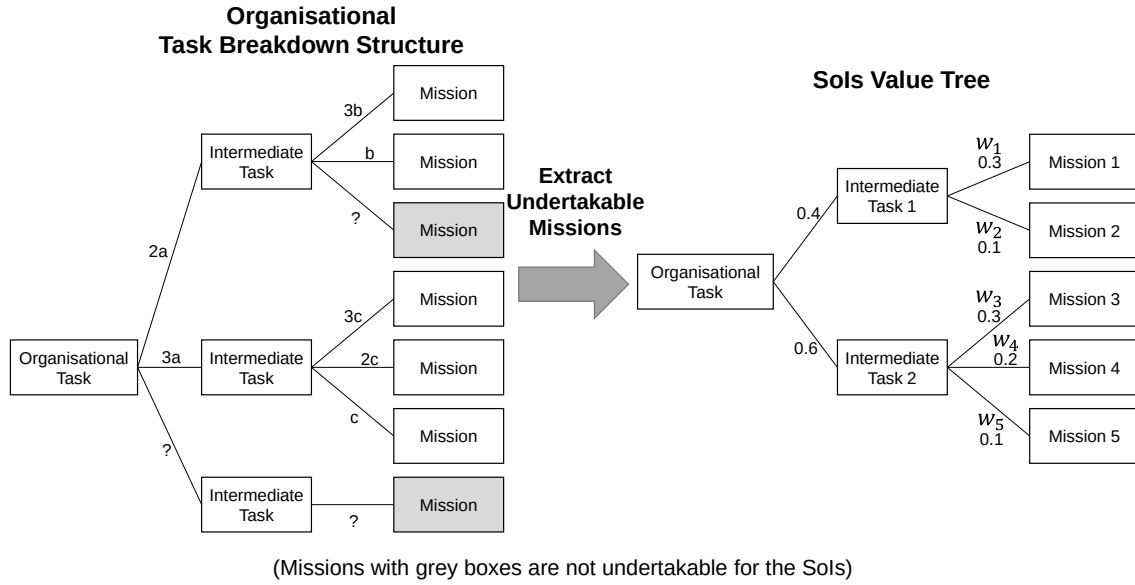


Figure 3 Formation of Sols Value Tree

It should be noted that in the context of value-driven design and related literature, establishing a link between value and monetary cost is often suggested. For example, in (Ross et al., 2010), value is defined as the “aggregation of (non-monetary) benefits relative to the monetary cost of obtaining those benefits”. That aspect is beyond the scope of this framework. To consider value for money, cost needs to be estimated outside this framework and compared with the benefit, which is given as the value.

Using the resulting value function, alternatives can be evaluated in relation to the customer needs. The function can be viewed to map the realised capability levels to value in accordance with the customer needs, which is characterised by the set of missions and their weights.

Value-Based Evaluation among Temporal Changes

In the previous section, an approach for value-based evaluation of capability-based systems was presented. To account for temporal changes including design evolutions, the proposed approach was expanded by incorporating the methodology of the Epoch-Era Analysis (EEA)(Ross & Rhodes, 2008).

During the system lifecycle, the design of Sols may evolve over time. In this framework, different designs are assumed to be characterised by a finite set of **design variables**. Each design variable represents a decision which differentiates one design from another. A design of Sols is represented by an instance of **design vector, D** .

Furthermore, the external factors surrounding the Sols are subject to changes and uncertainties. These factors can be classified into needs and context. Factors related to the value perception of the customer organisation are a part of the needs. It is characterised by the value function, specifically the set of missions and their weights. Missions can be added or removed, and their priorities can change during the lifecycle. In the form of mathematical expression, what changes over time is the function f_{value} , which specifies which v_i are relevant to the value and how they are aggregated into the value. Thus, importance weights w_i are included in this function as parameters. Therefore, Equation (4) can be rewritten as:

$$V = f_{value}(v) \quad (5)$$

Alternatively, this value function can be viewed as a part of a model of customer capability needs.

Other external factors fall under the **context**. In this framework, those which affect the realised mission capability level is of primary concern. Factors regarding adversary systems and systems cooperating with Sols are typical examples. Like the design vector D , context can be characterized as **context vector C** consisting of **context variables**.

Missions can be viewed as models which take a pair of Sol design and context as the input, and output a realised capability level. Mathematically, it is expressed using D, C and v_i .

$$v_i = f_{mission i}(D, C) \quad (6)$$

Here, it is assumed that the form of the functions is invariant over time. If there is a significant change in mission itself and applying the same function becomes inappropriate, introduction of a new mission is required. For example, ballistic missile defence can be called as a subset of air defence with higher target performance, which could be characterised with slightly different C . However, many unique elements, such as space-based sensors, make a design space considerably different from that of ordinary air defence mission. In such a case, considering separate mission model will make sense.

Changes occurring within the Sol boundary (i.e., changes of design) are expressed as transition from one D to another, while changes occurring outside the boundary (i.e., changes of context or needs) are expressed as one pair of (C, f_{value}) to another pair.

In the EEA, an **epoch** is defined as a discrete time period with fixed context and needs, which together specify an external condition to the Sols. By sequencing epochs, a **scenario** which describes one possible future awaiting the Sols can be characterised. Likewise, an evolution trajectory of the Sols design can be characterised by a sequence of their designs. By combining them, a value trajectory of the Sols is finally calculated, enabling a value-based “what-if” analysis.

However, to use this relational structure in actual analysis, w_i need to be determined to reflect the customer needs, and $f_{mission i}$ need to be known. In the following sections, methods for these are described.

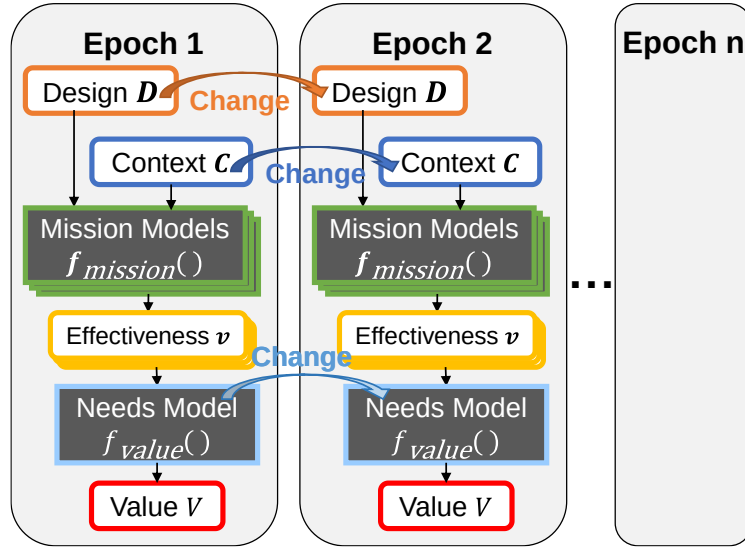


Figure 4 Domain Ontology as an Input-Output Relationships Model

3. Scope

This framework enables calculating the value trajectory of Sols with different designs, under different scenarios. Uncertainties regarding the future are considered by planning multiple scenarios, which are assumed to collectively capture the range of possibilities. Scenarios themselves are deterministic and therefore the analysis is on a “what-if” basis. Certainly, probabilities can be introduced as an addition to this framework (e.g., by assigning probability to each individual change or to each scenario). If the risk attitude of the customer can be modelled, it may also enhance the framework. However, these are kept beyond the scope of this research. While these may enable identifying the “best” options, the results are likely to be sensitive to additional assumptions regarding the future forecast, which are to be made by the analysts. As the aim of this framework is to provide information to aid decision making, rather than making decision by itself or optimising design, the framework should be still useful.

Also, in this framework, the value function is defined to solely indicate the benefit side of the Sol and the cost side is not covered. This is because cost is domain specific. By using a lifecycle cost model, and specifying a definition of cost effectiveness, the alternatives can be evaluated based on it.

Lastly, while it is assumed that the capability can be quantified as a capability level between 0 and 1, a method for doing it is excluded from the scope of this framework. This is because it is highly dependent on the mission and the customer. This problem is briefly discussed in Section 5.

4. Capability Needs Prioritisation

A needs model represents customer organisation’s value perception, which is intangible by nature. As this framework deals with capability needs, it is characterised by importance weights of tasks which are present in the value tree. The importance weight of one task is determined relatively to another task.

As a part of this framework, a method based on pairwise comparisons with a ratio scale is proposed. By relying on expert qualitative comparisons, weights can be

determined from a limited number of comparisons. The method requires that the user (the customer) answers a series of qualitative questions such as: “How important is the fulfilment of task Y compared to that of task Z, in order to perform their parent task X?”. The answers are to be chosen from verbal statements defined as the Fundamental Scale (Table 1). By applying an algorithm to the quantified intensities corresponding the statements, quantitative weights including w_i are given systematically.

Table 1 Fundamental Scale (T. L. Saaty & Vargas, 2012)

Intensity of importance	Definition	Explanation
1	Equal importance	Two activities contribute equally to the objective
2	Weak	
3	Moderate Importance	Experience and judgment slightly favor one activity over another
4	Moderate plus	
5	Strong importance	Experience and judgment strongly favor one activity over another
6	Strong plus	
7	Very strong or demonstrated importance	An activity is favoured very strongly over another; its dominance demonstrated in practice
8	Very, very strong	
9	Extreme importance	The evidence favoring one activity over another is of the highest possible order of affirmation
Reciprocals of above	If activity i has one of the above nonzero numbers assigned to it when compared with activity j, then j has the reciprocal value when compared with i	A reasonable assumption
Rationals	Ratios arising from the scale	If consistency were to be forced by obtaining n numerical values to span the matrix

Comparisons are carried out between each pair of tasks which share the same “parent” task in the value tree. High-level tasks are compared as well (Figure 5), and their results affect the importance weights of lower-level capabilities. This enables modelling changing customer capability needs at various levels, affecting the eventual value of Sols.

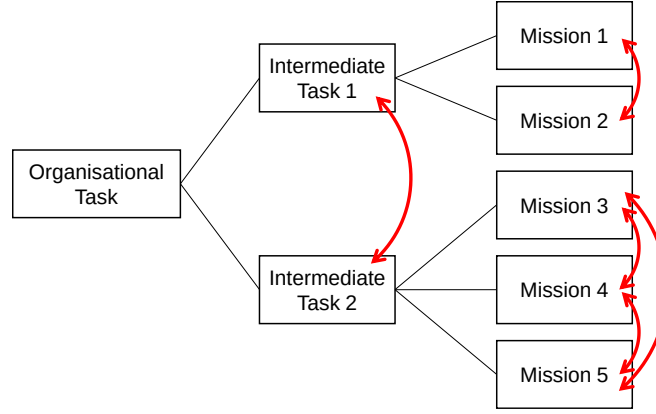


Figure 5 Pairwise Comparisons of Tasks

For scenario planning, one can create different scenarios for what-if analysis by making changes in some of the comparison results.

5. Mission Effectiveness Prediction

In this framework, missions are expressed as input-output models:

$$v_i = f_{mission\ i}(\mathbf{D}, \mathbf{C}) \quad (6)$$

Here the input is a design vector $\mathbf{D}$ and a context vector $\mathbf{C}$, while the output is mission capability level v_i . Assuming that v_i is given by applying aggregation function to a vector of mission **measures of effectiveness** (MoEs) $\mathbf{E}$, the function $f_{mission\ i}$ can be decomposed such that:

$$\mathbf{E} = f_{effectiveness\ i}(\mathbf{D}, \mathbf{C}) \quad (7)$$

$$v_i = f_{success\ i}(\mathbf{E}) \quad (8)$$

$$v_i = f_{success\ i}(f_{effectiveness\ i}(\mathbf{D}, \mathbf{C})) \quad (9)$$

Here, $f_{effectiveness\ i}$ represents an objective outcome of conducting the mission, while $f_{success\ i}$ would be defined by the customer organisation, as how they evaluate the mission successfulness. For example, when there are two MoEs which are enemy loss and friendly loss, one armed forces may prioritise minimising the latter to maximising the former, while another may do the opposite. While appropriately defining $f_{success\ i}$ reflecting the customer preference could be a problem worth pursuing, this will not be handled in this framework, and it is assumed that $f_{success\ i}$ is pre-defined.

In contrast, the underlying relationships of $f_{effectiveness\ i}$ are what needs to be investigated. For exploring design alternatives, it is desirable to be able to obtain the output from the input and vice versa reasonably and quickly. For this purpose, a method based on the Structural Equation Modelling (SEM) was developed. SEM fits empirical data onto a hypothetical causal path diagram specified by the analyst to identify a quantitative model (Figure 6). Because of the ability to allow

specifying causal relationships, it can yield parsimonious and interpretable model relating design, context, performance, and effectiveness in the mission. The developed method was applied to experimental data obtained from combat simulation software. The resulting model implies trade-off relationships between design and context factors, providing an indication on the possible means of realising desirable mission effectiveness in different contexts.

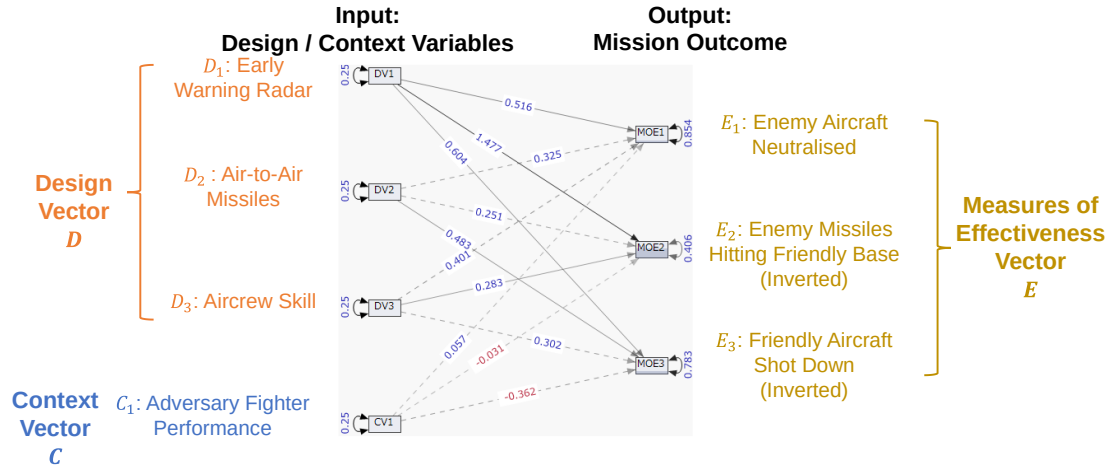


Figure 6 Fitted SEM Model

6. Implementation & Verification

An interactive computer programme for scenario planning was developed for demonstration and verification of the overall framework. Scenarios, which are characterised by external factors at each epoch, can be specified from the graphical user interface. The capability needs prioritisation method (Section 4) was integrated into it, allowing the user to prioritise capabilities by choosing pairwise comparison results. Mission models created by the proposed prediction method (Section 5) can be embedded in the program and can be used for estimating the realised capability level in each mission. By inputting them into the needs models identified for each epoch, the eventual value trajectory is calculated. The results are visualised for each design in the form of parallel coordinate plot. Any modifications on scenarios immediately invoke recalculation, and the visualisation is updated instantly. With this feature, it is expected that users can conduct the what-if analysis in an efficient manner.

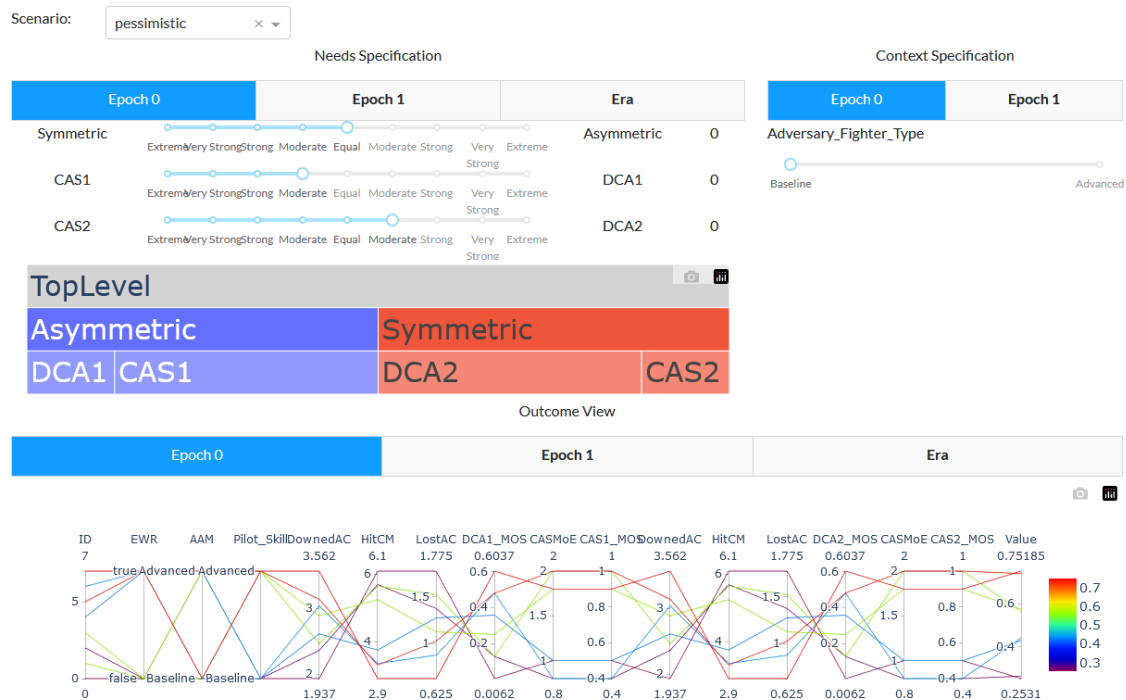


Figure 7 Software Implementation of Proposed Framework

Appendix: Terminology

Task: A clearly defined activity specifically assigned to an organisation. (adopted from (United States Department of Defense, 2013), as cited in (OUSD(R&E) USDOD, 2020))

Solution: An individual or organisation which performs a task.

Capability: The ability to complete a task under specified conditions and level of performance. (adopted from (OUSD(R&E) USDOD, 2020))

Capability Level: A quantity indicating the level of accomplishment of a task, as perceived by the customer.

Organisational Tasks: A set of tasks to which the customer organisation is assigned.

Organisational Capability: The ability of the customer organisation to complete organisational tasks under specified conditions and level of performance.

Value: A 1-D cardinal metric assigned to a solution indicating the preference of the customer.

Systems of Interest (Sols): One or more constituting entities of solution for one or more tasks whose alternatives are being analysed.

Design Variable: A variable representing a decision which differentiate one design of Sols from another.

Design Vector: A vector consisting of a set of design variables, which represents an instance of Sols design.

Sols Context: Factors external to Sols which are not related to the value perception of the customer.

Context Variable: A variable representing a factor which differentiates one Sols context from another.

Context Vector: A vector consisting of a set of context variables, which represents an instance of Sols context.

Epoch: A discrete time period with fixed context and needs.

Scenario: A sequence of epochs which describes one possible future awaiting the Sols.

Measure of Effectiveness: A metric indicating one aspect of the outcome of the mission execution, defined independent of solution.

Use Case

A use case for the framework was formulated for a demonstration. It is intended to be simple, yet relevant to all the issues which this framework is intended to cover, including: (1) planning of changing high-level needs, (2) changing mission needs, and (3) changing context in the form of scenarios, and (4) value-based evaluation of Sols against the planned scenarios. In this section, the problem setting and the process for formalising it into scenario planning are described stepwise.

Step 1

Nation A is planning the future capability development for their air force in a 10-year time frame. They have identified candidates of future actions as follows:

- Acquire and field early warning radar.
- Upgrade air-to-air missiles.
- Train air crew.

Their alternatives range from doing nothing to executing all three, making eight alternatives in total. To identify the most promising option, they decided to carry out an analysis using value-based scenario planning.

Step 2

First, their current needs were clarified. Their high-level tasks are:

- Conflict with nation B, which currently has limited air power.
- Conflict with nation C, which currently has advanced air power.

To accomplish these high-level tasks, the air force is expected to execute the following missions:

- Defensive counter air (DCA, protection from air attack)
- Close air support (CAS, ground attack)

They assessed current capability needs by giving answers to following questions:

- How important is the conflict with nation B, compared to the conflict with nation C?
- How important is DCA, compared to CAS, for the conflict with nation B?
- How important is DCA, compared to CAS, for the conflict with nation C?

The following are the results:

- Conflict with nation B is moderately more important, compared to that of nation C.
- For conflict with nation B, CAS is moderately more important than DCA.
- For conflict with nation C, DCA and CAS are equally important.

Based on this assessment, current capability needs were quantitatively expressed as a needs hierarchy.

Step 3

Concurrently, the input-output structures of the missions were determined. The input factors of the missions, which are design and context variables, were chosen to enable analysing their impact on the outputs, which are MoEs. In this use case, the three possible design changes of Sols identified in the step 1 were included in the design variables. Also, as the analysts anticipate that adversary fighter aircraft may be upgraded, its performance was included as a context variable.

However, some factors have only negligible influence on some missions. From analyses (e.g., modelling and simulation), influential factors on each mission effectiveness were identified as follows:

Design / Context		DCA	CAS
Design	Presence of early warning radar	✓	
Design	Performance of air-to-air missiles	✓	
Design	Skill level of air crew	✓	✓
Context	Performance of the adversary fighter aircraft	✓	✓

Based on this result, mission models were created using the proposed method.

Step 4

Anticipating the possible changes, two scenarios were formulated based on their future forecast on capability needs and context as follows:

“Optimistic” scenario:

- Conflict with nation B will be moderately more important compared to that of nation C. (Unchanged)
- For conflict with nation B, CAS will be strongly important than DCA.
- For conflict with nation C, DCA will be moderately more important than CAS.
- Nation C will improve the performance of their fighter aircraft.

“Pessimistic” scenario:

- Conflict with nation C will be strongly more important compared to that of nation B.
- For conflict with nation B, DCA and CAS will be equally important.
- For conflict with nation C, DCA will be strongly more important than CAS.
- Nation B and C will improve the performance of their fighter aircraft.

These scenarios were expressed in a form of sequences of parameterised needs and context.

Step 5

Using the proposed framework, value trajectories are calculated for each evolution alternative against each scenario.

REFERENCES

- Abbas, A. E., & Cadenbach, A. H. (2018). On the Use of Utility Theory in Engineering Design. *IEEE Systems Journal*, 12(2), 1129–1138. <https://doi.org/10.1109/JSYST.2016.2602562>
- Collopy, P. D., & Hollingsworth, P. M. (2012). Value-Driven Design. *Journal of Aircraft*, 48(3), 749–759. <https://doi.org/10.2514/1.C000311>
- Henshaw, M. J. de C., Lister, P., Harding, A. D., Kemp, D., Daw, A. J., Farncombe, A., & Touchin, M. (2011). 6.3.2 Capability Engineering – An Analysis of Perspectives. *INCOSE International Symposium*, 21(1), 712–727. <https://doi.org/10.1002/J.2334-5837.2011.TB01238.X>
- JCS USDoD. (n.d.). *Universal Joint Task List*. Retrieved 2 June 2023, from <https://www.jcs.mil/Doctrine/Joint-Training/UJTL/>
- OUSD(R&E) USDoD. (2020). *Mission Engineering Guide*. https://ac.cto.mil/wp-content/uploads/2020/12/MEG-v40_20201130_shm.pdf
- Ross, A. M., Gregory O'Neill, M., Hastings, D. E., & Rhodes, D. H. (2010). Aligning perspectives and methods for value-driven design. *AIAA SPACE Conference and Exposition 2010*. <https://doi.org/10.2514/6.2010-8797>
- Ross, A. M., & Rhodes, D. H. (2008). Using natural value-centric time scales for conceptualizing system timelines through epoch-era analysis. *18th Annual International Symposium of the International Council on Systems Engineering, INCOSE 2008*, 4(1), 2250–2264. <https://doi.org/10.1002/j.2334-5837.2008.tb00871.x>
- Saaty, T. L., & Vargas, L. G. (2012). *Models, Methods, Concepts & Applications of the Analytic Hierarchy Process* (Vol. 175). Springer US. <https://doi.org/10.1007/978-1-4614-3597-6>
- Taliaferro, A. C., Gonzalez, L. M., Tillman, M., Ghosh, P., Clarke, P., & Hinkle, W. (2019). *A Guide to Capability-Based Planning (CBP)* (Defense Governance and Management: Improving the Defense Management Capabilities of Foreign Defense Institutions). https://www.ida.org/research-and-publications/publications/all/d/de/defense-governance-and-management_improving-the-defense-management-capabilities-of-foreign
- United States Department of Defense. (2013). *Joint Publication 1: Doctrine for the Armed Forces of the United States, March 25, 2013, Incorporating Change 1, July 12, 2017*.

Appendix B Evaluation Questionnaire



Value-based planning for capability evolution

Research evaluation session

Researcher: Takashi Shibui (PhD Student)

Supervisors: Prof. Marin Guenov and Dr. Atif Riaz

Department: Centre for Aeronautics, School of Aerospace, Transport and Manufacturing

Date: 15/06/2023

Introduction

The purpose of this evaluation session is to help the researcher obtain feedback on the usefulness and industrial relevance of their work. For a brief overview of the research, please refer to the abstract and outline provided in the following sections.

Additionally, a demonstration of the framework developed will be presented through an application to a use case. More details about this use case are available in its corresponding section of this information pack. During the evaluation section, the application of the novel methods to the use case will be presented to the participants.

Objectives

The objectives of the session are:

- To introduce my research to a group of practitioners.
- To obtain feedback from this group regarding:
 - The validity of the assumptions.
 - The validity of the methods.
 - The usefulness and industrial relevance of the approach.
 - Suggestions for further improvement.

Format

The session is intended to be completed within two hours. The outline is envisioned to be as follows:

- Introductory presentation and use case demonstration (30 min)
- Questions and discussion (30 min)
- Questionnaire (15 min)
- Open questions / further discussion (45 min)

Required from participants

Before the session

Kindly read and complete the Informed consent form and the Contact details forms, which are included in this document.

After the presentation/demonstration

Please ask questions if something was not clear or if you require more information. Discussion is most welcome.

When completing the questionnaire

Please answer the questions to the best of your knowledge. If a question is not clear, please feel free to ask for further explanation.

After the session

Please do contact me if you would like to provide or receive more information. You can also contact me to exercise your right to withdraw as stated in the Informed consent form.

Questionnaire form

Part 1 – Likert type questions

Please indicate to what extent you agree or disagree with the following statements by marking the appropriate box.

	Strongly Disagree	Disagree	Neither agree or disagree	Agree	Strongly agree
I. Validity of assumptions					
1. It is reasonable to assume that a finite number of representative deterministic scenarios collectively capture the range of future uncertainties (as an alternative to a probabilistic approaches).					
2. It is reasonable to identify customer needs with the capability to perform their organisational tasks.					
3. It is reasonable to assume that a set of representative subtasks can be appropriately derived from a supertask so that the subtasks' fulfilment collectively indicates the fulfilment of the supertask.					
4. It is reasonable to evaluate the fulfilment of a supertask based on the fulfilment of its subtasks and their importance over each other.					
5. It is reasonable for the customer to compare solution alternatives based on their contribution to the fulfilment of the organisational tasks.					
6. It is reasonable to measure a complex mission achievement by aggregating measures of effectiveness.					
7. It is reasonable to estimate measures of effectiveness achieved by a given solution from empirical data (e.g., simulation results).					

II. Usefulness of individual methods					
8. A hierarchy of tasks (i.e., from strategic to tactical level) adequately describes customer needs and their changes.					
9. Pairwise semi-quantitative comparisons for task prioritisation help to translate intangible customer needs into formal value functions.					
10. Estimated causal models of missions help to understand underlying relationships between solutions and their mission effectiveness.					
III. Usefulness of overall framework					
11. The proposed framework enables effective identification of solution alternatives with a high value under a given future scenario.					
12. The proposed framework is capable of adding value to the company by providing useful knowledge for product development.					

Part 2 – Open questions

1. What possible difficulties and limitations do you anticipate in applying the proposed framework to real-world problems? (e.g. If you strongly disagree with one of the assumptions).

2. Regarding the proposed capability needs prioritisation method, what are the advantages and disadvantages in relation to the current industrial practice?

3. Regarding the proposed mission effectiveness prediction method, what are the advantages and disadvantages in relation to the current industrial practice?

4. Is there anything you found particularly useful within the proposed framework?

5. Is there anything you found particularly not useful within the proposed framework?