Managing Epistemic Uncertainties in the Underlying Models of Safety Assessment for Safety-Critical Systems

Chris Wai Kiat Leong

Doctor of Philosophy

University of York
Computer Science

Jun 2018

Abstract

When conducting safety assessment for safety-critical systems, epistemic uncertainty is an ever-present challenge when reasoning about the safety concerns and causal relationships related to hazards. Uncertainty around this causation thus needs to be managed well. Unfortunately, existing safety assessment tends to ignore unknown uncertainties, and stakeholders rarely track known uncertainties well through the system lifecycle.

In this thesis, an approach is described for managing epistemic uncertainties about the system and safety causal models that are applied in a safety assessment. First, the principles that define the requirements for the approach are introduced. Next, these principles are used to construct three distinct steps that constitute an approach to manage such uncertainties. These three steps involve identifying, documenting and tracking the uncertainties throughout the system lifecycle so as to enable intervention to address the uncertainties.

The approach is evaluated by integrating it with two existing safety assessment techniques, one using models from a system viewpoint and the other with models from a component viewpoint. This approach is also evaluated through peer reviews, semi-structured interviews with practitioners, and by review against requirements derived from the principles. Based on the evaluation results, it is plausible that our approach can provide a feasible and systematic way to manage epistemic uncertainties in safety assessment for safety-critical systems.

Contents

ABSTRACT ... 3

CONTENTS ... 5

LIST OF TABLES ... 9

LIST OF FIGURES ... 11

ACKNOWLEDGEMENTS ... 13

AUTHOR'S DECLARATION ... 14

CHAPTER 1 – INTRODUCTION ... 15

1.1 The Danger of Uncertainties in Safety Assessment ... 15
1.1.1 CHALLENGER Space Shuttle Explosion ... 15
1.1.2 PIPER ALPHA Disaster ... 16
1.1.3 Safety Assessment in the RSAF ... 17
1.1.4 Implication of These Uncertainties ... 18

1.2 Risk, Confidence and Uncertainty in Safety Assessment ... 18
1.2.1 Risk ... 19
1.2.2 Confidence ... 19
1.2.3 Uncertainty ... 20

1.3 Research Strategy ... 25

1.4 Thesis Proposition ... 29

1.5 Thesis Structure ... 29

CHAPTER 2 – SURVEY OF RELATED LITERATURE ... 33

2.1. Concepts in Uncertainty Management ... 34
2.1.1. Classifications of Uncertainty ... 35
2.1.2. Qualitative Modelling of Uncertainty ... 38
2.1.3. Quantitative Modelling of Uncertainty ... 45

2.2. Concepts in Safety Management ... 48
2.2.1. Safety Assessment ... 48
2.2.2. Safety Assurance ... 56
2.2.3. Treatment of Uncertainty in Safety Standards ... 58

2.3. Modelling of Safety-Critical Systems (SCS) ... 61
2.3.1. Growing Challenges of Uncertainty in SCS ... 62
2.3.2. Types of SCS Model ... 66
2.3.3. Analysis of SCS Models ... 71
2.3.4. Through-life Management of SCS ... 75

2.4. Summary ... 80

CHAPTER 3 – PRINCIPLES AND MODELS TO MANAGE EPISTEMIC UNCERTAINTIES 81

3.1. Introduction 81

3.2. Principles to Manage Epistemic Uncertainties 82
3.2.1. Principle 1: Epistemic uncertainties should be identified 83
3.2.2. Principle 2: Epistemic uncertainties should be documented. 88
3.2.3. Principle 3: Epistemic uncertainties should be tracked and addressed in a systematic manner through-life. 95

3.3. Underlying Models to Manage Epistemic Uncertainties 99
3.3.1. Conceptual Model of System Safety 100
3.3.2. Causal Relationship Model of Safety Assessment 110

3.4. Summary 115

CHAPTER 4 - AN APPROACH TO IDENTIFY, DOCUMENT, TRACK AND ADDRESS EPISTEMIC UNCERTAINTIES 117

4.1. Introduction 117
4.1.1. Overview of the Target-Analysis-Goal (T.A.G.) Approach 119
4.1.2. When Is T.A.G. Approach Applicable in Safety Assessment? 120

4.2. Initiative 1: A Taxonomy of Causal Mechanisms to Identify Uncertainty Target 122
4.2.1. Primary Causal Condition: Human 125
4.2.2. Primary Causal Condition: Organisation 126
4.2.3. Primary Causal Condition: Technology 127
4.2.4. Primary Causal Condition: Process 128
4.2.5. Primary Causal Condition: Information 129
4.2.6. Primary Causal Condition: Environment 130
4.2.7. Apply the Taxonomy to Recognise Uncertainty 131

4.3. Initiative 2: A Systematic Prioritisation Process to Document Uncertainty Analyses ..134
4.3.1. Process to Integrate Uncertainty Management in Existing Techniques 135
4.3.2. Factors to Prioritise Uncertainty Analyses 141

4.4. Initiative 3: A Structured Method to Track Action Plan and Address Goal 152
4.4.1. The GQTA Approach 153
4.4.2. Factors to Consider in Action Plan 154
4.4.3. Possible Outcomes When Addressing Uncertainty 158

4.5. Initial Evaluation and Summary 161

CHAPTER 5 – APPLICATION OF T.A.G. APPROACH IN SAFETY ASSESSMENTS FROM SYSTEM VIEWPOINTS 163

5.1. Introduction 163

5.2. Safety Assessment using STPA 165
5.2.1. Yongwen Railway Safety Control Structure 166
5.2.2. Unsafe Control Actions between TCC and ATP 169
5.2.3. Control Flaws leading to the Unsafe Control Actions 169

5.3. Integrate T.A.G. approach into STPA 169
5.3.1. Process of Tagging Uncertainties in STPA 170
5.3.2. Step 1a: Identify and Tag Uncertainties from SCS 170
5.3.3. Step 1b: Identify and Tag Uncertainties from UCAs and CFs 174
5.3.4. Step 2: Document and Prioritise Uncertainty Analyses 178

5.3.5. Step 3: Generate Track and Address Action Plan 182

5.4. Initial Evaluation and Summary 187

CHAPTER 6 – APPLICATION OF T.A.G. APPROACH IN SAFETY ASSESSMENTS FROM COMPONENT VIEWPOINTS 189

6.1. Introduction.. 189

6.2. Safety Assessment using FTA 191

6.3. Integrate T.A.G. approach into FTA 194

6.3.1. Process of Tagging Uncertainties in FTA 194

6.3.2. Step 1: Identify and Tag Uncertainties from Fault Tree.. 195

6.3.3. Step 2: Document and Prioritise Uncertainty Analyses 197

6.3.4. Step 3: Generate Track and Address Action Plan 198

6.4. Safety Assessment using FMEA.. 200

6.5. Integrate T.A.G. approach into FMEA... 203

6.5.1. Process of Tagging Uncertainties in FMEA.. 203

6.5.2. Step 1: Identify and Tag Uncertainties from FMEA 204

6.5.3. Step 2: Document and Prioritise Uncertainty Analyses 207

6.5.4. Step 3: Generate Track and Address Action Plan 208

6.6. Initial Evaluation and Summary 210

CHAPTER 7 - EVALUATION 213

7.1. Introduction.. 213

7.1.1. Structured Argument 213

7.1.2. Goal 1 (G1): Thesis Proposition.. 213

7.1.3. Strategy 1 (S1): Verification 215

7.1.4. Strategy 2 (S2): Validation... 216

7.2. Evaluation against Initial Requirements 217

7.2.1. Verify Contributions to Identify Epistemic Uncertainty 217

7.2.2. Verify Contributions to Document Epistemic Uncertainty 219

7.2.3. Verify Contributions to Track and Address Epistemic Uncertainty............................... 220

7.3. Peer Review – Sharing at Conferences and Workshops................................. 221

7.4. Peer Review – Semi-Structured Interview.. 223

7.4.1. Interview Methodology.. 223

7.4.2. Interview Analysis 228

7.5. Preliminary Industrial Adaptation.. 238

7.5.1. Applying T.A.G. approach in Multi-Perspective Hazard Analysis 239

7.5.2. Context-Specific HOT-PIE Taxonomy 240

7.5.3. Review of Prioritisation Factors... 241

7.5.4. Action-Item Table for Track Action Plan 242

7.6. Conclusions.. 243

CHAPTER 8 – CONCLUSIONS 247

8.1. Summary of Thesis Contributions.. 247

8.1.1. Contributions to Identifying Epistemic Uncertainties 247

8.1.2. Contributions to Document Epistemic Uncertainties 248
8.1.3. Contributions to Track and Address Epistemic Uncertainties 248
8.1.4. Overall Conclusions 249

8.2. Future Work 250
8.2.1. Complementing Quantitative Techniques 250
8.2.2. Complementing with Database 251
8.2.3. Integrating with other Safety Assurance Initiatives 251
8.2.4. Integrating with other Safety Assessment Initiatives 251
8.2.5. Refining through Pilot Implementation 252
8.2.6. Applying beyond Safety 252

8.3. CODA 253

ANNEX A – LITERATURE REFERENCED IN HOT-PIE TAXONOMY 255

ANNEX B – GOAL SETTING GUIDE 263

ANNEX C – SEQUENCE OF EVENTS FOR YONGWEN RAILWAY ACCIDENT 265

ANNEX D – CHAPTER 5 STPA-TAG: PRIORITISATION OF OBSERVATIONS 2, 3 & 4..267

ANNEX E – CHAPTER 6 FTA-TAG: PRIORITISATION OF OBSERVATIONS 2 275

ANNEX F – WHEEL BRAKE SYSTEM DESCRIPTION 277

ANNEX G – CHAPTER 6 FTA-TAG: ACTION PLAN OF OBSERVATION 2 279

ANNEX H – CHAPTER 6 FMES-TAG: PRIORITISATION OF OBSERVATIONS 3 283

ANNEX I – CHAPTER 6 FMES-TAG: ACTION PLAN OF OBSERVATION 3 285

ANNEX J – EVALUATION QUESTIONNAIRE 289

ANNEX K – T.A.G. APPROACH USERS GUIDE 293

GLOSSARY 305

REFERENCES 307

List of Tables

Table 1. Different Perspectives Regarding Risk and Uncertainty 23
Table 2. The Four Phases under the Research Methodology 26
Table 3. Research Methods to Tackle 'What', 'Where' and 'How' 27
Table 4. Comparison between Data, Information and Knowledge [26] 34
Table 5. Decision Making in Different Context under Cynefin Framework 40
Table 6. Uncertainty Factor (extracted from Aven) 47
Table 7. Terminology of Safety 49
Table 8. Cycle of Key Safety Assessment Processes (extracted from DEF-STD 00-56)........ 50
Table 9. Comparing Traditional and Emergence Engineering (extracted from Ulieru) 64
Table 10. Types of Emergent Hazard (extracted from Redmond) 64
Table 11. Types of Model to Represent Safety-Critical Systems 66
Table 12. Comparison of Development and Operation Phases 75
Table 13. Association between Safety Elements and System Elements of the same Type 105
Table 14. The T.A.G Approach to Proactively Manage Epistemic Uncertainty 120
Table 15. Descriptions of the Six Primary Causal Conditions in HOT-PIE Taxonomy 124
Table 16. HOT-PIE Primary and Secondary Causal Conditions 124
Table 17. Causal Mechanisms under the Human Primary Causal Conditions 125
Table 18. Causal Mechanisms under the Organisation Primary Causal Conditions 126
Table 19. Causal Mechanisms under the Technology Primary Causal Conditions 127
Table 20. Causal Mechanisms under the Process Primary Causal Conditions 128
Table 21. Causal Mechanisms under the Information Primary Causal Conditions 129
Table 22. Causal Mechanisms under the Environment Primary Causal Conditions 130
Table 23. Literature Survey on the Characteristics of Uncertainties 142
Table 24. Example of Expected Efforts Calculation to Analyse Uncertainties 147
Table 25. Guided Questions to Prioritise Uncertainty Analyses 151
Table 26. Literature Survey of Ways to Track and Address Uncertainties 154
Table 27. Guide to Set Goals for Action Plan 157
Table 28. Summary of Chapters that Apply T.A.G. Approach 163
Table 29. Safety Constraints for Railways System Components 167
Table 30. Unsafe Control Actions for the TCC and ATP 169
Table 31. Control Flaws for the TCC and ATP........ 169
Table 32. Summary of Epistemic Uncertainties Identified using HOT-PIE Taxonomy 177
Table 33. Responses to Guided Questions for Uncertainty Observation 1 178
Table 34. Scores based on Prioritisation Factors........ 180
Table 35. Summary of Prioritisation for the Uncertainty Observations 181
Table 36. Guide to Develop Action Plan for Uncertainty 3 182
Table 37. Summary of Chapters that Apply T.A.G. Approach 189
Table 38. Summary of Epistemic Uncertainties Identified in WBS FTA 197
Table 39. Summary of Prioritisation for the Uncertainty Observations from WBS FTA 197
Table 40. Summary of Epistemic Uncertainties Identified in WBS FMES 207
Table 41. Summary of Prioritisation for the Uncertainty Observations from WBS FMES 207
Table 42. GSN Elements used in the Evaluation 213
Table 43. 2nd level Goals under Verification........ 215
Table 44. Types of Evaluating Approach........ 217
Table 45. Summary of Participants 225
Table 46. Thematic Analysis Approach 227
Table 47. Summary of Adaptation of the T.A.G. Approach 239
Table 48. Multi-Perspectives Hazard Analysis 239
Table 49. Context-Specific HOT-PIE under Scenario Based HA for Engineering Testing 240
Table 50. Example of Track and Address Table 243
Table 51. Responses to Guided Questions for Uncertainty Observation 2 267
Table 52. Responses to Guided Questions for Uncertainty Observation 3 269
Table 53. Responses to Guided Questions for Uncertainty Observation 4 272
Table 54. Answers to Guided Questions for Uncertainty Observation 2 275
Table 55. Guide to Develop Action Plan for Uncertainty 2 279
Table 56. Answers to Guided Questions for Uncertainty Observation 3 283
Table 57. Guide to Develop Action Plan for Uncertainty 3 285

Table 58. The HOT-PIE Taxonomy of Causal Mechanisms 294
Table 59. Guided Questions to Prioritise Uncertainty Analyses 297
Table 60. Prioritisation Table for Uncertainty Action...298
Table 61. Guide to Set Goals in Action Plan 300

List of Figures

Figure 1. Association between ram-Uncertainty and Safety Risk 21
Figure 2. Scenarios to Consider both Risk and Uncertainty in Safety Assessment 24
Figure 3. Complementary Divergent and Convergent Loops in Our Research 26
Figure 4. Underlying Models of Safety Assessment* 28
Figure 5. Three Focus Areas in Literature Survey 33
Figure 6. Johari Window 37
Figure 7. Example of a Fault Tree (extracted from ARP 4761) 53
Figure 8. FMEA Methodology (extracted from Ericson [65]) 54
Figure 9. Example of FMEA Worksheet (extracted from ARP 4761) 54
Figure 10. Typical Control Flaws for a Controlled Loop (extracted from STPA Primer) 55
Figure 11. Example of CPS in Aviation (extracted from Sampigethaya) 68
Figure 12. Ministry of Defence Architectural Framework (MoDAF) (extracted from Benade) 71
Figure 13. Typical Process Flow for CBR 72
Figure 14. A Product-Line Engineering Approach [79] 73
Figure 15. Concept of Modular Certification (extracted from Trapp) 74
Figure 16. Categorisation of System Acquisition Lifecycle in Defence 75
Figure 17. Example of Integrating Safety and System Lifecycles 77
Figure 18. Key Relationships between Elements: Association, Inheritance and Aggregation 82
Figure 19. Relations between Uncertainty and Assumption 89
Figure 20. MoD CADMID Cycle 92
Figure 21. MOD’s CADMID vs Design Lifecycle 93
Figure 22. The RSAF Island Air Defense System [135] 96
Figure 23. The two Underlying Models of Safety Assessment 100
Figure 24. Conceptual Model of System Description according to IEEE 42010 101
Figure 25. Example of Conceptual Model of System Description 103
Figure 26. Interactions between System and Safety Elements 104
Figure 27. Example of Interactions between System and Safety Elements 106
Figure 28. Conceptual Model of System Safety 108
Figure 29. Example of CMSS applied on Weapon System Safety Assessment 110
Figure 30. The 'Condition' Concept adopted from Wilson's Safety Data Model 111
Figure 31. Causal Relationship in Safety Causal Model 112
Figure 32. Coleman’s Boat of Causal Pathways 113
Figure 33. Example of a Coleman’s Boat of Causal Pathway 114
Figure 34. Causal Relationship Model of Safety Assessment 114
Figure 35. Illustration of CRMSA 114
Figure 36. Integrating T.A.G. Approach into Existing Safety Assessment 121
Figure 37. T.A.G. Approach First Step: Identify Target 122
Figure 38. Illustration of a Human Related Uncertainty represented with CRMSA 126
Figure 39. Illustration of an Organisation Related Uncertainty represented with CRMSA 127
Figure 40. Illustration of a Technology Related Uncertainty represented with CRMSA 128
Figure 41. Illustration of a Process Related Uncertainty represented with CRMSA 129
Figure 42. Illustration of an Information Related Uncertainty represented with CRMSA 130
Figure 43. Illustration of an Environment Related Uncertainty represented with CRMSA 131
Figure 44. Causal Relationship Model of Safety Assessment 133
Figure 45. T.A.G. Approach Second Step: Document Analysis 134
Figure 46. MoD CADMID Cycle 135
Figure 47. Generic Process Flow for Through-Life Safety Assessment 136
Figure 48. Process Flow for Zonal Safety Analysis 137
Figure 49. Process Flow to Integrate T.A.G. Approach in Safety Assessment 139
Figure 50. Process Flow to Integrate T.A.G. Approach in Zonal Safety Analysis 140
Figure 51. Factors to Prioritise Uncertainty Analyses 143
Figure 52. Prioritisation Matrix for Uncertainty Analysis 147
Figure 53. Prioritisation Matrix Example 149
Figure 54. An Example of Prioritisation for Uncertainty Analysis 150
Figure 55. T.A.G. Approach Third Step: Track and Address Goal 152
Figure 56. T.A.G. Approach Step 3: Track and Address Goal 152
Figure 57. Focus Areas in Uncertainty Management Action Plan 155

Figure 58. Possible Outcomes when Addressing the Uncertainty..160
Figure 59. Process Flow for STPA Safety Assessment..166
Figure 60. Basic Railway Safety Control Structure in the Yongwen Accident..................................168
Figure 61. Process Flow to Tag Epistemic Uncertainties in STPA Safety Assessment......................172
Figure 62. Uncertainty about Fatigue among Workers 173
Figure 63. Uncertainty about Infrastructure Support by Shanghai Railway Bureau 174
Figure 64. Uncertainty about Software Algorithm during Signal Abnormality 175
Figure 65. Uncertainty about Human Intervention during Signal Abnormality..................................176
Figure 66. Uncertainty Observation 1 Plotted on Prioritisation Matrix Chart 180
Figure 67. Uncertainty Observation 1 to 4 Plotted on Prioritisation Matrix Chart 181
Figure 68. Possible Scenarios when Addressing Uncertainty in Software Algorithm 186
Figure 69. Safety Assessment Techniques Applied on FHA, PSSA and SSA....................................191
Figure 70. Process Flow for FTA Safety Assessment..192
Figure 71. WBS FTA: Unannunciated Loss of All Wheel Braking (extracted from ARP 4761) 193
Figure 72. Process Flow to Tag Epistemic Uncertainties in FTA Safety Assessment........................194
Figure 73. Uncertainty about Reaction Time due Limited Vision 195
Figure 74. Uncertainty about the Logic in the BSCU Command Messages 196
Figure 75. Uncertainty Observation 1 and 2 Plotted on Prioritisation Matrix Chart...........................198
Figure 76. Process Flow for FMEA Safety Assessment 201
Figure 77. Wheel Break System (WBS) Components Breakdown (extracted from ARP 4761) 201
Figure 78. Wheel Brake System FMES of Braking Failure Effect at SSA 202
Figure 79. Process Flow to Tag Epistemic Uncertainties in FMEA Safety Assessment.....................204
Figure 80. Uncertainty about Mechanical Faults at Control Valves 205
Figure 81. Uncertainty about Task Overload during Emergency..206
Figure 82. Uncertainty Observation 3 and 4 Plotted on Prioritisation Matrix Chart...........................208
Figure 83. Structured Argument to Evaluate Thesis Contributions using the GSN 214
Figure 84. Trending Based on Experience 231
Figure 85. Trending Based on Industry...232
Figure 86. Thematic Map from Semi-Structured Interview..234
Figure 87. Refined Prioritisation Factors for NwSoS 242
Figure 88 Uncertainty in the Infrastructure Support by Shanghai Railway Bureau............................267
Figure 89 Uncertainty Observation 2 Plotted on Prioritisation Matrix Chart 269
Figure 90 Uncertainty in the Role of Software during Signal Abnormality..269
Figure 91 Uncertainty Observation 3 Plotted on Prioritisation Matrix Chart 271
Figure 92 Uncertainty about Human Interrupt during Signal Abnormality 272
Figure 93 Uncertainty Observation 5 Plotted on Prioritisation Matrix Chart 274
Figure 94 Uncertainty Observations Plotted on Prioritisation Matrix Chart.......................................276
Figure 95. Wheel Break System (WBS) Components Breakdown...277
Figure 96 Uncertainty Observation 1 Plotted on Prioritisation Matrix Chart 284
Figure 97. Process Flow to Integrate T.A.G. Approach in Safety Assessment...................................293
Figure 98. T.A.G. Approach First Step: Identify Target...294
Figure 99. T.A.G. Approach Second Step: Document Analysis...296
Figure 100 An Example of Prioritisation for Uncertainty Analysis..299
Figure 101. T.A.G. Approach Third Step: Track and Address Goal...300
Figure 102. Possible Outcomes when Addressing the Uncertainty...301
Figure 103 Conceptual Model of System Safety (CMSS) 302
Figure 104 Causal Relationship Model of Safety Assessment (CRMSA) 303
Figure 105 Illustration of CRMSA 303

Acknowledgements

I would like to thank my supervisors Tim Kelly and Rob Alexander for your guidance, assistance and encouragement. You have provided me the focus and insights when I have difficulties in appreciating complicated subjects such as uncertainty and knowledge. I would also like to thank the staff members at the University of York who have made my life in the university a wonderful experience.

I would like to thank the Republic of Singapore Air Force for providing me this opportunity to embark on a research that is close to my heart and important to the organisation. The Air Force has provided me with life-long opportunities and exposures to upgrade myself professionally as a soldier and technically as an engineer.

Finally, I would like to thank my family and friends for their love and support. I would not have persevered if not for your encouragements throughout this journey.

Author’s Declaration

Some parts of the research presented in this thesis have been previously published in the following papers and presentations:

- Paper: Leong, Kelly, Alexander, “Incorporating Epistemic Uncertainty into the Safety Assurance of Socio-Technical Systems”, Causal Reasoning for Embedded and Safety-critical Systems Technologies (CREST) 2017, Apr 2017.
- Paper: Leong, Kelly, Alexander, “A Structured Approach to Manage Model Uncertainty in Safety Assessment”, International System Safety Conference, Aug 2017.
- Presentation: Leong, “Managing Epistemic Uncertainties in the Underlying Models of System Safety Assessment for Socio-Technical Systems”, International Conference on System Safety and Cyber Security, Oct 17.
- Presentation: Leong, “Managing Epistemic Uncertainties in the Underlying Models of System Safety Assessment for Safety-Critical Systems”, System Safety Society (Singapore Chapter), Jan 18.
- Paper: Leong, “Managing Epistemic Uncertainties in the Underlying Models of System Safety Assessment for Safety-Critical Systems”, Singapore Aerospace Technology and Engineering Conference 2018, Feb 18.

I declare that this thesis is a presentation of original work and I am the sole author. This work has not previously been presented for an award at this, or any other, University. All sources are acknowledged as References.

Chapter 1 – Introduction

1.1 The Danger of Uncertainties in Safety Assessment

Safety assessment throughout the lifecycle for large-scale safety-critical system can be complicated. It involves multiple areas of expertise, resources and time to make informed decisions about the safety of the system not only for one safety assessment but multiple assessments throughout the system lifecycle. The knowledge that is needed to make a confident analysis depends on the quality of information available during the safety assessment and often depends on analyses conducted in previous safety assessments. In practice, there is a risk that safety analysts may make unsafe decisions when there is uncertainty about the information used during the safety assessment. The worst scenario could be a misplaced sense of confidence that the information available adequately represents the real world where the system is operating, which may lead to an unsafe situation during system development and operation. Unfortunately, such a misplaced sense of confidence has resulted in many catastrophic accidents and incidents as shown in the following examples.

1.1.1 CHALLENGER Space Shuttle Explosion

On 28 January 1986, the space shuttle CHALLENGER exploded 73 seconds after launch and killed all seven crew members. The cause of the explosion was the technical failure of the primary and secondary rubber seals (known as O-rings) at the solid rocket boosters to prevent hot propellant gases from escaping the boosters to other parts of the shuttle, such as the external fuel tanks. The temperature was so cold that the O-ring was not able to perform its role of sealing the joints. This O-ring breakthrough caused the catastrophic explosion that lead to the disaster.

In her research on the CHALLENGER space shuttle accident [1], the sociologist Diane Vaughan surfaces multiple problems that could be contributing to the misfit, such as the overconfident belief in the assumed safety of the shuttle launch through the years (she called this ‘normalised of deviance”), cultural and social context during the multiple meetings and reviews prior to the launch that discourage candid interactions, as well as the lack of complete information to make a safe decision to avoid the accident. In our research, we are concerned with the last problem - the lack of sufficient information on the performance of the O-ring that affect the subsequent decision made on the launch of the space shuttle.

The hazard we are concerned with is the O-ring breakthrough and the uncertainty is with the analysis of the O-ring performance specifically at low temperature. In her research, Diane wrote that while safety analysts had plotted the performance of the O-ring at high temperature based on field data of previous launchers, they did not have data for low temperature as there were no launching done at low temperature before. Hence, there was some data but insufficient to provide conclusive evaluation if the O-ring can still perform safely at low temperature. There was uncertainty about the safety assessment by the analysts and this affected the assessment on the actual day of the launch. As stated by Diane, on the day of the launch, the surface temperature fell to 80-degree Fahrenheit before the launch, but this was not measured and reported since it was not a launch criterion.

The lack of uncertainty monitoring was also sharply pointed out by American physicist, Richard Feynman, who was an appointed member of the Rogers Commission to investigate the cause of the CHALLENGER disaster. During the commission hearing, Feynman [2] reported that "*the fact that this danger did not lead to a catastrophe before is no guarantee that it will not the next time, unless it is completely understood. When playing Russian roulette the fact that the first shot got off safely is little comfort for the next*". He was implying that uncertainty about the O-ring performance in low temperature was not handled properly.

For the CHALLENGER space shuttle accident, this uncertainty about the low temperature performance of the O-ring was not tracked and highlighted timely to the higher management. The potential danger of the O-ring not able to seal at lower temperature was not raised as a critical issue. Instead, it became what Diane stated as "routine, mixed and weak signal" that the management eventually conclude as not important enough to terminate the launch. The management had missed a good chance to arrest the danger of launching at low temperature so as to prevent the accident.

In conclusion, we believe that while uncertainty about safety assessment may not directly cause an accident, proactive monitoring of such uncertainty could raise awareness to other issues that may prevent a hazard from becoming an accident. Our research wants to focus on monitoring such uncertainty systematically so as to increase the confidence in safety assessment.

1.1.2 PIPER ALPHA Disaster

In the PIPER ALPHA disaster on the evening of 6 July 1988, 165 lives out of the 226 people were lost on board the offshore oil platform. According to the Lord Cullen's public inquiry [3], several factors have been identified that could have contributed to the disaster. Some of the major factors include:

- Job pressure to complete the task leading to deviation from procedure and poor shift handover,

- Lack of communication among the workers to avoid the faulty pump,
- Poor documentation of the job completed and outstanding,
- Legacy design on the oil platform that didn't contain the fire after multiple explosions, and
- Lack of emergency preparedness and training leading to wasted time during rescue.

While Lord Cullen had the benefit of hindsight when he wrote the report, it was also discovered that many of these factors were known issues prior to the accident. For example, the legacy design of the oil platform was highlighted during a safety audit and it was concluded that more in-depth analysis needs to be conducted to assess the fire-fighting system at the production platforms where the fire brought out. From cost and risk perspectives, the analysis was not given priority based on the information available then and it was not followed up after the audit. Unfortunately, the analysis wasn't conducted even until the disaster happened, and the failure of the fire-fighting system was one of the contributing factors for not localising and containing the fire within the production platforms.

If the uncertainty associated with the safety assessment about the fire-fighting system within the production platform was flagged and managed with the right emphasis, it could have been tracked and followed up diligently through the system lifecycle. While this might not have guaranteed that the necessary analysis would have been conducted in time, it could at least have ensured that such plausible but uncertain concerns were not ignored before it was given a chance to be investigated. Such uncertainty might turn out to be safety-critical when more information is solicited after an initial safety assessment.

1.1.3 Safety Assessment in the RSAF

In the Republic of Singapore Air Force (RSAF), we introduce new weapon systems regularly into our fighting forces, while continuing to maintain existing and legacy systems. Such new systems would go through a lifecycle similar to the MoD's CADMID cycle [4]. Safety assessments have to be conducted at various milestones and the system needs to be assessed to be safe before the acquisition process can proceed. Uncertainty about such safety assessment can occur when the relevant information is either not available or insufficient.

Scenarios where information is not available include:

- contractor does not have equipment specification for the equipment that is to be provided by sub-contractors,
- expertise not present during the safety assessment, and

- inability to predict the operation profiles owing to another system that is in the operating concept being yet to be developed.

Scenarios where information is not sufficient include:

- use of preliminary system design instead of the final one,
- use of initial operating concept before the final concept is available further down the lifecycle, and
- use of flight trial report that only consists of 3 flight profiles out of the desired 4, as the last flight profile was not conducted due to bad weather.

While thankfully, there has been no major accident in the RSAF that is due to uncertainty from safety assessment, this concern continues to be one of the top watch areas as the airforce moves toward operating in safety critical system-of-systems environment.

Moving forward, we will use running examples in this thesis that are based on uncertainties experienced while conducting safety assessments for new systems in the RSAF.

1.1.4 Implication of These Uncertainties

All the examples above represent complicated safety-critical systems with established risk assessment and management programs. However, the examples show that uncertainties continued to exist even in safety-critical systems with established safety assessment programs. When not managed well, these hazards that have uncertainties can become catastrophic and result in the loss of lives and property. In a safety assessment, there could be uncertainties that would need more time to investigate. If the uncertainties are identified early, it could have been tracked and follow-up diligently throughout the system lifecycle. The danger is when uncertainty that turns out to be safety-critical (when more information is solicited) gets discarded before it is investigated.

This thesis is concerned with the management of such uncertainties during safety assessment of complicated safety-critical systems. Since the danger of uncertainties in safety assessment is apparent, there is value to consider complementing existing safety assessment techniques with additional steps to better manage potentially safety-critical uncertainties.

1.2 Risk, Confidence and Uncertainty in Safety Assessment

Before discussing the management of uncertainty in safety assessment, it is important to differentiate the concept of risk and confidence in safety assessment. In safety assessment, risk and confidence can often be confused and mistaken to mean the same thing.

1.2.1 Risk

Traditionally, safety assessment has mostly been focused on risk, which is defined as a "*combination of the severity of the mishap and the probability that the mishap will occur*" [5]. It involves multiple activities such as hazard identification, risk analysis and risk management. Hazard identification involves discovering hazards that can potentially result in a mishap, risk analysis estimates the probability of occurrence and severity of a mishap and risk management proposes actions to reduce the risk associated with the hazards.

Stakeholders would use the arguments and evidence from the risk analysis and risk management to claim that a system is reasonably safe despite the presence of potential hazards. A safety argument explains how the risks due to each hazard are either not safety-critical or that that the risk can be mitigated to an extent that the system is reasonably safe. For example, the proposition that the system is safe for flight because there has been a successful flight trial is part of a safety argument. A safety evidence provides the facts or information to support the truth of a claim or argument about the safety of a system. For example, the flight trial result is a piece of evidence to support that argument that a system is safe by doing a flight trial.

Risk assessment is a fundamental activity in safety assessment for stakeholders to determine hazards and suitable measures to adequately address these hazards. There are many established principles and processes to conduct risk assessment in existing safety assessment techniques. For example, the "as low as reasonably possible" (ALARP) principle [6] is adopted in the U.K. to help stakeholders justify that risk has been reduced to a point whereby "*the cost involved in reducing the risk further would be grossly disproportionate to the benefit*". There are also many established techniques to identify hazards such as the Fault Tree Analysis (FTA) and the Failure Modes and Effects Analysis (FMEA) [7]. Stakeholders can also make use of a risk assessment matrix to analyse the probability and severity of a mishap. There are also many hazard logging tools and hazard tracking systems available for stakeholders to manage the hazards throughout the system lifecycle.

1.2.2 Confidence

To derive greater assurance in a safety assessment, it is important to focus on both risk and confidence during the analysis. This is echoed by Jelen [8], who states that there is a need to focus on getting sufficient confidence that one is satisfied with the safety assessment, besides focusing on only the assessment itself. If risk is used to quantify the extent that the system is safe during a safety assessment, then confidence measures the trust that one has on the safety assessment. For example, supporting a similar argument, the Federal Aviation Administration in the U.S.A. establishes a need for its Safety Management System to “systematically provide confidence that organizational outputs meet or exceed safety requirements” [9].

It is not easy to measure and analyse confidence as it is often implicit during safety assessment, rather than explicitly discussed and documented. Furthermore, there is often a lack of distinction between risk and confidence in a safety assessment. For example, a safety analyst may assess that a system is less risky because the safety assessment has been conducted by an experienced and reputable individual. However, being experienced and reputable should only increase the confidence in the evidence provided. It is the findings from the safety assessment that should determine the level of risk associated with a system.

Such lack of distinction between risk and confidence can make it difficult to spot incompleteness and imperfection when justifying the level of risk and the level of confidence in a safety assessment. Since risk is often the key focus, there is a danger that important considerations that affect confidence get ignored or discarded because of large volume of risk arguments and evidence presence in the assessment.

1.2.3 Uncertainty

In safety assessment, the greater the uncertainties, the less confidence we have with the results of the assessment. Uncertainty can broadly be classified as aleatory or epistemic [10]. Aleatory uncertainty is about randomness (e.g. the outcome of throwing a dice); epistemic uncertainty is due to a lack of knowledge (e.g. lack of details in the initial design document). In this research, the focus is on the management of epistemic uncertainty. Epistemic uncertainty can be due to things we know we do not know (known uncertainties), or things we do not know we do not know (unknown uncertainties) [11]. In practice, safety analysts never completely know if they have all the information about the system and its environment when modelling them during safety assessment. Hence, both known and unknown epistemic uncertainties are unavoidable.

> *“When an agency is evaluating significant adverse effects on the human environment in an environmental impact statement and there are gaps in relevant information or scientific uncertainty, the agency shall always make clear that such information is lacking or that uncertainty exists.”*

Extract from "Policy Issues Related to Worst Case Risk Analyses and the Establishment of Acceptable Standards of De Minimis Risk", Miller B. Spangler, U.S. Nuclear Regulatory Commission” -

The above extract from the US Nuclear Regulatory Commission illustrates the importance of acknowledging and managing uncertainty when evaluating adverse effects on the human environment. In a separate study on the role of interdisciplinary analysis on technical risk assessment, Spangler [12] similarly highlights the demand for knowledge - "*It is quite obvious that the assessment of such a wide diversity of beneficial and adverse consequences involves knowledge and scientific methods from a large number of scientific disciplines*" and the need to focus on managing uncertainty when our scientific knowledge is limited - "*The greater the limitations of science, for whatever reasons, to provide reliable answers, the greater is the need for wisdom to carry the burdens of analysis and translate the resulting balance of knowledge, judgment, and uncertainty into a meaningful form for decision-making*".

In safety assessment, such adverse effects are measured in terms of technical safety risk. Similarly, we are concerned with uncertainty associated with such technical risk assessment, which we name it as ra-uncertainty for short in this thesis. And the specific ra-uncertainty that we are focusing on is the uncertainty associated with models that are used to predict safety risk, which we name it as risk assessment model uncertainty, or ram-uncertainty for short.

For this section, we address the association between such ram-uncertainty and safety risk. First, we explain how ram-uncertainty is considered in relation with safety risk. Next, we contrast our understanding of ram-uncertainty and safety risk against another researcher that also uses the terms uncertainty and risk.

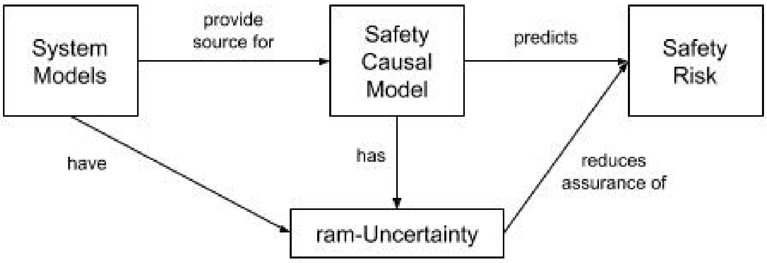


Figure 1. Association between ram-Uncertainty and Safety Risk

The relationships between ram-uncertainty and safety risk are summarised in Figure 1. From the diagram, we can make three major observations:

- The figure shows that in safety assessment we gather together relevant models of the system, from which we can derive a safety causal model (e.g. a model that establishes the causal relationships that exist between faults, failures, hazards and accidents). This causal model (which we will describe in detail under section 3.3.2) provides us with the basic information from which we attempt to predict the safety risk associated with those hazards (from an understanding of the causes and consequences of the hazards).

- Whilst this is a well-understood concept of safety assessment, in this thesis we explicitly recognise that it involves inherent uncertainties. Firstly, there can be uncertainties associated with the underlying system models (e.g. do they accurately and sufficiently capture all the relevant design dependencies and influences?). Secondly, both the uncertainty associated with these 'source' models, and the subjectivity present in the creation of the safety causal models (e.g. the subjectivity of scoping assumptions used in hazard identification) inevitably leads to uncertainty in the safety causal models that are established. This uncertainty is the ram-uncertainty we defined at the start of the section.
- As a result of such ram-uncertainty, safety analysts will be less confident of the safety risk that has been predicted from the safety causal model. In order words, the ram-uncertainty reduces the assurance in this area.

The relationships between system models and safety causal models, as well as with the ram-uncertainty, are described in detail under section 3.3.1.

For example, in the CHALLENGER space shuttle accident, the safety analysts needed to predict the safety risk of having a blow-through of the O-ring (i.e. the safety risk of causing damage or life lost if the blow-through happen). The associated question about the knowledge they have related to the blow-through (e.g. accuracy of the rocket booster model and the significance of temperature on the behaviour of the O-rings) would determine the ram-uncertainty associated with the safety risk. In this accident, those making the decision were clearly aware of the blow-through safety risk, but due to the lack of knowledge (i.e. uncertainty) underestimated this safety risk.

Of course, it is true that the safety causal models that are constructed in safety assessment often attempt to consider the aleatoric uncertainty (i.e. uncertainty related to quantifiable randomness) that is associated with real-world events through a probabilistic treatment. For example, models (such as from operational experience) can be used to predict failure rates associated with physical failure events, which in turn can lead to probabilistic estimates of consequential outcomes. However, this is not the focus of this thesis. Instead, we are concerned with the epistemic uncertainty (i.e. relating to knowledge) associated with the risk assessment models (ram-uncertainty) we have created, the potential impact that such ram-uncertainty may have on decisions based on such models, and the necessity of identifying, documenting and tracking this ram-uncertainty.

For comparison, we look at how Knight [13], a researcher from the economic domain, associate uncertainty with risk (which we shall call economic risk). In his seminal work on

"Risk, Uncertainty and Profit", Knight differentiates economic risk and uncertainty by the presence or absence of knowledge regarding the possible outcomes to a given situation, such that one can accurately measure the probability of an outcome in a quantitative manner. When all possible outcomes and their associated probabilities of occurrence are known, Knight considers it as a situation where economic risk can be assigned and the odds calculated with a probabilistic treatment. In safety assessment, this is akin to treating a real-world event with aleatory uncertainty such as calculating the probability of an outcome when rolling a dice. As for the term - uncertainty, Knight uses it to describe the opposite situation whereby one is unable to know everything needed in order to calculate the odds. Knight believes that economic risk cannot be assessed in such situation since it will involve subjective judgement rather than using probability. We summarise Knight's and our understanding of risk and uncertainty in Table 1.

Table 1. Different Perspectives Regarding Risk and Uncertainty

	Type of risk we are predicting	How is aleatory (random) uncertainty being treated?	How is epistemic (knowledge) uncertainty being treated?
Our Perspective	Safety Risk, such as in aviation domain. Predicting safety risk from a set of chosen models that represent the real world, which are inherently uncertain.	Use quantitative probabilistic treatment to calculate safety risk from models with aleatory uncertainty. (Example: predicting failure rate from statistics of physical failure events)	Either (1) not being treated or (2) use assumption subjectively to scope the situation when predicting safety risk. (Example: with insufficient knowledge about a system design, either (1) ignore it or (2) make assumption such that the missing knowledge is not important when predicting system safety risk.)
Knight's Perspective	Economic Risk, such as in insurance domain. Predicting economic risk using models that have perfect knowledge of all possible outcomes	Use quantitative probabilistic treatment to calculate economic risk from models with aleatory uncertainty. (Example: predicting major illness occurrence rate from statistics of actual illness)	Treat as uncertainty instead of as economic risk. (Example: insurance usually does not cover extreme sports as there are too many unknowns regarding the danger of such activities)

When there is epistemic uncertainty, Knight chooses not to consider it as a situation suitable to predict the economic risk. This is perhaps plausible in the economic domain, such as when an insurance agency decides on the finite situations that it wants to insure (usually those with perfect knowledge as much as possible). This is not the same as predicting safety risk where a safety analyst does not have control over the level of knowledge available during the safety assessment and is typically being asked to make a priori predictions about future events. Furthermore, the ram-uncertainty we are concerned with is not determined by the amount of information that is (theoretically) knowable about a situation or a system. Instead, it depends on the actual knowledge that safety analysts have at the point of conducting the safety assessment.

For example, in 1995, eight Chinook helicopters were ordered by the UK MoD from Boeing but the MoD did not contractually purchase the required software source code in order to assess the safety risks that were needed to validate the airworthiness of the helicopters [14]. While this information is theoretically knowable (i.e. it could be provided by Boeing), it was not actually available to the safety analysts in the UK MoD. However, the lack of software source code did not cause the helicopter to be more unsafe. Instead, it created more uncertainty when predicting the safety risk, which affect the confidence in the safety assessment.

To conclude, safety analysts often do not have perfect knowledge of the underlying models (system and safety causal models) that are used to predict safety risk. Hence, there will inherently be epistemic uncertainty. In our research, we are pragmatically concerned with what the safety analysts know (and do not know) at the point of predicting the safety risk. In other words, the focus is on the quality of the knowledge that the safety analysts have when assessing the safety risk.

In our research, we consider two possible scenarios where uncertainties can be managed together with risk (see Figure 2).

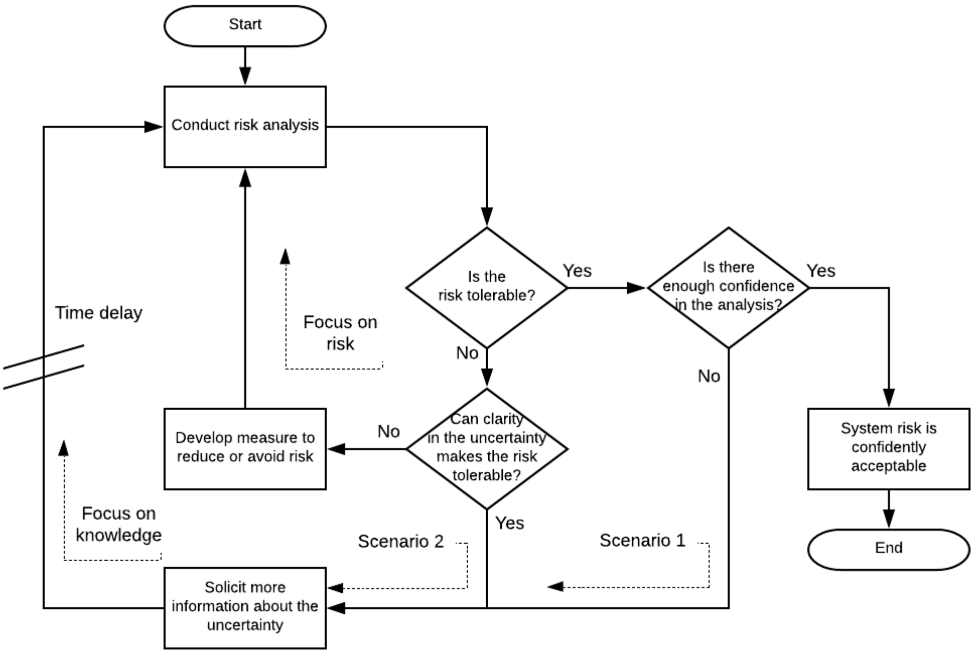

Start
Conduct risk analysis
Is the risk tolerable?
Yes
No
Is there enough confidence in the analysis?
Yes
No
System risk is confidently acceptable
End
Can clarity in the uncertainty makes the risk tolerable?
No
Yes
Develop measure to reduce or avoid risk
Solicit more information about the uncertainty
Time delay
Focus on risk
Focus on knowledge
Scenario 2
Scenario 1

Figure 2. Scenarios to Consider both Risk and Uncertainty in Safety Assessment

In scenario 1, while the assessed risk is tolerable, there may not be enough confidence in the risk analysis due to known uncertainties associated with the safety assessment. Hence, it is important to solicit more information about the uncertainties so that the system risk can be confidently accepted. For example, there could be an acceptance test result that shows a system is safe. However, the equipment used for the testing may not be certified and this reduces the confidence in the analysis. Efforts may be needed to verify that the result generated from the test equipment can indeed be trusted.

In scenario 2, the assessed risk may not even be tolerable due to either the presence of known uncertainties or the concern about uncertainties that the safety analysts are not even aware of (i.e. unknown uncertainties). Hence, there may be a need to solicit more information to discover plausible uncertainties that the safety analysts may not have considered before. For example, a system is expected to operate with another legacy system but there is insufficient information about the legacy system to make a comprehensive safety assessment. There could be uncertainties about how both systems are going to interoperate safely and efforts would be needed to identify and track these uncertainties until there is sufficient information.

Uncertainties should not be managed retrospectively after the safety assessment as it may not be possible or efficient to solicit comments from the various stakeholders about the uncertainties associated with the assessment. Furthermore, the stakeholders involved in the safety assessment may have already forgotten their rationales for accepting or not accepting a condition based on its uncertainties. This is made worse as uncertainties during safety assessment are often not documented and there will be few records of such considerations after the safety assessment for any retrospective tracing. Hence, an uncertainty should be managed as soon as it is being discovered during a safety assessment. While this may sound trivial, it underlines the importance of tracking uncertainties in parallel with risk analysis.

1.3 Research Strategy

According to Kothari [15], in his book on research methodology, there are two basic approaches to research - quantitative or qualitative. In our research, we depend largely on soliciting qualitative concerns from safety analysts and subject matter experts to develop a practical proposal to support safety assessment. As described by Kothari, qualitative research focuses on '*subjective assessment of attitudes, opinions and behaviours'*. Our qualitative research can be described as having two complementary loops (divergent and convergent) comprising four phases: observe, assemble, construct and evaluate. Such divergent-convergent framework is not new and supported by researchers, such as the work by Aviña et al [16]. In their research on 'Engineering a better future', Aviña et al emphasised that '*interactive divergent (idea generation) and convergent (idea test and selection) thinking are the fundamental processes underlying research*'. The outcomes and key activities for these four phases are elaborated in Table 2. The common factor that binds both loops is the hypothesis, which in our research becomes the thesis proposition (see Figure 3).

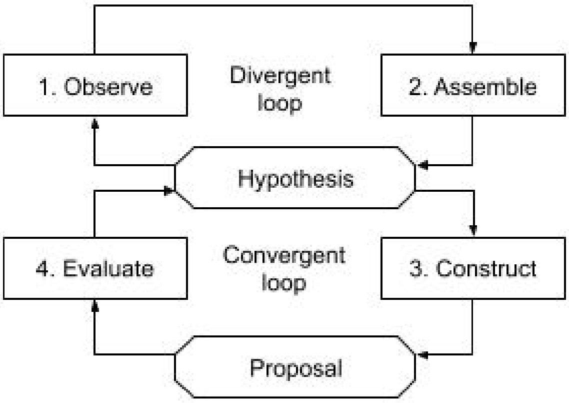


Figure 3. Complementary Divergent and Convergent Loops in Our Research

Table 2. The Four Phases under the Research Methodology

Phase	Outcomes	Key Activities
Observe	Observations to frame the problem space	Construct initial problem definition Conduct preliminary literature survey Interview stakeholders
Assemble	Ideas to scope the hypothesis	Capture ideas on the gaps in safety assessment Capture ideas on the gaps in managing uncertainty
Construct	Approach to implement the proposition	Develop the steps, processes and models that support the approach
Evaluate	Feasibility and future growth for the approach	Conduct qualitative feedback Conduct quantitative feedback

Divergent Loop: Observe and Assemble. As shown in Figure 3, the first iterative loop in our research is to observe the problem space and assemble possible focus areas, before we eventually hypothesise a thesis proposition. It is divergent in nature since the aim is to cast the considerations as wide as possible so as to appreciate the challenges facing safety analysts when they conduct safety assessments.

- **Observe**. The initial problem definition referenced the 2014 technical review from the Republic of Singapore Air Force. The review surfaced the need to develop expertise in managing safety assessment for air force systems that are highly integrated with each other. In parallel, we conducted literature surveys related to safety assessment and interviewed safety analysts to solicit feedback on the challenges they face. For example, during the interview, an air force safety analyst raised the concern that safety assessment is becoming more uncertain as it is getting more and more challenging to understand the complicated interactions between systems.

- **Assemble**. During the interview with the safety analysts, they were asked the question - '*What are the common concerns you have when conducting safety assessment?*'. Lack of knowledge or epistemic uncertainty has repeatedly been surfaced as one of the key concerns, as well as the lack of systematic approach to manage unknowns in a networked environment. Hence, we have chosen the management of epistemic uncertainty in safety assessment as our key focus in this thesis. Through the session, we have also further clarified that the uncertainty we are concerned with is not the uncertainty due to the absence of safety techniques, but rather it is about the uncertainty when conducting safety assessment with different safety techniques. It involves ram-uncertainty associated with not only the safety causal model, but also the system models that produces the safety causal model (as explained in section 1.2.3).

Convergent Loop: Construct and Evaluate. From the thesis proposition, we went into the second research loop (as shown in Figure 3) to construct our proposal by converging the observations into something useful and practical for the safety analysts. As we narrowed down on the area of uncertainty, we use the Kipling method [17] that advocates asking ourselves the 5W1H questions (i.e. what, who, where, when, why, how) in the domain of information gathering, to surface important questions we want to focus on. The *'why'* portion has been answered earlier through the rationalisation on the importance of uncertainty management in chapter 1, while the *'who'* and *'when'* are tactical questions downstream during implementation of initiatives. Hence, three important questions (what, where and how) remain that are related to the management of uncertainty. The methods used to tackle each question is summarised in Table 3.

Table 3. Research Methods to Tackle 'What', 'Where' and 'How'

Questions and Research Methods	Considerations
WHAT aspects of the epistemic uncertainties do we want to manage? Research Methods: Literature survey Observations from qualitative interviews	There are many aspects of epistemic uncertainty that we can focus on as described in detail under the literature survey in chapter 2. For example, under the research by Weick [18], he suggests to focus on the three areas of 'sensing', 'seizing' and 'transforming'. Another relevant research by Diana [19] looks at the areas of 'monitoring', 'understanding', 'deciding how to respond' and 'producing a response'. After our brainstorming session and interview with safety analysts, we have decided to narrow down to the major areas of identify, document, track and address - which mirrors closely the focus areas from the work of Weick and Diana. This grouping is intentional so as to simplify the concept to promote acceptance during implementation. The details will be elaborated in chapter 3.2.
WHERE do these epistemic uncertainties reside in safety assessment model? Research Methods: Modelling according to fundamental relationships and reference standards.	We aim to develop a common representation to consider epistemic uncertainty during safety assessment by defining the underlying models of safety assessment. The focus is on the epistemic uncertainties that can reside in both the system models in the system domain and the safety causal model in the safety domain (see Figure 4); and represent it using standard modelling relationships like association, inheritance and aggregation. External references have also been made which include the IEEE 42010 standard for architectural description [20], the 'Condition' concept developed by Wilson [21] and the Coleman's boat of causal pathway representation [22]. All these references help us to develop a way to represent epistemic uncertainty while complementing the current way of conducting safety assessment. The details will be elaborated in chapter 3.3.

HOW can we better manage the epistemic uncertainties as part of the system lifecycle? Research Methods: Literature survey Process development Goal Question Matrix (GQM) approach Refinement through case-study	The approach to manage epistemic uncertainties will take reference from the principles of identify, document, track and address. Different research methods are adopted for different principles. For identification, the focus is on doing literature survey to (1) solicit factors where uncertainty can reside and (2) how uncertainty can be categorised. In terms of documentation, track and address; there are two central research methods: process development and GQM approach. For documentation, the aim is to develop complementary processes to augment existing means of documentation during safety assessment (such as assumption documentation). During brainstorming, a concern was raised over the fear that the list of uncertainties being identified could be long. Hence, we need a systematic way of prioritising the list of identified uncertainties during documentation. As a result, the GQM approach [23] is adopted because it provides a structured way to develop factors to help in prioritisation. For track and address, similarly, we use both process refinement to complement existing processes of tracking safety related issues and the GQM approach to formulate questions for safety analysts to develop an action plan to conduct track and address of epistemic uncertainties. The details of how we propose to manage epistemic uncertainty will be elaborated in chapter 4. Further refinement to the approach is also carried out through case studies. These case studies are elaborated under chapter 5 and 6.

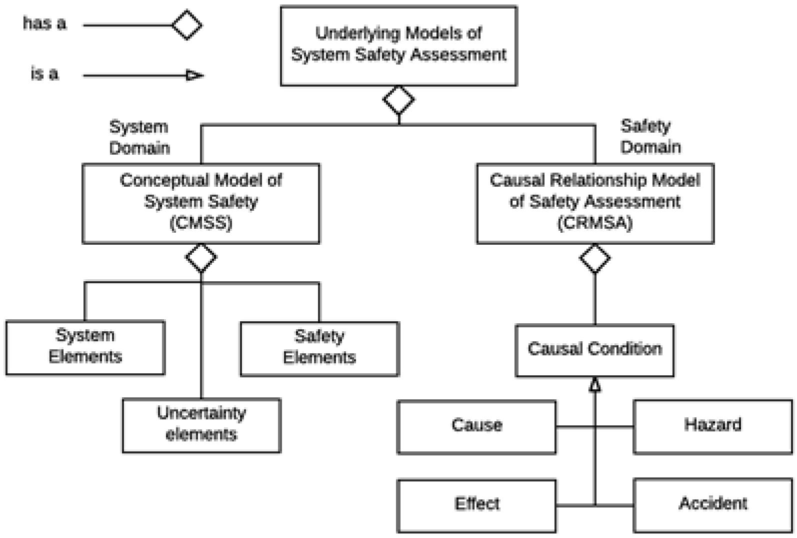

has a
is a
Underlying Models of System Safety Assessment
System Domain
Safety Domain
Conceptual Model of System Safety (CMSS)
Causal Relationship Model of Safety Assessment (CRMSA)
System Elements
Safety Elements
Uncertainty elements
Causal Condition
Cause
Hazard
Effect
Accident

Figure 4. Underlying Models of Safety Assessment*

* The CMSS concerns the system domain, while the CRMSA is specifically focusing on the safety domain. We will use these two models to recognise critical relationships in safety assessment where epistemic uncertainties may reside. The details of both models are captured in chapter 3.3.

Finally, we evaluate our proposal based on the thesis proposition to validate its feasibility to support safety assessment. The evaluation is based on two core activities of verification and validation. These two areas are proposed by the ISO 9001:2015 standard that is used to guide the evaluation if a product or service has met specify requirements. During the assessment, a semi-structured interview is chosen as the interview is very specific to a small pool of safety analysts. Hence, a set of open-ended questions, with a one-to-one interview, provides the flexibility to solicit qualitative feedback on the thesis proposal. Compared to a fixed set of questions, valuable feedback can be missing without a more in-depth discussion while conducting the questionnaire.

The remainder of this thesis is structured as follows: First, in the next section, we will state the hypothesis that is translated into the thesis proposition and explain the key terms in the proposition. In the following chapters of this thesis, we will describe the literature survey (chapter 2), theory (chapter 3), approach (chapter 4) and application (chapter 5 and 6) that are integral in constructing the eventual proposal. The final evaluation based on the thesis proposition is presented in chapter 7.

1.4 Thesis Proposition

Having considered the danger of uncertainties in section 1.1 and the relationships between risk and uncertainty in section 1.2, the thesis proposition is defined as follows:

> ***Epistemic uncertainties*** *in the* ***underlying models*** *of safety assessments for safety-critical systems can be* ***feasibly*** *and* ***systematically identified****,* ***documented*** *and* ***tracked through-life*** *in order to enable intervention to* ***address*** *potential risk.*

The following terms from the above proposition are worthy of explanation:

- **Epistemic uncertainties –** Uncertainties that are due to a lack of knowledge
- **Underlying models –** Representations of the real world for the purpose of analysis
- **Feasibly –** Able to apply the results within the context of existing industrial practice
- **Systematically –** Methodically, according to principles and concepts
- **Identify –** Establish or indicate what something is
- **Document –** Provide information or evidence that can be recorded for traceability
- **Track –** Follow the movement or development of something
- **Through-life –** From conceptual design to operation, to the end of life
- **Address –** Begin to deal with an issue

1.5 Thesis Structure

Chapter 2 surveys the literature that are relevant to this research, especially two important concepts that serve as important background knowledge: uncertainty and safety. Whilst there is various literature in each of these separate domains, the aim is also to find out the extent that uncertainty is being surfaced and managed in safety assessment (i.e. dealing concurrently with both domains of uncertainty and safety). The survey also focuses on the modelling of safety-critical system, which is the type of system that this research will be concerned with.

Chapter 3 describes the theoretical foundation to manage epistemic uncertainties in safety assessment after soliciting the knowledge and observations from the literature survey. It answers two questions: 1) "*WHAT aspects of the epistemic uncertainties do we want to manage?*" and 2) "*WHERE do these epistemic uncertainties reside in safety assessment models?".* To answer the first question – "WHAT", this chapter defines the aspects that are important to focus on in managing epistemic uncertainties by stating three principles that guide the development of the eventual approach: identify, document, track and address. The reasons that these principles are important are also presented. To answer the second question – "WHERE", two underlying models are constructed to describe the locations where epistemic uncertainties can reside in system and safety causal models. We use the uncertainties experienced while conducting safety assessments for new systems in the RSAF as a running example in this chapter, as well as subsequent ones.

Chapter 4 provides an approach to manage epistemic uncertainties in safety assessment based on three desired considerations: comprehensive, effective and feasible. The approach comprises three steps that are guided by the three principles mentioned earlier in chapter 3. In the first step, to help identify uncertainty, a taxonomy of causal mechanisms is consolidated to recognise known and unknown uncertainties not covered in the original safety assessment. In the second step, to help document uncertainty, process and factors to prioritise epistemic uncertainties analyses are developed to document such uncertainties in existing safety assessment techniques. In the final step, to help track and address uncertainty, a method to develop actionable goals is introduced to manage epistemic uncertainties that can be tracked through-life and addressed when some thresholds are met.

Chapter 5 describes the application of the approach from chapter 4 on existing safety assessment techniques. To cover a broad range of safety assessment techniques, the evaluation will focus on safety assessments conducted from both system and component viewpoints. In this chapter, the focus is on safety assessments from system viewpoint. The approach to manage epistemic uncertainties is integrated with the Systems-Theoretic Process Analysis (STPA) hazard analysis technique. This is based on the STPA analysis on the Yongwen railway system by Song et al [24].

Chapter 6 describes the application of the approach from chapter 4 on existing safety assessment techniques conducted from a component viewpoint. The approach to manage epistemic uncertainties is integrated with the Fault Tree Analysis (FTA) and Failure Modes and Effects Analysis (FMEA) techniques. This is based on the aircraft design example in ARP 4761 [7]. Unlike the system viewpoint that consider broader system issues such as technology and processes, component viewpoints can often create models that are more targeted such as an electrical circuit diagram or physical design specifications. The challenge is to be aware of the scope and limitations of each model, as well as appreciate the relationships between models that potentially can affect each other and become safety critical.

Chapter 7 evaluates the approach proposed in the thesis against the requirements given in Chapters 3, 4 and 5. It then evaluates the approach against the original thesis proposition given in Chapter 1. To evaluate the effectiveness and efficiencies of the proposed approach, peer reviews during sharing at conferences and semi-structured interviews have been conducted. Feedback from preliminary adaptation of the approach in the industry has also been solicited.

Chapter 8 concludes the research presented in this thesis and states the direction of possible future work.

Chapter 2 – Survey of Related Literature

The danger of uncertainties in safety assessments has been mentioned in chapter 1 with examples from the RSAF, CHALLENGER and PIPER ALPHA case studies. It is thus important to focus on how we can better manage such uncertainties when conducting safety assessments on safety-critical systems. To do that, three domains are identified in this literature survey: uncertainty management, safety management and modelling of safety-critical systems (see Figure 5).

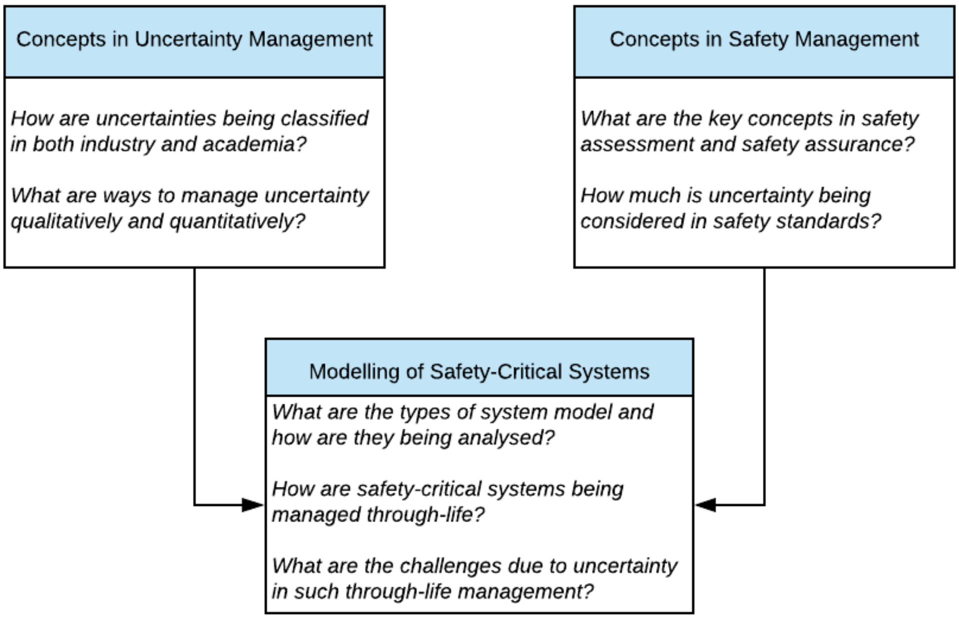

Concepts in Uncertainty Management
How are uncertainties being classified in both industry and academia?
What are ways to manage uncertainty qualitatively and quantitatively?
Concepts in Safety Management
What are the key concepts in safety assessment and safety assurance?
How much is uncertainty being considered in safety standards?
Modelling of Safety-Critical Systems
What are the types of system model and how are they being analysed?
How are safety-critical systems being managed through-life?
What are the challenges due to uncertainty in such through-life management?

Figure 5. Three Focus Areas in Literature Survey

Our survey has focused on the current developments and challenges related to these three domains. While our research focuses on uncertainty management in safety, there are useful literature related to our research in uncertainty management from other domains. We have therefore complemented our understanding by drawing upon definitions and concepts in other domains. In this aspect, we have widened the scope of this chapter by examining the literature on how other domains (e.g. economic, sociology) manage uncertainties. To begin with, we research into how uncertainty is being classified and managed currently (see section 2.1) and then we narrow our reviews on concepts related to safety management (see section 2.2). Finally, we survey the ways that safety-critical systems are currently being modelled (see section 2.3), keeping in mind the observations from the research on uncertainty and safety management.

2.1. Concepts in Uncertainty Management

The study of uncertainty is common in many domains, besides just safety assessment, as it occurs whenever we do not know something. While we acknowledge the presence of uncertainty, there isn't a common standard or approach to classify or treat uncertainty. For example, this is echoed by Walker [25] in his research into policymaking situation or scientific decision support. He states that while the existence of uncertainties is generally acknowledged, there is little focus on the "*different dimensions of uncertainty*" and "*different characteristics, relative magnitudes, and available means of dealing with them*". This section will provide an overview of the current classifications and treatments of uncertainty as a background for our research.

Before we embark on our research to manage the lack of knowledge, it is important to clarify the differences between the meaning of data, information and knowledge as used in this thesis. We based our definition closely to that provided by Jashapara [26] as follows (see Table 4):

Table 4. Comparison between Data, Information and Knowledge [26]

Type	Definition	Example
Data	known facts or things used as a basis of inference	(1) Physical observations from flight test (2) System behaviour at different temperature steps
Information	systematically organised data	(1) Flight test report that collates data from the flight test (2) System specification that states the temperature range for normal operation
Knowledge	actionable information	(1) Flight readiness review that uses information from the flight test report to identify hazards (2) System operating manual that uses system temperature specification to determine the cooling requirement for different operating scenario

In safety assessment, safety analysts have to make decisions both with the presence and absence of knowledge, which is actionable information being used in the context of the system being investigated. While our research focuses on the management of knowledge, it is also important to consider the impact when there is a lack of either data or information since both can affect the quality of the knowledge. Besides focusing on knowledge that we have, we also want to develop an approach that focuses on managing knowledge that we do not have (i.e. an identified epistemic uncertainty) during a safety assessment.

When relevant information about a system is not accessible to the safety analysts, they would run the danger of lacking system knowledge necessary to predict safety risk. For example, in his research on the NIMROD aircraft accident in September 2006 [27], Dogan highlighted many instances whereby there were safety documents (e.g. inspection and failure reports) available as information, which were not converted to practical knowledge as this information was either not sought out, inaccessible, or "*stored in places or database which are not readily accessible to those on Front Line*". Unfortunately, the right information wasn't provided to the right people at the right time to make the safety assessment more robust to prevent the accident.

2.1.1. Classifications of Uncertainty

In this session, we will highlight two common ways that uncertainty is being classified. The first is about the "aleatory vs epistemic" divide, while the second focuses on the differences between known and unknown uncertainties.

2.1.1.1. Aleatory versus Epistemic

Within the context of scientific computing, Roy and Oberkampf [28] classifies uncertainty into either aleatory or epistemic. He clarifies that *aleatory* is *"the inherent variation in a quantity that, given sufficient samples of the stochastic process, can be characterized via a probability density distribution",* while *epistemic* is the *"uncertainty due to lack of knowledge by the modelers, analysts conducting the analysis, or experimentalists involved in validation".*

For Roy's research in scientific computing, he mentions many sources of uncertainty such as *"model inputs, the form of the model, and poorly characterized numerical approximation errors"*. Such sources of uncertainty can either be aleatory, epistemic, or a mixture of both. Specifically, for epistemic uncertainty, Roy adds that the lack of knowledge can occur in various aspects such as lack of information about the system or its environment, during simulations or when collecting data to verify or validate the system through experiment.

Kiureghian **[29]** also classifies uncertainty as either aleatory or epistemic. He explains that the word aleatory is derived from the Latin word, *alea*, which refers to the "*rolling of dice"*. Following this argument, Kiureghian states that aleatory uncertainty refers to uncertainty that is due to the random nature of the subject of interest. On the other hand, the word epistemic is derived from the Greek word, *epistēmē*, which refers to "*knowledge"*. That's why epistemic uncertainty is uncertainty that is the result of a lack of knowledge.

Kiureghian cautions that it may not always be easy to distinguish between these two types of uncertainty. In his research, he mentions the challenges to distinguish between both types of uncertainty by mentioning that "*attempts to get a hold of the two concepts to make indisputable and unambiguous definitions seem to slip between the fingers"*. He further explains that it can be difficult to distinguish both types of uncertainties when modelling a system as it depends on the knowledge and experience of these who built the system model.

In summary, aleatory uncertainty refers to the inherent randomness in a system. According to Skinner et al [30], this form of uncertainty "*cannot be reduced, although additional research may help to better understand the complexities of the system of interest"*. Epistemic uncertainty refers to the imperfection of knowledge associated with a system. To Skinner et al, epistemic uncertainty may be "*quantified, reduced, and possibly eliminated, depending on the specific situation*".

Observations. In our research, we would assume the common definitions that aleatory uncertainty refers to the inherent randomness in the system of interest, whereas epistemic uncertainty is about the lack of knowledge at the point of analysis. As mentioned in chapter 1, we believe that it is possible to better manage risk during safety assessment by focusing on and being explicit about the knowledge that is lacking, i.e. the epistemic uncertainty. Hence, moving forward, our literature survey pays more attention to epistemic uncertainty.

2.1.1.2. Known versus Unknown

There is literature that present epistemic uncertainties by what are knowns and unknowns such as Logan [31], Daase and Kessler [32] and Chow [33]. Former United States Secretary of Defence Donald Rumsfeld made this classification of epistemic uncertainty famous during his news briefing in 2002. This was later captured in his memoir: Known and Unknown [34].

> *"Reports that say something hasn't happened are always interesting to me because as we know, there are known knowns; there are things we know we know. We also know there are known unknowns; that is to say we know there are some things we do not know. But there are also unknown unknowns – the ones we don't know we don't know."*

According to Rumsfeld, known knowns are *"facts, rules, laws we know with certainty"*. Known unknowns are *"gaps in our knowledge, but they are gaps that we know exist"*. The way to manage known unknown is to be able to query the status quo by asking the *"right question"*. Unknown unknowns are *"gaps in our knowledge, but they are gaps that we don't know exist"*. Rumsfeld quoted the 9-11 terrorist attack as an unknown unknown whereby no one would ever think about such a scenario before the tragedy strikes.

While Rumsfeld mentioned that it was impossible for anyone to expect the 9-11 terrorist attack, an aerial strike into a tall building could be a valid scenario being conceived by a team that is formed intentionally to identify possible terrorist strikes. An unknown unknown situation can be subjective since it depends on the amount of efforts and resources to focus on identifying the unknowns. Hence, we believe that by directing the attention to plausible unknown unknowns, we can help safety analysts to recognise unknowns that are related to the system under analysis so that they become known unknowns for further tracking.

Such an approach of classifying epistemic uncertainty has since become popular in many domains. For example, the German sociologists Daase and Kessler [32] agree with Rumsfeld by stating that the *“cognitive frame for political practice”* can be based on the relationship between *“what we know, what we do not know, what we cannot know”*. Besides the three categories, philosopher Slavoj Žižek [35] argues that there could be a fourth classification, unknown known, which represent knowledge that one *“intentionally refuse to acknowledge that we know”*. He believes that human beings have unknown knowns that represent *“beliefs, suppositions and practices that we pretend not to know about, even though they form the background of our public values"*.

One possible model where the known vs unknown classification could have been derived from is the Johari window [36]. The window is a 2x2 matrix with four perspectives (see Figure 6). It was designed to create self-awareness of the information being processed or communicated between multiple parties in an organisation.

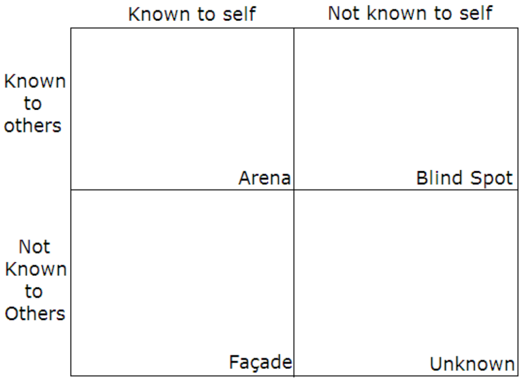


Figure 6. Johari Window

The first two perspectives on the left (Arena and Façade) are quadrants of knowledge that are known to oneself. Hence, these would most likely be the known knowns that do not lead to epistemic uncertainty. The top right quadrant, Blind Spot, refers to knowledge that is not known to oneself even though it may be known by others. This is like the known unknowns where epistemic uncertainty resides. The last quadrant on the bottom right refers to the unknown unknown situations where neither oneself or others are aware of the knowledge.

Observations. The Johari Window, while developed by Luft and Ingham [36] for self-help groups and corporate development, is a useful reference for our safety domain as it provides a matrix to articulate information that is known and unknown to an individual; as well as information that is known and unknown to others. Such a matrix is relevant for system safety as we depend on the information that is known to the safety analysts during the hazard identification to make our risk assessment. The differentiation between known unknowns and unknown unknowns is important in our research as it creates the awareness and need to manage both types of uncertainties. While known unknowns can be hazardous to a system, the unknown unknowns can potentially be hazardous too. Hence, it is important to engineer an approach to move potential hazards first from unknown unknown (i.e. totally not aware) into the known unknown (i.e. being aware of a situation but realise that one does not have sufficient knowledge about it) during safety assessment. Then, such known unknown needs to be tracked systematically until information is available to address the epistemic uncertainty such that the hazard becomes a known known in safety assessment.

2.1.2. Qualitative Modelling of Uncertainty

We begin the survey of existing qualitative approaches of modelling uncertainties by considering the key characteristics of epistemic uncertainty. This would follow by a survey of existing ways of managing uncertainty qualitatively. As a useful reference, we end the section by reviewing ways that uncertainty is being modelled in requirements engineering.

2.1.2.1. Characteristics of Uncertainty

In order to apply the right approach to manage epistemic uncertainty, one needs to make informed decision about uncertainty based on its characteristics.

In the context of industrial development, both McQuiston [37] and Johnston & Bonoma [38] refer to three attributes: complexity, novelty and importance to characterise uncertainty and determine its impact on the organisation. McQuiston refers to complexity as "*how much information the organisation must gather to make an accurate evaluation of the system*". Novelty is defined by him as "*the lack of experience of individuals in the organisation with similar situations.*", while Importance is considered as "*the perceived impact on organisational profitability and productivity*".

In their study of new product development for IT systems, Peng et al [39] and Novak & Eppinger [40] uses the Organisational Information Processing Theory (OIPT) to determine the characteristics of uncertainty. According to the OIPT, different tasks are expected to have different degree of uncertainty. It is further explained that "*higher uncertainty implies higher* *variability in and unpredictability of exact means to accomplish the task, in turn leading to poorer task outcomes*". In their study of inter-organisation supply chain, Bensaou and Venkatraman [41] explain that the OIPT implies that *"an organisation must design appropriate structural mechanisms and adopt the right technologies and practices to provide the information processing capabilities that meet the organisation's information processing needs".*

For project management, both Peng and Novak conclude that information processing needs are affected by product size and task interdependence. Product size refers to the number of parts in the product design and task interdependence refers to the influence of any given task on other tasks. Specifically, Peng [39] also adds a third factor – project novelty which includes novelty of product or process, lack of information about markets and customers and the ambiguity of project goals. Galbraith [42] also uses the Organisation information processing model to emphasize that the greater the task uncertainty, the greater the amount of information that must be processed to achieve a given level of performance.

Separately, in their study of product development projects, Tatikonda and Rosenthal [43] consider uncertainty to vary along two dimensions: technology novelty and project complexity. Tatikonda refers to technology novelty as the "*newness to the development organisation of the technologies employed in the product and process development effort". He defines project complexity as "the nature, quantity and magnitude of organisational subtask and subtask interactions posed by the project"*. He adds three attributes under project complexity, namely project difficulty (level of task performed), objective novelty (novelty of task objectives) and technology interdependence (interdependency of task units). Interestingly, Tatikonda classifies novelty as a component under project complexity which is different from the classification by McQuiston, which was explained earlier.

Also, in the domain of project management, Shenhar and Dvir [44] introduce the "*diamond framework*" to identify and manage uncertainty by considering four project characteristics: "novelty, technology, complexity and pace". They consider novelty as "*how intensely new are crucial aspects of the project"*, while complexity is about "*finding out how complicated are the product, the process and the project involved"*. They also mentioned that having an awareness of the characteristics is not enough if there is no time and resources to learn and improve.

Other categorisations include the "*learnability, multiplicity, temporality, complexity, uncertainty and sociability*" factors introduced by Svejvig & Anderson [45] under their rethinking project management initiatives. Separately, Saunders et al [46] refers to determinants of uncertainty based on "*environmental, individual, complexity, information, temporal and capability*". Like Shenhar, both Svejvig and Suanders agreed that having the resources to learn are important besides being aware of the characteristics of uncertainty.

Observations. From the above survey, characteristics such as complexity and novelty are the common factors exacerbating uncertainty. Such characteristics would be useful subsequently in our research when we attempt to qualify uncertainty in safety assessment.

2.1.2.2. Management of Uncertainty

In this section, we look at existing approaches to manage uncertainty.

Cynefin framework. One such approach is the Cynefin framework that was developed by Snowden and Biine [47]. The framework aims to help decision makers propose different strategies to address issues under different levels of uncertainty. The framework comprises five contexts to help decision makers manage uncertainty. The first four contexts are simple, complicated, complex, and chaotic. The fifth context, disorder, refers to the situation when it is not clear which of the other four contexts are the most dominant. These five contexts are summarised in Table 5.

Table 5. Decision Making in Different Context under Cynefin Framework

Context	Characteristics	Example
Simple	Clear cause-and-effect relationships that are easily observable by everyone, e.g. known-knowns	Loan payment processing
Complicated	Clear cause-and-effect relationships but not everyone is aware and may have multiple "right answer" (i.e. multiple causes leading to same effect), e.g. known-unknowns	Working of a Ferrari sports car
Complex	Cause-and-effect relationships may not be apparent to anyone, e.g. unknown-unknown	Rainforest ecosystem
Chaotic	Impossible to determine cause-and-effect relationships as there is no manageable pattern, e.g. Unknowable, (based on current knowledge)	Sept 11, 2011 attack on the World Trade Centre
Disorder	Difficult to recognise which of the other four contexts is predominant	Leaders arguing with one another

It is interesting to note that the September 11 incident is classified as chaotic, which is unknowable, by Snowden, whereas Rumsfeld considered the 9-11 terrorist attack as an unknown unknown in section 2.1.1.2. One reason could be that Rumsfeld may have knowledge that was not available to Snowden. Another reason could be that Rumsfeld did not create a fourth category known as "unknowable". This again demonstrates that there are multiple ways of categorising different situations depending on individual's knowledge and experience.

In the case of safety assessment, we have assumed that most of the causal relationships related to safety can be derived, given the right knowledge. However, such causal relationships may not be known to everyone, especially during the safety analysis due to time and resource constraints. Hence, we classified it under the complicated domain according to the Cynefin framework. However, as a system evolves, it is possible for its complexity to evolve such that it exhibit characteristics of different context, such as simple or complicated.

Pacing of Experiences. Under the Dynamic Capabilities Theory, Eisenhardt & Martin [48] advocate that "*resources and capabilities must be constantly reallocated and reoptimized to adapt to changing environment*". Using a "*resource-based"* view of organisation, Eisenhardt et al say that an organisation should constantly monitor changes due to uncertainty and be ready to shift resources using learning mechanisms such as "*practice, codification, mistakes, and pacing*". Specifically, the concept of pacing is worth elaborating. It refers to the pacing of experience so that it is conducive for learning and gaining knowledge. Argote [49] explains that "*experience that comes too fast can overwhelm managers, leading to an inability to transform experience into meaningful learning".* In the context of uncertainty management, if we equate experience to information gain, this could serve as a cautious not to overwhelm managers with a long list of uncertainties to the extent that they will not be able to analyse this massive information. Argote further adds that *"similarly, infrequent experience can lead to forgetting what was learned previously and so result in little knowledge accumulation as well*". For uncertainty management, if most uncertainties are not tracked, managers may not be exposed regularly to the presence of these uncertainties and could eventually forget about them later in the system lifecycle.

Managing Surprises. In a complicated system, confidence can be undermined by the element of surprise due to uncertainty. Aven [10] states that uncertainty leads to surprises and one of the ways to reduce surprises is to reduce the undesirable impact of surprises on safety. Traditionally, uncertainty is measured under risk assessment as part of the "likelihood" or probability of harm. Aven argues that this is incomplete as probability of harm is derived from a finite set of known past occurrences, which does not consider unknown events. He introduces the phrase "*ignorance of unknown events*" to describe a "*lack of understanding of how the consequences of the activity are influenced by the underlying factors".* While one's knowledge will neither be completely deterministic nor total ignorance [25], the lack of complete awareness and knowledge of a system and its environment will always limit the accuracy of safety assessment.

Aven assumes that the goal of a decision maker is to reduce the undesired impact of surprises due uncertainty, instead of expecting to eliminate them. The challenge, he added, was on how to communicate uncertainty between multiple stakeholders, especially between system engineers and management, so as to make more people aware of safety related issues. While we agree with Aven's pragmatic approach of acknowledging the presence of epistemic uncertainty when making decision, it is also the ethical responsibility of safety analysts and system engineers to reduce the epistemic uncertainty as much as possible. We want to develop a proactive approach to reduce the epistemic uncertainty to as low as reasonably practical. At the same time, strive for a balance between being aware of the uncertainty and reducing the uncertainty through life.

Sensemaking. Another area of research relevant to manging uncertainty is the domain of sensemaking. As explained by Weick [18], sensemaking in an organisation refers to the efforts to "*develop information processing mechanisms capable of detecting trends, events, competitors, markets, and technological developments relevant to their survival".* One of the frameworks to operationalise sensemaking in an organisation is introduced by Petit & Hobbs [50]. They build a conceptual framework to study project uncertainty by recommending three core activities for sensemaking. These three core activities are sensing, seizing and transforming.

- "sensing" refers to the processing of information in the environment related to an uncertainty,
- "seizing" refers to the efforts to identify and decide whether changes are needed with respect to what has been sensed earlier, and
- "transforming" refers to the actual action to change the "routines of the enterprise".

Sensemaking can be used to engineer the tracking of known unknowns until information is available to eliminate the uncertainty to the extent that it becomes a known known suitable for safety assessment. The three activities of "sensing", "seizing" and "transforming" are useful reference to develop the approach to manage the epistemic uncertainty through-life.

In another study, Saunders et al [51] also attempts to define key approaches to manage uncertainties under project management. In their research, they recommend three key approaches - *structural, behavioural* and *relational*, to help managers prepare for uncertainties in project life and describe ways to identify, analyse and act on the uncertainties. Separately, Diana [19] mentions that uncertainty would be the key impetus to drive the need to do sensemaking in the risk and security domains. He elaborates that "*risk signals that we can capture (sense) provide us with an opportunity to mitigate that risk (respond)*". The sensemaking processes proposed by Diana involve "*monitoring and observing, incessant rehearsing, understanding and interpreting data, deciding how to respond, and producing a response*". While most are intuitive, "*incessant rehearsing*" is an activity that is worth explaining. It refers to the continual discovery of possible scenarios and be prepared for them if they ever become reality.

Observations. Using the taxonomy from the Cynefin framework, we consider safety-critical systems as complicated systems whereby there are definite cause-and-effect relationships. The challenge then is to identify such cause-and-effect relationships in the presence of epistemic uncertainties. While the literature in this section bring out the areas to focus on to manage uncertainty (e.g. surprises, experiences, sensemaking), none of them focus on implementing methods to manage uncertainty. The closest to an implementable approach is the introduction of the three core activities (sensing, seizing and transforming) by Petit & Hobbs [50] but even that has been presented as a strategic framework, instead of a tactical method. Moving forward, we would consider customising the research under the domain of sensemaking and build a practical and systematic approach to manage uncertainty in our context. We will study this in greater detail in chapter 3 and 4.

2.1.2.3. Modelling of Uncertainty in Requirements Engineering

As we focus on the underlying models of safety assessment, it is useful to review research in the area of uncertainty modelling. In particular, the treatment of modelling uncertainty in requirement engineering presents some useful contributions. In particular, the following three areas of research present relevant insights: partial modelling, under specification and subjective modelling.

Partial Modelling. The concept of partial modelling is introduced by Famelis and Chechik [52] for model-based software development that involves multiple software design options that have yet to be finalised. The uncertainty in this context refers to the lack of knowledge or confirmation on which solution will be chosen out of the multiple ways of developing the software. It was highlighted that traditional software developers are "*comfortable with using models to express information about software but not good in expressing uncertainty and reasoning about uncertainty with such models*". In partial modelling, uncertainty is expressed by having multiple of such alternate partial models that represent other possible scenarios in view of uncertainty. For reasoning, annotations (TRUE, FALSE and MAYBE) are used to label the elements in the model to measure the level of uncertainty. Annotation TRUE or FALSE implies that the element is either chosen or not chosen based on the available information, while MAYBE refers to element that the developers are unsure if it will be chosen due to lack of knowledge. The partial model research by Famelis et al emphasises that besides identifying and annotating uncertainty, there must be an approach to reason about the uncertainty and integrate this approach into existing processes. There is similar concern in our research whereby we are also concerned with the way to express uncertainty in the underlying models for safety assessment, as well as the approach to reason about this uncertainty during hazard identification. The partial model uses a sequential process of "*construction,* *verification, diagnosis and refinement*" to reason and manage uncertainty. Similarly, our research will also consider developing an appropriate process to manage uncertainty for safety assessment.

Underspecification. Underspecification in requirements engineering can be observed in software policy and standards. This may lead to ambiguity. An example is the assessment by Ferrari et al [53] that there are over 18% of underspecified sentences in the Standard EN50128:2011 Software for railway control and protection systems. Reading across to our safety domain, we need to be equally mindful that safety policy and standard can also be ambiguous, and this can unintentionally lead to uncertainty when they are being referred to during safety assessment. In software engineering, another aspect of underspecification is when certain variables in the software model (e.g. in design choice, software configuration, implementation choice) are omitted either intentionally or unintentionally. Such omission could be due to a genuine lack of knowledge or intention to allow flexibility for subsequent interpretation. Papavassiliou and Mentzas [54] works on handling uncertainty due to underspecification in "weekly structured" software business processes. In their approach, work flows that may not be well articulated shall be labelled as black boxes using the common UML notation. When more information is available, the actual work flow will then be added to the model to complete the specification of the task in the model. Such method of tracking the uncertainty by underspecified the model and subsequently address the uncertainty by inserting the actual work flow are good reference for our research when we devise the approach to manage uncertainty throughout a system life cycle. The topic of underspecification will be revisited in section 3.2.2.2 of this thesis regarding the concern of documenting uncertainty.

Subjective modelling. A third area where uncertainty is being modelled is known as subjective modelling. In subjective modelling, fuzzy logic is used when there is uncertainty in the assessment of variables that are not definitive at time of decision making. For example, Wang [55] introduces subjective modelling in ship safety assessment as safety analysts often can only use subjective descriptors to describe the safety associated with an event. In his research, three parameters (failure likelihood, consequence severity and failure consequence probability) are assessed subjectively. Each parameter is described using "*subjective linguistics variables*" such as failure likelihood ("*highly frequent", "frequent", "reasonably frequent", "aver-age", "reasonably low", "low" and "very low*"); consequence severity ("*catastrophic", "critical", "marginal" and "negligible*"); and failure consequence probability ("*definite", "highly likely", "reasonably likely", "likely", "reasonably unlikely", "unlikely" and "highly unlikely*"). Each linguistic variable will be further described using a membership function, according to fuzzy theory. Wang states that such subjective safety analysis "*provides marine safety analysts with flexibility in articulating judgements produced* *by multiple safety analysts*". Another similar example is the research by Ozdamar and Alanya [56] on software project scheduling. Ozdamar et al explains the need to manage uncertainty due to a lack of precise knowledge in software development project. They chose fuzzy set theory as it "*enables modelling the uncertainty associated with vagueness, with imprecision, and/or with lack of information about the system*". As seen by both examples, subjective modelling is useful for analysis where events, scenarios and activities are mostly known but there are uncertainties to either which will occur or in what severity or frequency. In other words, these are known uncertainties and the focus is on deriving fuzzy terms to reduce linguistics ambiguity. While this is important, our research is also concerned with the lack of knowledge to identify such events, scenarios or activities in the first place. There are unknown uncertainties we want to focus on in our research, beyond minimising any linguistic ambiguity.

2.1.3. Quantitative Modelling of Uncertainty

Besides qualitative modelling of uncertainties, there are also attempts to quantify uncertainties. While qualitative assessment focuses on describing the characteristics of the uncertainties, quantitative assessment focuses on calculating the amount of uncertainties based on certain measurement. In this section, we have surveyed the use of mathematical formula and matrix to quantify uncertainties.

Use of Mathematical Formulae. The use of statistical probability in risk assessment is one of the common types of mathematical formulae in the safety domain. This is also used in treating uncertainty quantitatively. For example, in the research by Mensing [57], he uses statistical approximation to model epistemic uncertainties quantitatively as part of his efforts to measure the risk of explosion at a military site. He derives mathematical formulae by representing the epistemic uncertainties of specific parameters using lognormal distributions. For example, he modelled the variation in the quantity of explosive, E, using a lognormal probability distribution:

$$E = Eo^{*}\ \delta e$$

In this equation, Eo refers to the median daily number of exposures, and δe is a lognormal random variable.

However, in order to do that, there is a need to have a *"mathematical/probabilistic model of the environment and risk source of interest"* or what Mensing terms as the *"model of the world (MOW)"*. The MOW involves building models to represent the real world. For example, for the explosion case, Mensing suggested the need to define MOW for *"physical characteristics of the explosive site structures, temporal distributions of the quantity of explosives, number of* *personnel at the site, occurrence and magnitude of explosive events, severity of the explosive effects and occurrence of fatalities at the explosive site".*

In his concluding remarks, Mensing acknowledges that *"successful application of the methodology relies, significantly, on development of the MOW and modelling of the epistemic uncertainties associated with the MOW"*. He emphases that the success of estimating the risk relies on *"how well the MOW depicts the actual physical environment being analysed"*. Unfortunately, there were no further details on the way to "*develop the MOW or model the epistemic uncertainties*". There were no details in the research about how such MOW can be constructed and be considered as sufficient to represent the real world.

In their work on environmental and health risk assessment, Hammonds et al [58] presented guidelines for evaluating uncertainty using mathematical equations and computer models. They analyse uncertainty quantitatively to estimate the confidence that can be placed in a risk estimate. Their analysis focuses on two key aspects: 1) defining the parameters that form the uncertainty model and 2) specify the probability distribution function of these parameters. Hammonds et al advocate the use of numerical methods to develop the probability distribution for the uncertainty model. The suggested numerical techniques to estimate the uncertainty include "*variance propagation, Monte Carlo simulation, differential uncertainty analysis, non-probabilistic methods such as fuzzy logic and first-order analysis employing Taylor expansions*".

One of the key challenges of quantitative uncertainty analysis discussed by Hammond et al is the inability to define the boundary or extent of the environment that can possibly influence a system. For example, they mention that "*if the characterization of the nature and extent of the amount of contamination in a given environmental media at a site is inadequate to permit even a bounding estimate (an upper and lower estimate of risk), a quantitative uncertainty analysis cannot be performed*". Another difficult task in quantitative uncertainty analysis is to support the "*judgmental decisions that are made to obtain subjective probability distributions for the uncertain model parameters*". Hammonds et al suggest that soliciting the views from more experts would help to "*defensibly estimate parameter and model uncertainty*".

Another approach of modelling uncertainty quantitatively is to use Bayesian probability, such as the Bayesian Belief Network. As Goldstein [59] explains, Bayesian approach is to "*quantify your uncertainties as probabilities, for the quantities you are interested in, and conditional probabilities for observations you might make given the things you are interested in. When data arrives, Bayes theorem tells you how to move from your prior probabilities to new conditional probabilities for the quantities of interest.*" Such a network of Bayesian probabilities can be applied in practice and is commonly known as Bayesian Belief Network. One of such example is the work by Kinder [60] for system of systems, which is a type of complicated system. Kinder uses Bayesian probability to model the system and applied the Monte Carlo simulation to generate the probability so as to manage the uncertainty when historical data is not available.

Use of Matrix. An example of a matrix most safety analysts are familiar with is the risk assessment matrix that use probability and severity to measure the risk of a hazardous situation. In his work on quantifying uncertainty, Aven [61] proposes a similar matrix to measure "*uncertainty factor*" based on risk and vulnerability as shown in Table 6. If an uncertainty factor is considered to have significant effect on either risk or vulnerability, more efforts would be allocated to analyse and manage that uncertainty. Aven considers safety assessment as affected by two types of uncertainty: (1) uncertainty about the occurrence of events and (2) uncertainty about the consequences if an event occurs.

Table 6. Uncertainty Factor (extracted from Aven)

e.g. Risk of ignition	Effect of risk			Effect on vulnerability		
Uncertainty factor	Minor	Moderate	Significant	Minor	Moderate	Significant
gas concentration		x			x	
no. of persons		x				x

Aven further explains that the assessments on the effect of risk and vulnerability are based on a specific *"background knowledge*". The assessments themselves are not uncertain. Instead the uncertainty lies in the *"background knowledge*", which is closely related to the assumptions made when quantifying the uncertainty. For example, under Table 6, the probability that defines the risk of ignition is not the main concern of uncertainty. Instead, the uncertainties lie in the amount of information available to determine the gas concentration and the number of persons in the affected areas. According to Aven, "*surprises relative to the assigned probabilities could occur if the background knowledge on which the probabilities are conditioned turns out to be wrong*".

Observations. An in-depth analysis of the mathematical methods discussed in this section is beyond the scope of our research on managing epistemic uncertainty. While there is a similar intent of managing uncertainty, quantification of uncertainty usually occurs at a more localised level of looking at specific parameter, instead of looking holistically at a system model. For example, in Aven's case study, since gas concentration and the number of persons can be measured, quantifying such uncertainties would be a natural approach. However, in hazard identification, we may need to manage uncertainty beyond just a localised area of focus, which may not be possible to quantify. Furthermore, such quantitative analysis assumes hazards are either known knowns or known unknowns. There is no provision in the analysis to consider the presence of unknown unknowns that may invalidate the assessment.

While quantitative uncertainty analysis is important, we are more concerned in our research with appreciating the system environment that underlies the context from which such quantitative uncertainty is derived. Such context is defined differently in different literature. For examples, Mensing [57] defines it as the model of the world, Aven [61] calls it the qualitative assessment of the situations that constitute the “background knowledge” and Hammonds et al [58] refers to it as the qualitative analysis to characterise the environment. However, all of them agree that such context cannot be assessed quantitatively. Instead, they have to consider such context in a qualitative manner and epistemic uncertainty is expected as they will not have perfect knowledge of this context. Hence, it is evidenced that qualitative assessment of the context is crucial even in quantitative treatment of uncertainty. This motivates us to focus our research on the qualitative aspect of epistemic uncertainty.

2.2. Concepts in Safety Management

Besides uncertainty, the other key focus in our research is safety management. Two of the key concepts in safety management are safety assessment and safety assurance. In our context, we define safety assessment as the processes and techniques to conduct safety analysis, while safety assurance is concerned with the degree of confidence over the results obtained from the safety assessment. In this section, we will first describe the definitions, processes and techniques in safety assessment under section 2.2.1. Next, we would describe the concept of safety assurance in section 2.2.2. Finally, in section 2.2.3, we would survey the extent that uncertainty and its management are explicitly mentioned in existing safety standards.

2.2.1. Safety Assessment

One challenge of managing safety is the lack of standardisation in the definitions, standards and techniques to conducting safety. In this section, we introduce the common definitions in safety assessment. This will be followed by a survey of safety assessment processes across a system lifecycle. To conclude, we describe safety techniques that are currently being applied to identify hazards that are relevant to our research.

2.2.1.1. Definitions

An example of a lack of standardisation is the differences in safety terminology from different literature. For comparison, the terminology of safety from DoD MIL-STD-882E on System Safety [62] and the UK MoD Defence Standard 00-56 [63, 64] are presented in Table 7. It is interesting to note that different versions of a standard may define the same term differently. This is evidenced in the comparison of issue 4 and issue 6 of Defence Standard 00-56 (see Table 7).

Table 7. Terminology of Safety

Terms	Def Stan 00-56 (issue 6)	Def Stan 00-56 (issue 4)	MIL-STD-882E
Assurance	(not defined)	Adequate confidence and evidence, through due process, that safety requirements have been met	(not defined)
Safe	Freedom from unacceptable or intolerable levels of harm.	Risk has been demonstrated to have been reduced to a level that is ALARP and broadly acceptable or tolerable, and relevant prescriptive safety requirements have been met, for a system in a given application in a given operating environment.	Freedom from conditions that can cause death, injury, occupational illness, damage to or loss of equipment or property, or damage to the environment.
Risk	(same as issue 4)	Combination of the likelihood of harm and the severity of that harm.	A combination of the severity of the mishap and the probability that the mishap will occur.
ALARP (As Low As Reasonably Practicable)	(defined in separately document)	A risk is ALARP when it has been demonstrated that the cost of any further Risk Reduction, where the cost includes the loss of defence capability as well as financial or other resource costs, is grossly disproportionate to the benefit obtained from that Risk Reduction.	(not defined)
Safety Requirement	(same as issue 4)	A requirement that, once met, contributes to the safety of the product, service or system or the evidence of the safety of the product, service or system.	(not defined)
Harm	Adverse impact on people, including fatality, physical or psychological injury, or short or long term damage to health.	Death, physical injury or damage to the health of people, or damage to property or the environment.	Mishaps: An event or series of events resulting in unintentional death, injury, occupational illness, damage to or loss of equipment or property, or damage to the environment.
Accident	(same as issue 4)	An event, or sequence of events, that causes unintended harm.	
Incident	(same as issue 4)	The occurrence of a hazard that might have progressed to an accident but did not.	
Hazard	(same as issue 4)	Potential to cause harm, e.g. A physical situation or state of a system, often following from some initiating event that may lead to an accident.	A real or potential condition that could lead to an unplanned event or series of events (i.e. mishap) resulting in death, injury, occupational illness, damage to or loss of equipment or property, or damage to the environment.
Safety Case	(same as issue 4)	A structured argument, supported by a body of evidence that provides a compelling, comprehensible and valid case that a system is safe for a given application in a given operating environment.	
System	(same as issue 4)	A combination, with defined boundaries, of elements that are used together in a defined operating environment to perform a given task or achieve a specific purpose.	The organization of hardware, software, material, facilities, personnel, data, and services needed to perform a designated function within a stated environment with specified results.

From Table 7, we observe that many of the terms are related to each other. The key ones are summarised here:

- *Safety assurance* is determined by having adequate confidence and evidence that *safety requirements* have been met.

- *Safety requirement* is a requirement that, once met, contributes to the *safety* of the product.
- A system is considered *safe* if it is free from unacceptable or intolerable levels of *harm*.
- *Harm* or *mishap* is an event or series of events resulting in unintentional death, injury, occupational illness, damage to or loss of equipment or property, or damage to the environment.
- *Risk* is a combination of the likelihood of *harm* and the severity of that *harm*.
- An *accident* is an event, or sequence of events, that causes unintended *harm*; while an *incident* is the occurrence of a *hazard* that might have progressed to an *accident*, but did not.
- *Hazard* is the potential to cause *harm*, often following from some initiating event that may lead to an *accident*.

The definitions described in this section would serve as the foundation for our research when describing the safety assessment for safety-critical system.

2.2.1.2. Processes

Since safety assessment is to manage the risk that arises from hazards, safety assessment processes would be focusing on managing both risk and hazard. While different literature may describe such processes differently, the cycle of safety assessment processes is generally similar. We have referenced the UK MoD Defence Standard 00-56 [63, 64] and listed these key processes in Table 8.

Table 8. Cycle of Key Safety Assessment Processes (extracted from DEF-STD 00-56)

Processes	Definition
Hazard Identification	Identify and list the hazards and accidents associated with a system.
Hazard Analysis	Describe in detail the hazards and accidents associated within a system, and defining accident sequences
Risk Estimation	Systematic use of available information to estimate risk
Risk Evaluation	Systematic determination, on the basis of tolerability criteria, of whether a risk is broadly acceptable, tolerable or unacceptable, and whether it is ALARP or whether any further Risk Reduction is necessary
Risk Reduction	Systematic process for reducing risk
Risk Acceptance	Systematic process by which relevant stakeholders agree that risks should be accepted.

Epistemic uncertainty can emerge from any of the safety assessment processes in Table 8, from identifying the hazards to accepting the assessed risk. As observed in the definitions, none of the processes explicitly consider the presence and impact of epistemic uncertainty.

Hence, our approach can help to create this awareness and focus to manage epistemic uncertainty throughout the system lifecycle.

For a safety-critical system, safety assessment can be repeated multiple times through its developmental lifecycle. We have chosen to base our research on the ARP 4761 standard [7], which is an industrial standard for conducting safety assessment to certify civil aircraft. The standard describes three phases where safety assessment is applied across the system lifecycle: Function Hazard Assessment (FHA), Preliminary System Safety Assessment (PSSA) and System Safety Assessment (SSA). In other words, the safety assessment processes described in Table 8 are repeated whenever FHA, PSSA and SSA are conducted separately at different milestones of the system lifecycle.

Function Hazard Assessment (FHA). According to ARP 4761, the FHA is a safety assessment to "*a systematic, comprehensive examination of functions to identify and classify failure conditions of those functions according to their severity*". The FHA is usually conducted at the start of a system lifecycle, during early design phase. The aim of the safety assessment is to assess the risk of each failure condition (or hazard) and develop the relevant safety requirements in respond to the risk. In ARP 4761, the safety assessment processes in FHA is as follow:

- Identify all the functions associated with the system of interest
- Identify and describe failure conditions associated with these functions
- Determine the effects of the failure conditions
- Classify failure condition effects based in its level of risk
- Assign requirements to the failure conditions
- Identify supporting material required to justify the effect classification
- Identify method to verify compliance with the failure condition requirements

Preliminary System Safety Assessment (PSSA). According to ARP 4761, the PSSA is a "*top down approach to determine how failures can lead to the functional hazards identified by the FHA, and how the FHA requirements can be met*". The PSSA is normally conducted continuously on components throughout the preliminary design phase of the system lifecycle. It considers deriving safety measures to meet the safety requirements. The outputs from PSSA are feedforward as inputs to the SSA during system development. The main processes in PSSA according to ARP 4761 are:

- Complete the list of safety requirements.

- Determine whether the system architecture and design, can reasonably be expected to meet the safety requirements and objectives (the safety assessment processes would be adopted here)
- Derive the safety requirements for the design of lower level items

System Safety Assessment (SSA). According to ARP 4761, the SSA is a "*a systematic, comprehensive evaluation of the implemented system to show that relevant safety requirements are met*". The SSA is conducted to verify that a system that has been developed conforms to the safety requirements generated from the FHA and PSSA. It is usually conducted in the development phase of the system lifecycle, after the conceptual and design phases. The main processes of SSA according to ARP 4761 are:

- Verify that the design requirements established in the FHA are met
- Validate that the classification established for the system effects are justified
- Verify that the safety requirements called out in, or derived from design requirements and objectives are met

In our research, we focus on these three phases (i.e. FHA, PSSA and SSA) to illustrate the many safety assessments that could have been conducted across a system lifecycle.

2.2.1.3. Techniques

There are many techniques currently available to help safety analysts during safety assessment. Some are well established methods applied in industry, while others are part of ongoing research. For example, the ARP 4761 standard mentions many industrial safety assessment techniques such as the Fault Tree Analysis (FTA), Dependence Diagram (DD), Markov Analysis (MA), Failure Modes and Effects Analysis (FMEA), Failure Modes and Effects Summary (FMES), Zonal Safety Analysis (ZSA), Particular Risks Analysis (PRA) and Common Modes Analysis (CMA).

As it is impossible to describe all safety assessment methods, we have chosen to highlight the ones used in our research, namely FTA and FMEA. Both methods are selected because they are the techniques commonly used in FHA, PSSA and SSA. In addition, we also highlight the Systems-Theoretic Process Analysis (STPA) technique that is currently under research. STPA is selected as it is a useful technique to consider multiple viewpoints from component to system level.

Fault Tree Analysis (FTA). FTA is a "*deductive failure analysis which focuses on one particular undesired event and provides a method for determining causes of this event*" [7]. It is used to "*determine the root causes and probability of occurrence of a specific undesired event*" [65]. The analysis starts with an "*undesired top-level hazard event*" and systematically discovers "*next lower level which could cause this event*". This continues until a "*Primary Event*" is uncovered or until safety requirements for the "top-level hazard event" are satisfied. A Primary Event, according to ARP 4761, is defined as "an event which for one reason or another need not been further developed". An example of a fault tree is shown in Figure 7, where the element fails because there is a functional fault and the protective mechanism is inoperative. FTA will be applied on our research in chapter 6

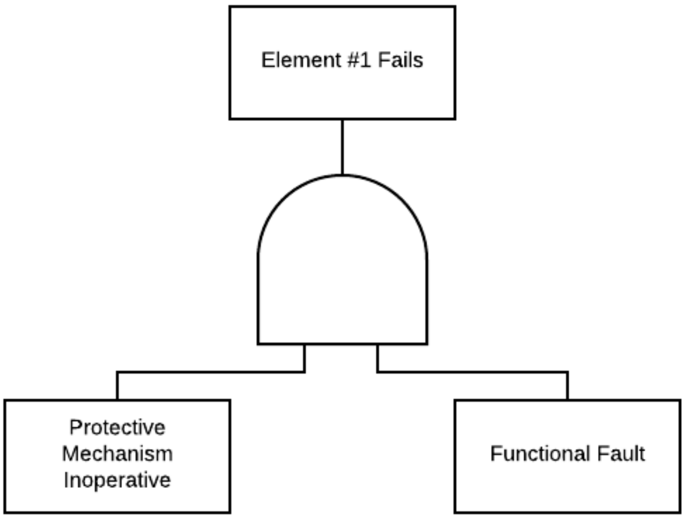
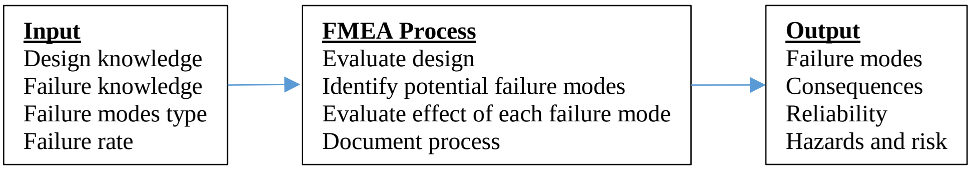


Figure 7. Example of a Fault Tree (extracted from ARP 4761)

The four main steps of FTA to construct a fault tree according to the standard are:

- State the undesired top-level event.
- Develop the upper and intermediate tiers of the fault tree that are minimum, immediate, necessary, and sufficient to cause the top-level event to occur.
- Develop each fault event until the root causes are established or until further development is deemed unnecessary.
- Establish probability of failure budgets or failure rate budgets, evaluate the ability of the system to comply with the safety objectives, and redesign the system if deemed necessary.

Failure Modes and Effects Analysis (FMEA). A FMEA is a "*systematic method of identifying the failure modes of a system, item, function, or piece-part and determining the effects on the next higher level of the design*" [7]. It is a bottom-up evaluation that looks at the impact of every failure mode identified. The results from the FMEA can be amalgamated to form the Failure Modes and Effects Summary (FMES). A FMES is a "*summary of lower level* *failure modes with the same effects from the FMEAs*". According to Ericson [65], the FMEA methodology can be summarised in Figure 8. The input and output from FMEA are as shown after running through the FMEA processes described in the middle. A typical worksheet that records the output from a FMEA is shown in Figure 9. FMEA will be applied on our research in chapter 6.

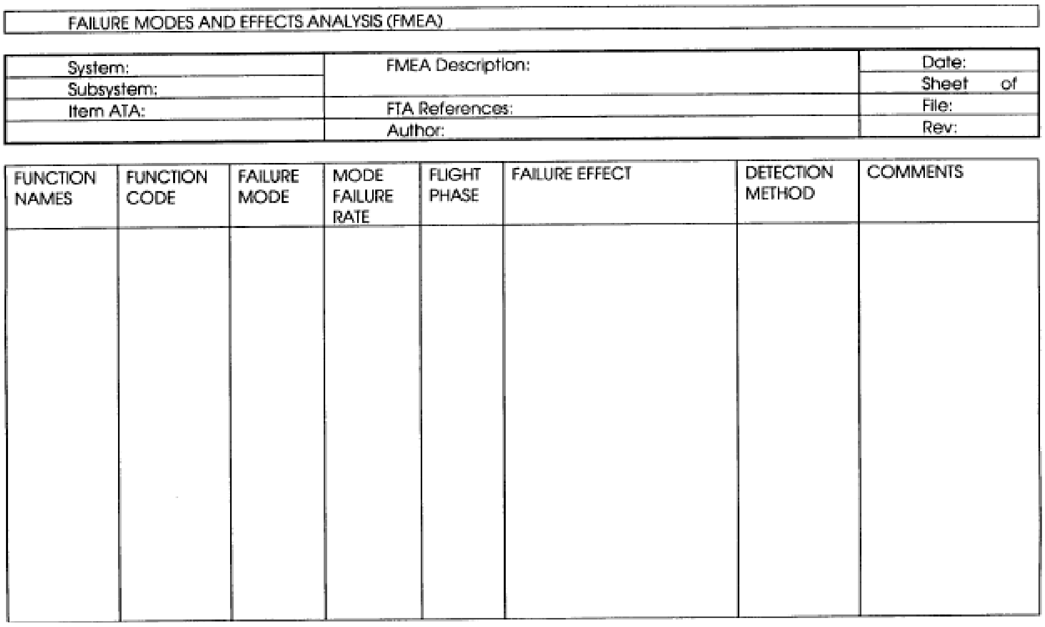


Input
Design knowledge
Failure knowledge
Failure modes type
Failure rate

FMEA Process
Evaluate design
Identify potential failure modes
Evaluate effect of each failure mode
Document process

Output
Failure modes
Consequences
Reliability
Hazards and risk

Figure 8. FMEA Methodology (extracted from Ericson [65])



FAILURE MODES AND EFFECTS ANALYSIS (FMEA)

System:	FMEA Description:	Date:
Subsystem:		Sheet of
Item ATA:	FTA References:	File:
	Author:	Rev:

FUNCTION NAMES	FUNCTION CODE	FAILURE MODE	MODE FAILURE RATE	FLIGHT PHASE	FAILURE EFFECT	DETECTION METHOD	COMMENTS

Figure 9. Example of FMEA Worksheet (extracted from ARP 4761)

Systems-Theoretic Accident Model and Processes (STAMP). STAMP is an accident causality model that is based on the study of systems that are interdependent of each other. It is applied in safety assessment by considering factors such as software, human, technology and organisation. These are factors that directly or indirectly affect the safety of social-technical system. According to Leveson [66], the basic activity in STAMP is to find constraints in the system. In a STAMP model, the cause of accident is considered as "*the result of a lack of constraints imposed on the system design and on operations*".

The hazard analysis technique that is based on STAMP is known as Systems-Theoretic Process Analysis (STPA). STPA uses control loops to identify constraints that, when violated, leads to unsafe situations. There are many applications of the STPA model and one of them is the research on safety-critical UAV system by Chen and Lu [67]. They summarise the STPA process as follows:

- Identify hazards and high-level Safety Constraints
- Develop hierarchical Safety Control Structure to identify Control Actions
- Assess each control action to discover unsafe behaviour as Unsafe Control Actions
- Identify the potential causes leading to the Unsafe Control Actions

The possible faults in a typical control loop, or what is known as Control Flaws (CFs) for STPA, are listed in Figure 10. STPA will be applied on our research in chapter 5.

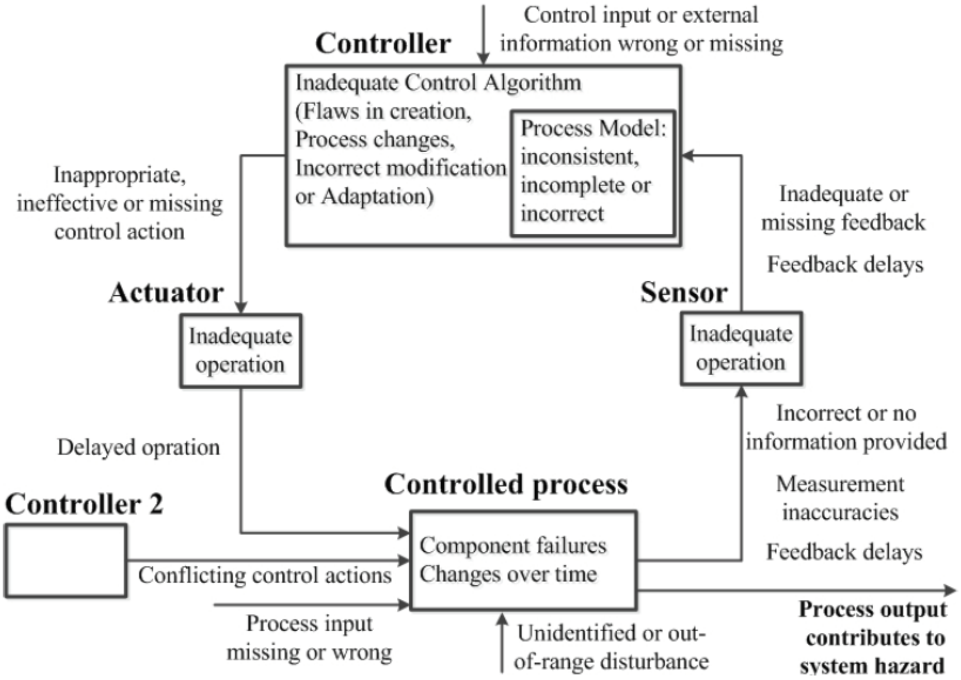

Control input or external information wrong or missing
Controller
Inadequate Control Algorithm (Flaws in creation, Process changes, Incorrect modification or Adaptation)
Process Model: inconsistent, incomplete or incorrect
Inappropriate, ineffective or missing control action
Inadequate or missing feedback
Feedback delays
Actuator
Inadequate operation
Sensor
Inadequate operation
Incorrect or no information provided
Delayed opration
Measurement inaccuracies
Feedback delays
Controlled process
Controller 2
Component failures
Changes over time
Conflicting control actions
Process input missing or wrong
Unidentified or out-of-range disturbance
Process output contributes to system hazard

Figure 10. Typical Control Flaws for a Controlled Loop (extracted from STPA Primer)

STPA provides a useful model to consider multiple viewpoints from component to system level. By using control loops, it allows the user to explicitly annotate causal relationships that can potentially lead to hazardous situations. However, like many safety assessment techniques, STPA assumes that these causal relationships are known relationships and does not provide the flexibility to include epistemic uncertainties (i.e. known unknowns) that may need to be tracked through-life due to lack of information at the point of assessment. While STPA consider a spectrum of factors such as software, human, technology and organisation, the discovery of plausible causal relationships from the unknown unknowns is not systematic as it is still dependent on the experience and expertise of the users during the analysis. The desire to use STPA to consider multiple viewpoints may even exacerbate uncertainty during the safety assessment. Hence, our research into managing epistemic uncertainty can be a plausible approach to complement the STPA when considering multiple viewpoints.

As mentioned at the beginning of the section, since FTA, FMEA and STPA are applied widely in either the industry or safety research, we would also be using them as examples of safety assessment techniques in our research to manage uncertainty. Besides being commonly used, we have chosen these three techniques as they are models that we can apply at component and system level.

2.2.2. **Safety Assurance**

In chapter 1, we have briefly mentioned risk, confidence and uncertainty in safety assessment. Risk evaluation in safety assessment depends on two factors: probability of the expected mishap occurring and the severity of the mishap if it occurs. Usually, the severity of a mishap is well understood, such as personal injuries or infrastructure damages. The probability of occurrence depends on understanding the possible causes and how the causes can lead to the mishap. This is affected by the ability to know the causes, to be aware there could be other unknown causes and to know where to find the unknown causes (which many of these could be impossible to quantify). In this section, the focus is on the measuring and managing of confidence in safety assessment, which is also known as safety assurance.

2.2.2.1. Measuring Confidence

The impact of uncertainty on risk evaluation was discussed by McDermid in his paper on software uncertainty [68]. Ideally, he mentions that "*safety engineering assumes that probabilities reflect aleatoric uncertainty, i.e. "randomness", which can be characterised by a stochastic model"* and that "*we implicitly assume ergodicity – that past failure behaviours are good predictors of the future."* With the assumptions of ergodicity and aleatory uncertainty, the probability of risk can be quantified using mathematics formulae such as probability density functions (PDF).

However, McDermid raises the concern that the *"shape of the PDF or even its mean"* may not be identifiable due to epistemic uncertainty or in his words *"the imperfect knowledge of the system or the stochastic model"*. He presents an interesting formula that was developed by Keynes [69] to explain epistemic uncertainty. Keynes introduced the concept of confidence between two propositions, a and b, which can be expressed as:

$$V(a/b) = Kr/(Kr + Ir)$$

He explains that "*Kr is the relevant knowledge about b, and Ir is the relevant ignorance, i.e. things we would like to know, but currently don't. The bigger V the greater confidence there is that proposition a follows from premise b*". While it looks simple, the formula provides a means to explain the relationship between epistemic uncertainty and confidence. The lower the epistemic uncertainty (i.e. the lower the ignorance Ir), the greater the confidence about the proposition since V increases. Based on this argument, McDermid states that there is a need to invest efforts to *"learn more about the models"* and gain more confidence so as to reduce the epistemic uncertainty.

As a conclusion to his research, McDermid suggests that *"from an ethical perspective, professional safety engineers and the research community need to embrace the difference between aleatoric and epistemic uncertainty".* He advocates that stakeholders involved in developing a system have the duty to ensure uncertainties about the system are explicitly identified and managed.

Observations. McDermid argues that other than quantifying the risk, the confidence in the risk assessment is also an integral part of the safety assessment. In safety, confidence in the risk assessment is often referred to as safety assurance.

2.2.2.2. Managing Confidence

To understand safety assurance, we revisit the concept of confidence that has been discussed in section 1.2. According to Jelen [8], assurance refers to the "*degree of confidence that needs are satisfied*". Applying this to safety, Jelen states that safety assurance focuses on getting sufficient confidence that one is satisfied with the safety assessment, rather than the assessment itself. In other words, assurance serves as a "*measure of confidence in the accuracy of a risk measurement*." Hence, if the safety assessment in the previous section aims to answer the question – "*How safe am I*?", then safety assurance seeks to answer the question – "*How confident am I with the answer to 'how safe am I'*?" Supporting a similar argument, the Federal Aviation Administration (FAA) defines safety assurance as activities under its Safety Management System that "systematically provide confidence that organizational outputs meet or exceed safety requirements" [9].

To gain confidence, one must have sufficient knowledge during safety assessment to make a credible argument to claim that a system is safe. Hence, the lack of such knowledge, or epistemic uncertainty, has direct influence on the confidence in a safety assessment. In their survey on safety assurance [70], Duan et al from the University of Minnesota and Pennsylvania state that it is necessary to deal with uncertainty in order to gain confidence in a safety assessment, which is defined as assurance case. They proceed to categorise the ways that "*researchers have reasoned about uncertainty in assurance cases*" into qualitative and quantitative approaches. According to Duan, the qualitative approach to manage epistemic uncertainty is to "reason it away" and the quantitative approach is to design actions that increase confidence so as to reduce uncertainty.

For qualitative reasoning, Duan brings out the work by Kelly [71] on “argumentation structure for assurance cases” and the use of Goal Structuring Notation to help describe assurance case. In such an approach, epistemic uncertainty is managed by reviewing and restructuring the assurance case to provide reasons for removing the uncertainty. Following similar qualitative approaches introduced by Kelly, Hawkins et al. constructed the “*assured safety argument*" whereby an assurance case can be divided into safety assurance case and confidence case [72]. The confidence case serves as a mean to argue about the confidence in specific portions of the safety assurance case where more justifications are needed.

For quantitative approaches, Duan introduces the research on Baconian Probability by Goodenough et al. [73]. Goodenough et al identify “*sources of doubt*” that could reduce confidence in assurance case and called them defeaters. The approach is to either remove such defeaters or rationalise that they are not affecting confidence on the assurance case. Duan et al assess that using a single number of probability in such quantitative approaches can be “*too coarse of an approach*” since it does not consider “*the subtle nuances in reasoning about uncertainty*”.

In a separate study, Cyra and Gorski [74] caution that the level of uncertainty can be so high that it prevents the ability to make decisions. They state that confidence can be so low until “*it is not a strong reject or accept on the decision scale, but it also has a fairly high level of uncertainty, casting doubt onto any decision that could be made*”. As a result, the decision maker may have to wait until more information is available to resolve the epistemic uncertainty and increase the confidence in the analysis.

Observations. While there are both qualitative and quantitative approaches to manage safety assurance, the assumption seems to be that there are ways to either “reason the uncertainty away” or create immediate solution to reduce the uncertainty (thereby increases the confidence). In complicated safety-critical systems, uncertainty may not be reducible. Even if it can be reduced, it may not be immediate as critical information may not be available. By focusing on defeaters, there is an implicit assumption in the literature that the uncertainties are known unknowns. The potential of unknown unknowns is not considered even though such epistemic uncertainties can be safety-critical and potentially impact the confidence in the safety analysis.

2.2.3. Treatment of Uncertainty in Safety Standards

As a form of assurance, confidence of the safety analyst with a safety assessment can be increased by showing compliance with certain safety standards. In our survey, we have reviewed three safety-related standards that are related to the military and airborne systems safety management: DEF STAN 00-56, MIL-STD-882E and DO-178C. We search for either consideration or management of uncertainty while carrying out safety assessment. We also look for any mentioning of key words such as unknown, change, deviation, identification and tracking.

DEF STAN 00-56. For DEF STAN 00-56 [63], Safety Management Requirements for Defence System, is used by UK MOD to manage safety assessment and assurance. It comprises two portions: requirements (part 1) and guidance (part 2). The standard provides guidance to contractor to comply with safety requirements during system acquisitions. While there are no explicit mentioned of uncertainty, the following are sections that are of interest:

- Section 7.1 – Deviation from General Requirement. It is mentioned that "*for any intended deviations, the tenderer should indicate how their approach will meet the intent of this Standard*". However, uncertainty would often lead to unintended deviation and the management of unintended deviation is not covered in the Standard.
- Section 11.4 – Hazard Tracking. The Standard dictates that contractor needs to implement a hazard log to track hazards. While this could be the mean to track uncertainty, it wasn't explicitly included.
- Section 11.7 – Failure Modes. Although it was mentioned that "*a normal function or previously identified safe failure mode … is used in different context, can lead to emergent hazardous behaviour*", there was no follow up explanation of managing such emergent hazards.
- Section 14 – Supply and Change Management. Under the section on change control, it is stated that "*the contractor shall define in the SMP, a change control system so that the safety impact of any planned or unplanned change can be identified and assessed*". The section proceeds to emphasis on the need to proactively identify, address, plan, monitor and incorporate changes to the system. While such changes could be due to uncertainty, there wasn't any direction on how such uncertainties can be identified in the first place.

MIL-STD-882E. MIL-STD-882E [62], System Safety, is applicable to agencies within the US Department of Defence (DoD). It is used for "*identifying hazards and assessing and mitigating associated risks encountered in the development, test, production, use, and disposal of defense systems*". It is the equivalent of DEF STAN 00-56 used by UK MOD. The following are observations regarding the Standard:

- Paragraph 4 – General Requirement. Under managing lifecycle risk, the Standard states that risk efforts should consider "changes, but not limited to, the interfaces, users, hardware and software, mishap data, mission, and system health data". Specifically, for software safety requirements, the Standard advocates that if there are expected tasks to be performed to meet a certain level of rigour but were unspecified or incomplete, the risk associated with them shall be documented explicitly for tracking. However, there is no requirement to explicitly document these incomplete tasks or what we considered as uncertainties.

- Task 101/103/106/202 – Hazard Management. In task 101, hazard identification is stated but no mention of incomplete knowledge. Although there are attempts to guide the identification of hazards such as (1) the possible "*functional disciplines*" in task 103 to include "*system safety, range safety, fire protection engineering, environmental engineering, explosive and ordnance safety, chemical and biological safety, …*" and (2) the potential contribution to mishap during preliminary hazard analysis in task 202 "*e.g. system components, energy sources, interface and control, COTS, operating environment and health hazards*", the list seems to be disorganised and only served as examples. Hazard tracking is described in task 106, which can serve as the mean to track uncertainty. It is also stated that any "*newly recognised hazards and significant changes in controlling the risk of known hazards*" should be documented in the periodic progress report.
- Task 304 – Request for Deviation. It is stated that "each hardware and software deficiency report to identify potential new hazards or modifications to existing risk levels". Again, while deviation is being considered, it is not monitoring scenarios where hazards are still not established due to uncertainty in the assessment.

DO-178C. DO-178C [75], Software Considerations in Airborne Systems and Equipment Certification, guides the software production for airborne systems and equipment to carry out its function "*with a level of confidence in safety that complies with airworthiness requirements*". In other words, the Standard focuses on the safety assurance of the software production.

- Section 2 – System Aspects Relating to Software Development. This section mentions that system safety assessment process, which comprises the functional hazard assessment, preliminary system safety assessment and system safety assessment, would identify the hazards that translate to system requirements that are related to software. Such requirements would also be under "safety monitoring" to "protect against specific failure conditions" throughout the software lifecycle. While such requirements would influence the software development, there isn't any explicit considerations when the hazards or requirements are not well defined due to uncertainty.

- Section 7 – Software Configuration Management. While there are formal processes to manage software configurations as part of software change review and control, there is no established processes to manage changes to the software requirements due to incomplete knowledge about the system. To elaborate, epistemic uncertainty about a system may result in unspecific software faults. Without a clear definition of the software fault, it is difficult to define the software requirements to mitigate the fault.
- Section 8 – Software Quality Assurance. In the software quality assurance process, audit should be conducted to ensure that any "*deviations from the software plans and standards are detected, recorded, evaluated, tracked and resolved*". It is also emphasised that "*early detection of process deviations assists efficient achievement of software life cycle process objectives*". Like other standards, the deviations are not explicitly referring to uncertainty in the system.

Observations. There is no explicit mention of uncertainty management in the safety standards that we have surveyed. Since there are recommended processes in each standard to track changes and hazards, we can consider enhancing these processes to track uncertainty throughout the system lifecycle. However, there could be potential trade-offs, like having additional time and manpower, to track such uncertainty. It is also important to make the distinction between hazard and uncertainty, as based on the knowledge available, one would be more confident about the risk associated with a hazard vis-à-vis the risk associated with an uncertainty.

2.3. **Modelling of Safety-Critical Systems (SCS)**

As technology advances, more and more systems in domains like defence, air traffic management, railway transport, nuclear power plant, offshore drilling and health care are becoming highly networked and complicated. These systems can be considered as safety-critical systems as "any failure can potentially lead to the loss of life and damage to property or environment" [76]. For example, in the defence industry, networks of large-scale safety-critical systems (often being referred to as Network Centric Warfare or System-of-Systems[1][77]) have been revolutionising the applications of military technology as machines and computers are used to carry out complicated and time-critical tasks. These machines are

[1] The System-of-Systems (SoS) refers to a set of systems that are cooperating for a common purpose while simultaneously working as independent entities.

inter-connected to form larger inter-dependent systems and continuing to expand in numbers and complexity as the armed forces attempt to accomplish more challenging missions.

In this section, we focus on three key concerns regarding safety-critical system that are important to our research. First, we survey the growing challenges due to uncertainty in such systems (see section 2.3.1). Next, we present common models that are constructed to represent the real world where safety-critical systems reside (see section 2.3.2). Besides modelling the system, it is equally important to appreciate ways to analyse safety-critical systems (see section 2.3.3). Lastly, we are interested in the ways that such systems are being managed through-life, from development till operation (see section 2.3.4).

2.3.1. Growing Challenges of Uncertainty in SCS

As we focus on epistemic uncertainty, it is important to appreciate what are the challenges in modern systems that make such uncertainty a growing concern. We focus our discussions on two key characteristics that are commonly associated with systems having epistemic uncertainty, namely adaptive and emergence.

2.3.1.1. Adaptive

Systems with uncertainties do not remain static. They evolve with time. Managing uncertainty will be challenging and unpredictable especially when the system is evolving. According to Trapp and Schneider in their study of "*Open Adaptive System*" [78], changes in adaptive systems could either refer to the environment or the internal states of the system.

Change in Environment. The environment can refer to any factors (e.g. social, natural, political, economic) outside a system that can potentially have influence on it. Even if the system does not change its structure or configuration, the uncontrollable change in the environment can impact the safety of the system (e.g. weather change). Monitoring the environment in safety assessment is crucial as illustrated by Habli [79]. In his research, he monitored changes in the environment as part of the assurance model of safety-critical product-line. Another way to model changes in the external environment is to make use of agent-based simulation.

Another author who acknowledges the impact of environment is Endsley [80]. She introduces the 3-level Situational Awareness (SA) model that comprises 3 processes to manage environmental changes: "(1) *Perceive the elements in the environment within a volume of time and space to assess the level of relevancy, (2) Comprehend their meaning to assess the level of significance, and (3) Project their status in the near future to provide prediction*."

Separately, to monitor the environment, Nwiabu [81] suggests a template to track six relevant items of information (goal, plan, identity, location, distance and time) as the system evolves. He mentions that the lack of awareness of changes in the environment is a key challenge when analysing a system. However, Nwiabu has not elaborated further on how his template can be integrated as part of system analysis.

Change within the system. Another form of adaptation involves the change within a system, such as its structure, pattern or behaviour. Subsystems or components within a system can evolve and change the state of the system. There could also be changes in the way a system is being configured. The following are interesting studies on managing uncertainties due to changes within a system.

- In analysing software risk of adaptive components, Kajtazovic [82] recommends a three-step process using concepts in contract binding, consistency analysis and dynamic deployment. The three steps involve "*define the attributes or components that affect safety, allow changes to an extent (e.g. constraint programming) to which there are no safety impact, allow the changed attribute or components to be loaded only into the final version of the software*".
- The Integrated Modular Avionics (IMA) concept under DO-297 [83] for avionics system is another component-level approach to manage the uncertainty from adaptive components. It focuses on certifying only modular components that have been modified, without the need to re-certify all the modules in the system.

2.3.1.2. Emergence

In chapter 1, we have provided several examples of disasters due to uncertainty. While the faults for each of these disasters are unique, one common observation is that it is extremely difficult to narrow down to a specific failure mode. Unlike simple systems where traditional safety analysis method can use linear reductionist approach to deduce the root causes, the failure modes in complicated safety-critical systems may not be easily identifiable. Reiman [84] observes that effects from such complicated system have "*several parallel contributing factors, instead of one or few causal chains as in linear systems*". Dekker [85] also believes that "*the behaviour of such complex system cannot be reduced to an aggregate of the behaviour of its constituent components*". Hence, even if one root cause has been identified, decision makers may face the frustration of not being able to fully comprehend the entire chains of casual relationships that can lead to a mishap.

Systems with uncertainties can behave differently at system level compare to individual level. This is known as having emergent behaviour at the system level, which does not exhibit itself at the component level. Emergent behaviour may not be predictable due to the presence of epistemic uncertainty. Ghorbani [86] describes emergence as a global behaviour that is "*being more than the sum of the individual components"*. Since emergence cannot be attributed to a single component, reductionist methods like modularisation and reconfiguration may not be effective in analysing such behaviour.

Ulieru [87], in his study on engineering emergence, provides a comparison between traditional and emergence engineering approach (Table 9).

Table 9. Comparing Traditional and Emergence Engineering (extracted from Ulieru)

Traditional Engineering	Emergence Engineering
System has well-defined design, production and functional phases	System has blurred boundaries between different phases with adaptable architectures
System's performance must be specified	Consider performance as an emergent property that cannot be predefined. Besides functional performance, the ability to stabilise the system through feedback as a result of emergent behaviour should be monitored.
Consider system's emergence as an undesirable "threat"	Focus on influencing global emergent behaviour around desired performance
Use top-down deterministic approach to design distributed system	Focus on the interaction among subsystems and the environment towards a collective emergent behaviour with a higher purpose, which cannot be identified in the behaviour of individual parts

Applying emergence to safety, it implies that emergent behaviour can lead to potential safety hazard as such behaviour cannot be detected easily with reductionist analysis. To manage emergent hazards, Redmond [88] attempts to define the different types of emergent hazard in a system as a reference for safety assessment (Table 10). While the list may not be comprehensive, it serves as a basis when considering emergent behaviours. Redmond's categorisation exemplifies the effort to help decision makers pay attention to emergent behaviour during safety assessment so as to reduce surprises.

Table 10. Types of Emergent Hazard (extracted from Redmond)

Types of Hazard	Definition
Reconfiguration	Results from the transfer of a complex system from one state to another
Interface	One subsystem causes a mishap in another subsystem by transferring a failure or partial performance over a defined interface, possibly through another subsystem.
Resource	Results from insufficient shared resources or resource conflicts
Proximity	Caused by the operation, failure or partial performance of another subsystem that is transferred to the victim subsystem by a means other than a defined interface
Interoperability	Command, response or data of one subsystem is interpreted by a second subsystem in a manner that is inconsistent with the intent of the first subsystem

Observations. For a safety-critical system going through the acquisition lifecycle, adaptive and emergent behaviours can create epistemic uncertainties throughout the lifecycle. Being aware of the challenges due to these two behaviours would help in our research to manage the epistemic uncertainties.

2.3.1.3. Illustrations

An example of a complicated safety-critical system is Air Traffic Control (ATC) at an airport. ATC is part of a bigger safety-critical system that comprises the ATC tower, the airport terminals, ground management systems and the numerous aircrafts departing and arriving at the airport. These systems must integrate closely with each other to ensure the safe operation of the airport. The following two incidents at Swanwick ATC illustrate the impact of unforeseen technical failures in a highly complicated airport operation.

- Incident One. On 7 Dec 2013, a technical failure in the voice communication system (VCS) at the Swanwick ATC tower resulted in a major disruption of close to 300 cancelled flights, hundreds of delayed flights and thousands of stranded passengers. National Air Traffic Services (NATS)[2] reported that the VCS failure causes the Swanwick area control not being able to switch from night (5 airspace sector) to day configuration (20~25 sectors) in order to handle the heavier day-time air traffic [89]. The investigation concluded that there were gaps in both the engineering and operation responses when unforeseen failures occurs in interconnected systems operating in complicated environment [90].
- Incident Two. Approximately one year later on 12 Dec 2014, another air traffic disruption at the same Swanwick ATC also resulted in numerous flight cancellations across Heathrow, Gatwick and London City [91]. This time, the System Flight Servers failed when more workstations were being brought online during the transition between normal and standby operation [92]. This affected air traffic control as it was impossible for controllers to access aircraft flight plans. NATS has announced that the failure is unprecedented in its 13 years of operations.

Observations. Both incidents highlight the difficulties in managing system safety in complicated environment. There will be intense pressure when failures occur and rightfully so since safety-critical systems are utilised in situations where there are severe consequences (e.g. danger to passenger and disruption to commercial flights) when the systems fail. It is also possible that certain failures may never manifest themselves during system development as it is impossible to predict and conduct safety assessment on all operational scenarios (e.g. overloading of the ATC System Flight Servers). One way of mitigating such risk could be to create an open and active sharing of uncertainties about safety throughout the system lifecycle from development to operation so as to better anticipate and identify such "*blind-spots*" during safety assessment.

[2] National Air Traffic Services (NATS) is a commercial entity that provides air traffic navigation services to aircraft flying through UK controlled airspace and it operates the ATC tower at Swanwick.

2.3.2. Types of SCS Model

Models are used to represent the structure, pattern and behaviour of a system. While there are many ways to model safety-critical systems, we have classified them into three model types and illustrated each type with examples. A summary of the three types of model is listed in Table 11. The challenges of considering epistemic uncertainties in safety-critical systems will be highlighted in each of the model.

Table 11. Types of Model to Represent Safety-Critical Systems

Model Types	Constructs	Examples
Domains	Different domains that a complex system operate in such as social, technical, cyber.	Socio-Technical System Cyber-Physical System
Part-whole	Duality behaviour of simultaneously represent a system as *part* of a greater system, or as a *whole* with its constituent of subsystems	System-of-Systems Holon model
Architecture Framework	Large-scale and multiple views of complex military system using matrices and diagrams	MoDAF DoDAF

2.3.2.1. Domain-Specific Model

In domain-specific models, a safety-critical system is represented by domains that it is operating in, such as social, political, technical, physical and cyber. We will illustrate with two common domain-specific models, namely socio-technical system and cyber-physical system.

Socio-Technical System. A common safety-critical system that is modelled based on the domain construct is the Socio-Technical System (STS). According to Baxter [93], STS was introduced by Emery and Trist in 1960 to describe a system that involves a "*complex interaction between humans, machines and the environmental aspects of a work system*". This implies that people, machines and context must be considered when developing such system. Recently, there are two variants of STS being introduced to describe large-scale enterprise. These are the Large-Scale Complex IT System (LSCITS) [94] and the Ultra-Large-Scale Systems (ULSS) [95]. The LSCITS considers large-scale independent STS (such as financial systems) that are owned and managed by different organisations but might come together as a "*coalition of systems*". Separately, the ULSS refers to a system with "*interdependent webs of software-intensive systems, people, policies, cultures, and economics*".

Modelling the social aspect of STS is particularly challenging. Consider modelling a through-life acquisition of military system as a STS. Such a system involves defence acquisition where multiple stakeholders collaborate to develop and operate machines that interoperate with one another throughout the system lifecycle. Through our research, authors have raised their challenges in modelling such systems, especially when defining the interactions between the social aspect (i.e. human element) and the technical machines. For example, Sommerville [94] cautions that technical specifications and system requirements that are referenced to build components under a STS can be incomplete and incorrect, as users would interpret and adapt the system unpredictably in practice.

There are also reports that raise concerns of unsafe behaviour because of unknown interactions between stakeholders due to uncertainties in the STS. To illustrate, we highlight three examples raised by various authors of the uncertain situation surrounding STS that can result in unsafe behaviour.

- *Uncertainty due to Diverse Opinions*. Atkinson [96] mentions that, in a multi-stakeholders STS, different parties may perceive and manage a situation differently from their own perspectives often to their own benefits. This may lead to unsafe behaviour which is unexpected. He also says that such uncertainty would be more pronounced for long-term and large-scale system acquisition which has higher chance of having more stakeholders. Since uncertainty is difficult to quantify, Atkinson proposes that it is necessary for "*management flexibility and tolerance of vagueness*" in order to manage such uncertainty. In other words, making decision that is flexible will be more suitable in a highly uncertain system in anticipation of downstream surprises. While such decisions may not be the most optimal, it can still be timely and acceptable to meet safety requirements.
- *Uncertainty due to Automation*. Johnson [97] uses the Uberlingen and Linate incidents to highlight the danger of miscommunication between operators, system engineers and contractors. In his examples, information regarding maintenance activities and system configurations of highly integrated avionics systems was not communicated accurately and it resulted in major failures in air traffic management. Johnson believes that increased automation causes the operators to lose situation awareness as they may not be adequately expose to application processes. In addition, high-level system integration also prevents operators having the opportunity and ability to understand diverse subsystems. Ideally, multiple parties (e.g. operators, engineers and maintenance officers) must work closely together even with increase automation since none of the stakeholders has complete knowledge of the whole system. Unfortunately, this may not be the practice in actual operation due to various reasons like geographical and time constraints.
- *Uncertainty due to Risk Homeostasis*. When modelling a STS, it is widely assumed that human beings would want to reduce safety risk whenever possible. However, Wilde [98] argues that, in any activity, people may "*accept a certain level of subjectively estimated risk to their health, safety, and other things they value, in exchange for the benefits they hope to receive from that activity (e.g. drug use,*

recreation or sports)". This is the theory of Risk Homeostasis. Wilde believes that "*people alter their behaviour in response to the implementation* of *health and safety measures, but the riskiness of the way they behave will not change, unless those measures are capable of motivating people to alter the amount of risk they are willing to incur*." It is important to appreciate the danger of risk homeostasis and be proactive to solicit and manage the risk appetite of different stakeholders.

Cyber Physical System. Another type of domain construct to model safety-critical system is the Cyber Physical System (CPS). CPS can be considered as a system that "*integrate the dynamics of the physical processes with those of the software and communication, providing abstraction and modelling, design, and analysis techniques for the integrated world*" [99]. According to Sampigethaya [100], the physical domain refers to the physical subsystems and environment, while the cyber domain can be a "*potential mix of digital computing, storage, software and data networks". He illustrated this construct using an aviation system (Figure 11). The cyberspace elements interact with the physical world (e.g. infrastructure, hardware, human, processes) through an interface of controller, sensors and actuators.

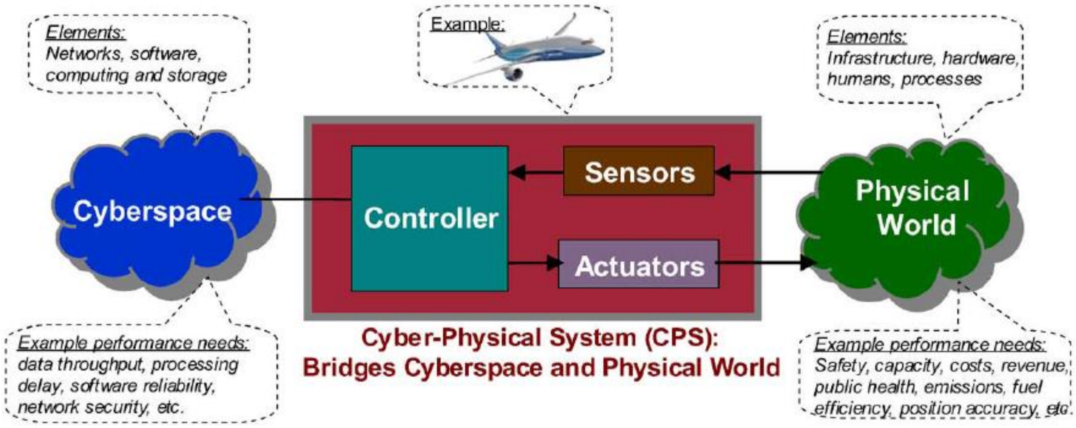

Elements:
Networks, software, computing and storage
Example:
Elements:
Infrastructure, hardware, humans, processes
Cyberspace
Controller
Sensors
Actuators
Physical World
Example performance needs:
data throughput, processing delay, software reliability, network security, etc.
Cyber-Physical System (CPS):
Bridges Cyberspace and Physical World
Example performance needs:
Safety, capacity, costs, revenue, public health, emissions, fuel efficiency, position accuracy, etc.

Figure 11. Example of CPS in Aviation (extracted from Sampigethaya)

Using a similar CPS modelling approach, Banerjee [101] introduces a mapping between the computing subsystems and physical subsystems to extract the interactions between the subsystem interfaces. In his example, the computing subsystems represent software modules while the physical subsystems are sensors that detect the physical environments. Banerjee also introduces two terms to describe the physical interactions of these subsystems: region of impact (intended interaction between subsystems) and region of interest (unintended interactions that may lead to side-effect). He used such a model to analyse the safety of medical devices, such as the Body Area Network, that were used to monitor and communicate vital signs of patients.

Observations. Both STS and CPS modelling have their challenges when it comes to modelling uncertainty. For the STS, we have already highlighted concerns raised by various authors of the uncertain situation surrounding STS that can result in unsafe behaviour. For the CPS modelling, while his approach is systematic, Banerjee's mapping is more relevant to physical properties that can be mapped using mathematical abstractions (e.g. temperature fluctuation and thermodynamic) that link computing units to physical phenomenon. It is difficult to apply in system where the physical dynamic may not be so well-defined using formulae and physic laws. In addition, the physical domain is described qualitatively with incomplete information. Uncertainty about the system is expected even though there are attempts to quantify its characteristics such as temperature and location.

2.3.2.2. Part-Whole Model

In part-whole model, subsystems in a safety-critical system can be considered as "*a component of one or more higher level systems or as whole which has other lower level components as parts*" [102]. We will illustrate with two common part-whole models, namely system-of-systems and holon model.

System-of-Systems. One popular part-whole construct for safety-critical systems is the System-of-Systems (SoS) model which describes a set of subsystems that are cooperating while simultaneously working as independent entities [77]. Such subsystem can in turn represents a system that is formed by a collection of components or sub-systems. According to Maier [103, 104], SoS is described as having the following unique characteristics: operational and managerial independence of individual subsystem, geographic distribution, emergent behaviour and evolutionary development. As SoS is used to describe large-scale system that is highly dispersed, it is understandable that geographical distribution is highlighted as a key property. Besides geographical distribution, the other characteristics are similar to the two challenges highlighted earlier in section 2.3.1: agile and emergent.

Harvey [105] cautions that "*as subsystems of the SoS can be independently operated and managed, there would be a delicate and dynamic balance of responsibilities at individual subsystem level, across boundaries of subsystems and upward at the global level for the entire complex system*". Harvey's concern raises the issue of ownership and responsibility for outcomes like safety for the SoS as subsystems leave and join the SoS dynamically.

Holon Model. Another part-hole construct that is less commonly applied in industry is the holon model. Two of such models are the part-whole state diagram and the phantom system model.

- *Part-whole State Diagram.* Pazzi [102] uses the holon model to partition and simplify subsystems representation. He models two behaviours in separate diagrams:

individual behaviour within each subsystem and global behaviour due to interactions among subsystems.

- *Phantom System Model.* Haimes [106] introduces the phantom system model that models a system with the assumption that there would be certain specific relationships or "shared states" between any two or more subsystems. Examples of such "shared states" include a shared database, the decision of one subsystem that has impact on another subsystem or a collective decision from the collaboration between stakeholders of different subsystem. Haimes recommends passing the knowledge of the specific relationships (e.g. commonalities, interdependencies, interconnectedness) from these "shared-states", as well as states that are not shared between subsystems, to the overall *"meta-model"* to "*explore and learn"* about the overall complicated system.

Observations. Both models do not receive wide spread application in the industry as it is a tedious and abstract approach to derive the models. However, both models do surface uncertainties faced by a subsystem, such as due to the dynamic behaviour of its embedded components (when considering the subsystem as a *whole*) and the external influence when integrating with other subsystems in a complicated environment (when considering the subsystem as a *part*).

2.3.2.3. Architectural Framework

A formal way to model large-scale military safety-critical systems is the use of architectural framework. Specifically, in the military, the US Department of Defence Architecture Framework (DoDAF) and the UK Ministry of Defence Architecture Framework (MoDAF) are preferred models in the military that use matrices and diagrams to capture the different perspectives or views of a system. These views "*represent different stakeholders in the military procurement, planning and implementation of military system*" [107].

Both frameworks are similar except for the inclusion of some customised views in each of them. These frameworks provide a standardised way to organise system architecture into consistent views that aim to complement each other in describing the system. Some of the more common views include the Operation (OV), Systems (SV) and Technical (TV) views. Benade [108] provides an illustration of the MoDAF architectural views as shown in Figure 12.

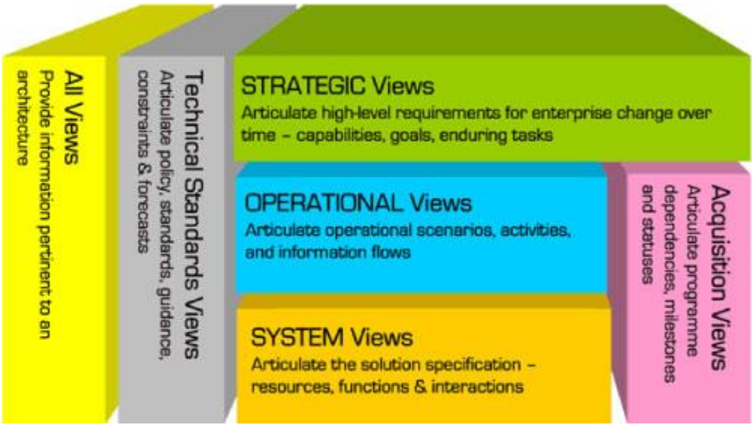

All Views
Provide information pertinent to an architecture
Technical Standards Views
Articulate policy, standards, guidance, constraints & forecasts
STRATEGIC Views
Articulate high-level requirements for enterprise change over time – capabilities, goals, enduring tasks
OPERATIONAL Views
Articulate operational scenarios, activities, and information flows
SYSTEM Views
Articulate the solution specification – resources, functions & interactions
Acquisition Views
Articulate programme dependencies, milestones and statuses

Figure 12. Ministry of Defence Architectural Framework (MoDAF) (extracted from Benade)

Observations. While MoDAF/DoDAF provide a consistent architecture to model a large-scale system, they lack the maintenance and traceability capabilities to track system behaviour throughout its lifecycle. For example, Bartolomei [109] argues that these frameworks capture complementary snapshots of the system but do not focus on the interdependencies and relationships between activities and processes. Similarly, Wrigley [110] states the concern that the DoDAF OV view is only showing a single operational context, without considering that the collection of subsystems only support a capability temporarily and a subsystem may be supporting multiple capabilities requirement simultaneously.

These frameworks also do not consider uncertainties in the system. Such architectural views would require time and effort to construct, which are more feasible during system development. Constructing such views for safety assessment during operation may not be suitable as decisions must be made rapidly in an operating environment.

2.3.3. Analysis of SCS Models

In this section, we focus on two approaches that are being applied to analyse SCS models. We have categorised them into the use of past experiences and modularisation.

2.3.3.1. Use of Past Experiences

The first approach uses past experiences as a reference to predict future behaviour of a safety-critical system. Some common techniques involve the use of cases, patterns and blueprints.

- *Case-Based Reasoning*. Case-Based Reasoning (CBR) is a decision support tool that "*compares the current problem with similar previously experienced concrete problems and solutions*" [111]. Such an approach is used when there is a lack of domain knowledge to provide formal representation of the different scenarios or system configurations. A typical representation of the CBR process is shown in Figure 13. CBR can be used in safety assessment by comparing new case (e.g. a new configuration or new safety case) with those stored in the case repository. If there is a match, the retrieved case from the repository would serve as a reference either to draw lessons from or to help in the assessment of the new case.

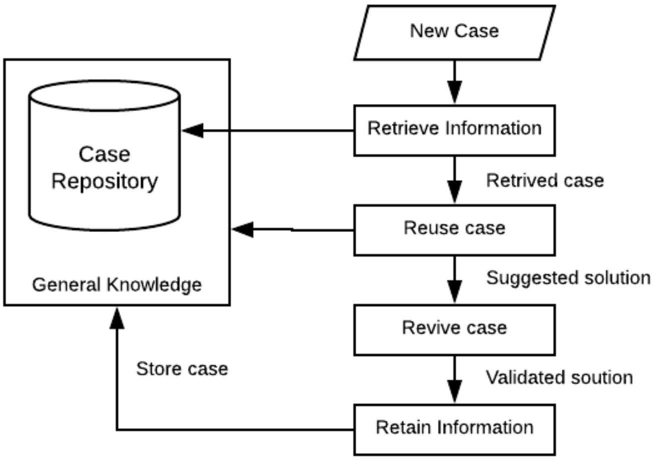


Figure 13. Typical Process Flow for CBR

- *Pattern*. Another technique to capture past experiences is to use patterns. Alexander [112] describes pattern as "*abstracted solutions to recurring design problems in a given context*". Rauhamki [113] introduces several functional patterns in machine and industrial process control that aim to "*document solutions and approach to implementation by capturing explicit and tacit knowledge"*. Other examples include Kazman's [77] integration pattern for new system in a SoS and Wei [114] software safety requirements pattern to be used for safety analysis.

- *Blueprint*. Blueprint is a specific type of pattern that can include information such as hardware modules, software applications and system configuration. For example, Jolliffe [115] implements a system blueprint for IMA described in section 2.3.1.1. It was developed by choosing the "*best bit"* from each hardware, software and configuration blueprints based on a set of mapping and optimisation rules. The selected system blueprint needs to be certified safe, before being "loaded" into the system based on certain integration rules.

Observations. Relying on past experiences is important as it harnesses collective wisdom from other experts and is not limited by the knowledge of an individual doing the assessment. The use of past experiences to support and guide safety analysis is attractive but there is a danger of believing that an outcome from a previous experience that matches the current event will occur again. For example, a used case from the past may have epistemic uncertainty associated with it, that makes it different from the current event and results in different outcome. Such uncertainty may not be documented and the safety analysis may lead to a wrong assessment.

2.3.3.2. Modularisation

A second approach is the use of modularisation to analyse safety-critical system. This approach decomposes a system into independent modules that can represent either a subsystem or a combination of subsystems with a given structure, pattern or behaviour. These modules will be added to, or removed from, the system depending on the changing conditions. We shall illustrate two separate studies that utilise modularisation, namely product-line engineering and modular certification.

- *Product-line engineering*. In a safety-critical system, there could be subsystems or a group of subsystems that are permanent features (e.g. the control tower in the air traffic management system) and there are subsystems that joined the complicated system temporarily (e.g. aircraft arriving and departing the airport). One approach is to define a reference architecture where subsystems are pre-defined as core assets and the relevant assets will only be selected at some specific "*binding time*" later to represent the safety-critical system. Such an approach is known as product-line engineering. The SEI-CMU[3] describes such a software product-line in a similar manner as "*a set of software-intensive systems that share a common, managed set of features satisfying the specific needs of a particular market or mission and that are developed from a common set of core assets in a prescribed way*".

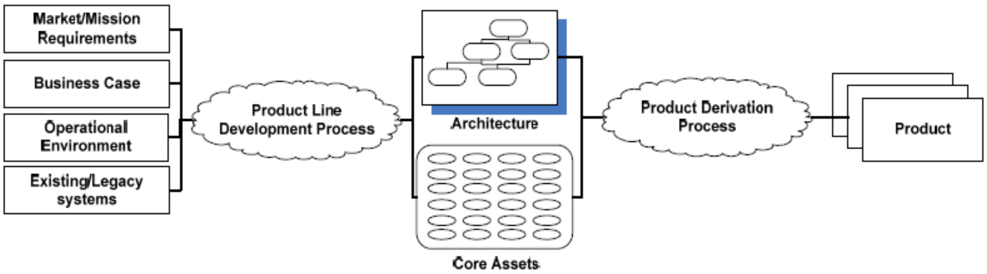


Figure 14. A Product-Line Engineering Approach [79]

Habli [79] introduces a similar approach to assure safety in a safety-critical product line (see Figure 14). He derives reusable core assets from environmental context, system feature and reference architecture. However, this approach has its limitation as the binding of the assets into the "*product*" must occur during the "*product derivation process*", which is within the developmental phase rather than during operation. Furthermore, the configuration or "*permitted variation*" for the product-line must be

[3] Software Engineering Institute, Carnegie Mellon University

pre-defined. It may not be possible to derive such an exhaustive set of configurations for a complicated safety-critical system.

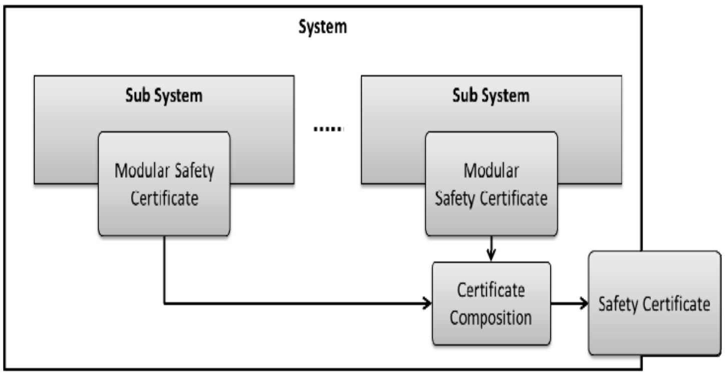


Figure 15. Concept of Modular Certification (extracted from Trapp)

- *Modular Certification*. Modular certification can be applied to safety-critical system whereby subsystems are modelled as independent modules and certified safe at the module level before being aggregated to be certified at the system level. This approach is highlighted by Trapp [78] as a concept to model adaptive system for safety assurance (see Figure 15). The Industrial Avionics Working Group has also proposed a modular and incremental approach for safety certification [116] for aviation components. It assumes that the "*cost of re-certification of change is related to the size and complexity of the system being changed"*. By modularising a system, it is hoped that the scale and complexity of change can be reduced, which makes the re-certification cheaper.

Observations. Modularisation can help to provide flexible solutions to meet changing context of a system where boundaries between subsystems are well defined. The key challenge is when the boundaries between subsystems became blurred and not well defined. For example, in section 2.3.1.1, both Kajtazovic's recommendation and the IMA concept aim to ensure change is incremental and does not necessitate a re-evaluation back to the development phase of the software. The intent is to develop timely and acceptable solution in anticipation of future uncertainties. However, this assumes that components are modular and there are minimal interactions between modular components.

For safety, modularisation may increase the risk of losing critical information or knowledge of the system as a whole. A system can comprise perfectly safe subsystems that may not be designed to operate together. This would potentially make the system unsafe. Such failure is not easily spotted and tracked, as it emerges only when the system is observed as a whole rather than as individual subsystem. Such emergent behaviour has been elaborated in section 2.3.1.2.

2.3.4. Through-life Management of SCS

In this section, we describe the typical lifecycle phases that a safety-critical system would experience, as well as the kind of known challenges in managing such system throughout its lifecycles. Both are important when we construct an approach to manage epistemic uncertainty throughout the system lifecycle.

2.3.4.1. *Phases in Through-Life Management*

A system can be considered as a "*combination, with defined boundaries, of elements that are used together in a defined operating environment to perform a given task or achieve a specific purpose*" [63]. To achieve such a task or purpose, a system often needs to go through multiple phases in time as part of its lifecycle. Consider the military, a system is realised usually by following the system acquisition lifecycle. A full-scale system acquisition lifecycle includes multiple phases, milestones and decision points that shape the development of a system till its operationalisation. Two examples of military acquisition lifecycles are shown in Figure 16. They are the UK MoD CADMID acquisition cycle and the US DoD Defence Acquisition Process.

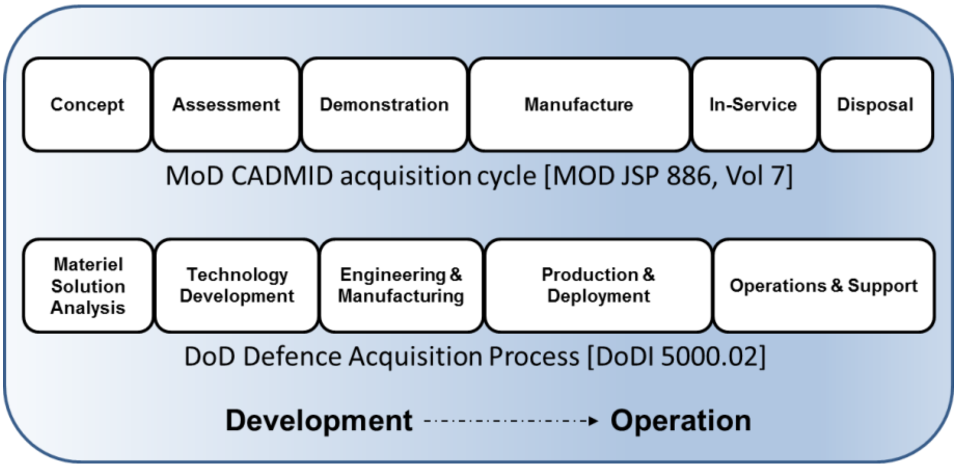

Concept
Assessment
Demonstration
Manufacture
In-Service
Disposal
MoD CADMID acquisition cycle [MOD JSP 886, Vol 7]
Materiel Solution Analysis
Technology Development
Engineering & Manufacturing
Production & Deployment
Operations & Support
DoD Defence Acquisition Process [DoDI 5000.02]
Development
Operation

Figure 16. Categorisation of System Acquisition Lifecycle in Defence

In general, a system moves from the development to the operation phases as it matures throughout its lifecycle. Both phases are subjected to uncertainty while exhibiting different system characteristics. The following table provides a general comparison of the two phases according to MOD JSP 886 and DoDI 5000.02.

Table 12. Comparison of Development and Operation Phases

Characteristics	System Development	System Operation
Typical processes	Design, plan, production, testing, and deployment	Operation, maintenance and support, retirement, phase-out and disposal
Decision-cycle	Long. Decision may be made in days or months after several trials and reviews	Short. Usually demands real-time analysis under time pressure.

Characteristics	System Development	System Operation
Challenges in information management	Potential information overload due to numerous paperwork, activities and milestones. Potential incorrect information due to lack of proper handover of system knowledge	Incomplete or imperfect information to make time-critical decision.
Human Dynamics	Multiple stakeholders involve with different responsibilities. Stakeholders change as system transits to different stages of the life-cycle.	Usually same group of individuals (e.g. operators and maintenance engineers) that operates the system. Mostly operators with field experience, rather than system experts with in-depth technical knowledge
Example of challenges	System design and requirements evolve throughout the system lifecycle Difficult to test interoperability with other systems which could already been in operation. Challenge of scalability due to resource constraints in order to consider all operational scenarios	Trade-offs between performance and risk-adverse options in safety analysis. Manage surprise due to emergent behaviour that will only surface when the systems interoperate dynamically with each other. Verification carried out in static development is not comprehensive enough to be used for real-time dynamic environment.

There are other flexible ways of acquiring capabilities that have been introduced either due to urgent need to meet operational requirements or a lack of budget for full-scale capability development. Such alternate approaches are commonly known as agile acquisition. For example, the MoD acquisition handbook [117] mentions that, besides full-scale acquisition, one can conduct system acquisition in the following ways:

- *Incremental acquisition*. This means that the system is developed in stages, with those components "*providing the most benefits or at an acceptable risk"* being first to be introduced.
- *Evolutionary acquisition*. Instead of developing the system elements in stages, one can also develop the capabilities of a system in phases. It aims to "*provide rapid acquisition of mature technology for the user, while recognising the need to improve future capability*". This approach demands a "*consistent and continuous definition of requirements through active feedback and the exploitation of mature technology*".

For example, the MOD Light Protected Patrol Vehicle program was managed as a form of agile acquisition known as Urgent Operational Requirement (UOR), whereby an "*80% solution"* was targeted to meet the need of rapid acquisition [118, 119]. While there wasn't a clear definition of this "*80% solution"* in the report, it demonstrated the pressure to deploy a capability urgently to meet operational requirement. The report did emphasise that such urgency might indirectly lead to increase safety risk when the project team streamlined some of the system engineering processes as a trade-off for urgent implementation.

Observations. When selecting an acquisition approach, there is a tendency to select an efficient solution by identifying a minimal acceptable level of acquisition activities to deliver the operating system in view of resource limitations. Such acquisition activities usually focus on ensuring system performance first, rather than system safety. Hence, it is common that uncertainty observed during safety assessment would not be tracked throughout the acquisition lifecycle.

2.3.4.2. *Integrating Safety and System Lifecycles*

Within a system acquisition lifecycle, there could be multiple iterations of the safety management lifecycle. Bozanno [120] mentions that there are two complementary roles for safety management in system acquisition, the early "*constructive*" role during system development to guide design activities; and the later "*destructive*" role during system integration and operation to determine if system is indeed safe. Ideally, these "*constructive*" and "*destructive*" roles are carried out using similar safety assessment processes as shown in Table 8 throughout the system acquisition lifecycle.

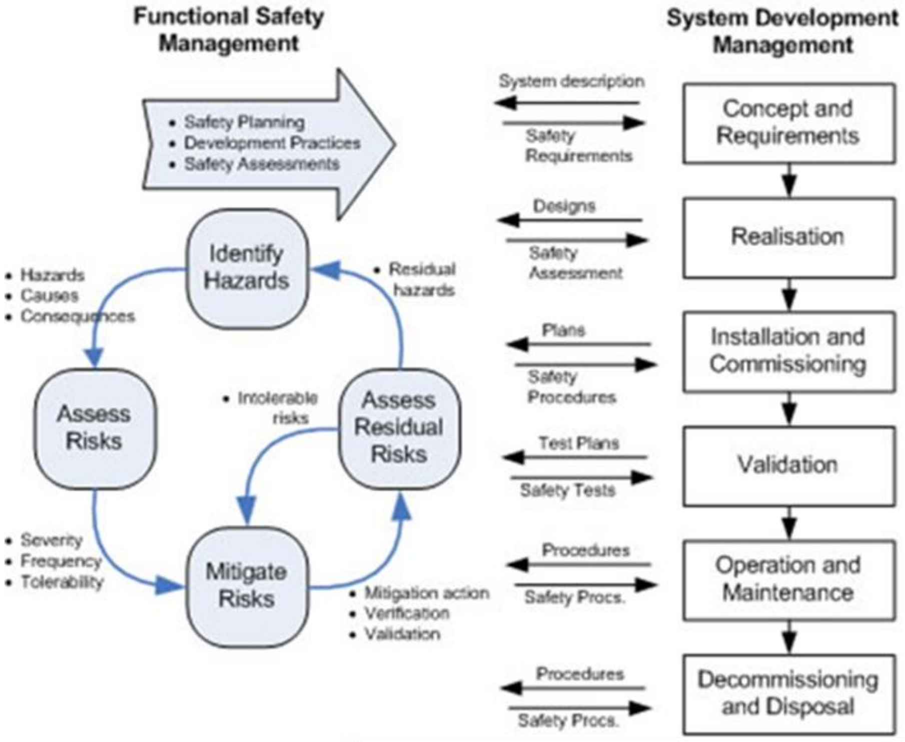


Figure 17. Example of Integrating Safety and System Lifecycles

Based on IEC61508[4], a suitable diagram to summarise the relation between safety assessment and system acquisition lifecycles is extracted from Chambers & Associates[5] [121] (see Figure 17). In the diagram, the system development management is similar to the system acquisition

[4] ISO/IEC 61508: Functional safety of electrical/electronic/programmable electronic safety-related systems

[5] Chambers and Associates (C&A) Pte Ltd is a "*systems engineering consulting, contracting and training organisation providing project services to a wide variety of industry sectors*".

lifecycle. The figure shows the safety assessment processes being called upon multiple times throughout the system acquisition from development to operation.

Separately, Barlow [122] also reiterates the importance of integrating acquisition and safety processes for a successful acquisition. In addition, he mentions the potential "*lock-in*" risk especially in the early stages of the project. As industry partners are usually ahead in the understanding of the system design in the early stages, they may seek to tie down (or "*lock in*") the operators' expectation of the system and limit the system capability early in the design stage. This will make it difficult for operators to modify the system design during system development downstream. Hence, he advocates that safety must be considered early and through-out the system lifecycle. For the complex Next-Generation Air Transportation Systems (NextGen), Baheti [123] also supports the position that it is important to consider safety throughout the system lifecycle, especially during initial design, implementation, maintenance and modification.

Observations. While there is evidence of established safety processes with materials such as safety reports, standards and safety manuals to support system acquisition; there are still challenges in integrating safety and acquisition throughout the entire lifecycle. This is especially challenging for large-scale and long-term acquisition projects which usually involve many stakeholders with different perspectives and experiences. This is made worse with the presence of epistemic uncertainty throughout the system lifecycle due to a lack of knowledge among the stakeholders during safety assessments.

2.3.4.3. Uncertainty in Through-Life Management

In this section, we will illustrate common situations that may result in epistemic uncertainty during through-life management of safety-critical system. One example is the diverse list of possible causes of epistemic uncertainty stated by Atkinson [96] in his research on project management:

- Novelty of design and technology
- Diverse and conflicting stakeholders' expectations
- Failure to anticipate concurrency of activities and capture dependency relationships
- Ineffective communication and knowledge management with changing stakeholders throughout system lifecycle
- Lack of continuity in personal and responsibilities when managing different interoperating systems

- Incomplete and imperfect information
- Lack of systematic process to capture corporate knowledge and lessons learned

Taking healthcare as another example, the need to pay greater attention to uncertainty for safety-critical system was highlighted in the ECRI's[6] report of "*top ten health technology hazards for 2014*" [124]. In the report, the top ten hazards in health care include failures in technical systems (e.g. data integrity failures in IT systems) and neglecting change management for technical network. The key recommendations to counter such failures include raising awareness among different medical stakeholders on how their actions can affect other "*operations, patient care and work processes*" and engendering different stakeholders to work together and mitigate the risk when using complicated networked medical devices.

Separately, in his comparison of large-scale healthcare systems across US and UK, Johnson [125] mentions several challenges that may lead to uncertainties related to the system:

- Tensions between optimisation at global versus local levels,
- Potential hazards due to system failure when staff increasingly rely on complicated technical system, and
- Lack of accurate and shared situational awareness among stakeholders to make informed decision.

Lipsitz [126] sums up the uncertainty in health care as a "*non-linear, dynamic and unpredictable*" system that needs greater attention compared to simple linear system. Such network of health care components includes the hospital, patience, medical devices, rehabilitation centres, etc., that may interact independently and yet produce unintended adverse consequences like drug reactions, infections and death. Like health-care, Lipsitz suggests that other safety-critical systems can first consider explicitly the challenges facing the system and then use it to create the awareness of potential hazards in such circumstances.

In another study to discover the uncertainties from project management, Saunders [46] conducted a survey with project managers in UK nuclear and civil aerospace industry. He generated a list that includes environmental, individual, complexity, information, temporal, capability and regulatory factors. While the sample size was small and the survey was only conducted in the UK, the survey result did reinforce the diverse challenges that are facing project managers. As part of his conclusive remark, Saunders argues that there is a lack of "*orientate and focus*" techniques for decision makers to be better aware of the impact due to

[6] The Emergency Care Research Institute (ECRI) carries out applied scientific research in healthcare to improve patient care in US and Europe.

the uncertainty. The right information was not collected timely to detect potential changes in a project and reduce the element of surprise.

Observations. To summarise, safety-critical system experiences multiple phases and each phase brings with it many epistemic uncertainties that may be cascaded throughout the system lifecycle. Such uncertainties may directly or indirectly impact the safety of the system. As shown in this section, the diverse list of uncertainties makes it difficult or almost impossible for any individual to be aware of all these uncertainties that can affect a safety-critical system. Borrowing Rumsfeld's definitions in section 2.1.1.2, there will be many unknown unknowns that could be discovered which are beyond the knowledge of any individual. There is a need to tap on the collective wisdoms from other experts as a leverage during hazard analysis to identify potential hazards even with the presence of epistemic uncertainty (i.e. known unknowns).

2.4. Summary

In this chapter, we have surveyed the literature from two important concepts that affect our research: uncertainty and safety. While we find out more on each domain separately, we also want to survey literature that deal concurrently with both domains of uncertainty and safety (e.g. the extent that uncertainty is being surfaced and managed in safety assessment). To put into the context that we have surfaced in chapter 1, we scope our survey to focus on safety-critical system, which is also the type of system that we will be dealing with in our research.

Through our survey, we have also decided to focus on the qualitative assessment of uncertainty even though quantitative assessment is equally important. This is because many uncertainties are difficult to quantify, and qualitative assessment is usually required to support quantitative values to provide a more comprehensive assessment (e.g. to provide the context). Our focus on qualitative assessment may potentially be the foundation for any uncertainty management.

With the observations in this chapter, we would be better prepared to develop the principles and models to manage epistemic uncertainties in safety assessment. Moving on, we would focus on modelling such epistemic uncertainties in the next chapter, before we propose the approach to manage these uncertainties.

Chapter 3 – Principles and Models to Manage Epistemic Uncertainties

3.1. Introduction

As mentioned in the introduction and subsequently elaborated in the literature survey under chapter 2, uncertainty can be classified as aleatory or epistemic [10] – aleatory uncertainty is about randomness; epistemic uncertainty is due to a lack of knowledge. Epistemic uncertainty can be due to things we know we do not know (known uncertainties), or things we do not know we do not know (unknown uncertainties) [11]. In practice, stakeholders never completely know if they have full data or knowledge about the system and its environment when modelling them.

Although epistemic uncertainty is unavoidable, we should reduce the undesired effects of known and unknown uncertainties during safety assessment. We have argued in chapter 1 that if epistemic uncertainties were not tracked and addressed appropriately, it can eventually lead to accidents. If these uncertainties were managed better during safety assessment, these accidents might had been averted.

As explained in section 1.3, we will suggest answers to the following three pertinent questions that would affect how epistemic uncertainties can be managed:

Q1. **WHAT** aspects of the epistemic uncertainties do we want to manage?

Q2. **WHERE** do these epistemic uncertainties reside in safety assessment model?

Q3. **HOW** can we better manage the epistemic uncertainties as part of the system lifecycle?

We elaborate on each question here:

- **Focus on 'WHAT': Three Principles to Manage Epistemic Uncertainties**. The first step in studying epistemic uncertainties is to define **what** are important to focus on in managing epistemic uncertainties in safety assessment. We have defined three important principles to develop the eventual approach. The reasons that these principles are important are also presented. These will be elaborated in section 3.2.
- **Focus on 'WHERE': Two Underlying Models to Manage Epistemic Uncertainties**. For system engineering, it is common to use system models such as circuit diagram and system operational envelope to represent a system. It is also common to use safety causal models such as Fault Trees and Swiss Chess model

during risk analysis. However, these models are mostly isolated models that do not directly show the relationship or influence between system and safety-related elements. Hence, we will introduce two underlying models to identify **where** are the critical relationships among system and safety-related elements that epistemic uncertainties can be found and managed. The Conceptual Model of System Safety (CMSS) can be used to represent relationships between system and safety-related elements, whereas the Causal Relationship Model of Safety Assessment (CRMSA) can be used to locate epistemic uncertainties among causal relationships in safety assessment. Both models will be described in section 3.3.

- **Focus on 'HOW': Concrete Approach to Manage Epistemic Uncertainties**. An approach on **how** to manage epistemic uncertainties as part of the ongoing through-life system engineering process is proposed based on the three guiding principles and two underlying models. It is known as the Target-Analysis-Goal (T.A.G.) approach which will be elaborated in chapter 4.

Notations to Represent Relationships. In this chapter, we will frequently refer to the following three common relationships[7] – association, inheritance and aggregation. The notations to represent these relationships are shown in Figure 18.

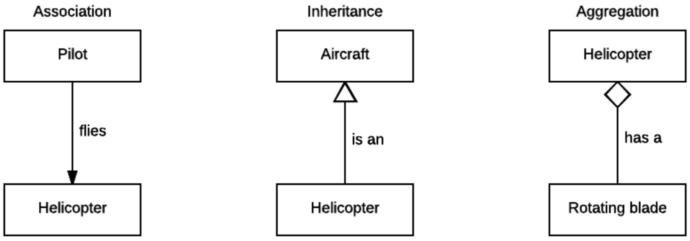


Figure 18. Key Relationships between Elements: Association, Inheritance and Aggregation

Association represents dependency between objects such as a pilot flies the helicopter or a soldier operates the missile. Inheritance refers to a 'is a' relationship, such as a helicopter is an aircraft, a missile is a weapon. Aggregation refers to a 'has a' relationship, such as a helicopter has rotating blades as a part, a missile has warhead as a part.

3.2. **Principles to Manage Epistemic Uncertainties**

In this section, we will describe and justify the importance of the three principles to develop the eventual approach. These three principles are presented below:

[7] These notations are adopted from the UML modelling language that is used to visualise the design of a system.

1. Epistemic uncertainties in the underlying models of safety assessment should be identified (justification in section 3.2.1).
2. Epistemic uncertainties in the underlying models of safety assessment should be documented (justification in section 3.2.2).
3. Epistemic uncertainties in the underlying models of safety assessment should be tracked and addressed in a systematic manner through-life (justification in section 3.2.3).

3.2.1. Principle 1: Epistemic uncertainties should be identified

We have established three reasons why it is important to **identify** the epistemic uncertainties in the underlying models of safety assessment. These reasons are summarised here and will be elaborated in the following section.

- Epistemic uncertainty is inherent in safety assessment
- There is no commonly practised approach of identifying epistemic uncertainty in safety-critical system
- Safety assessment is more complete by recognising and identifying plausible uncertainty than assuming that there is no uncertainty

3.2.1.1. Epistemic uncertainty is inherent in safety assessment

As systems get more complicated, it is increasingly challenging to be aware or anticipate every characteristic and behaviour of the systems as they interact with other systems and the environment.

For safety assessment, it is common to use models to represent a system and its environment as the analysis becomes complicated. System designs, safety assessment techniques, engineering theories and concept of operations are some of the common but diverse types of models that safety assessment depends on. In our research, we consider model to be abstraction of the real world for the purpose of analysis.

The statistician George Box once said, "*All models are wrong but some are useful*." [127]. This may sound misleading depending on how the word "*wrong*" is being interpreted. We believe that what Box means is that a model is "*wrong*" because it cannot completely represent the real world (since epistemic uncertainty is inherent in safety assessment, especially for complicated safety-critical systems[8] [128]); not that it is not suitable to use a model in an

[8] The modelling of safety-critical system is described in chapter 2.3.

analysis. He follows up by explaining that while a model cannot represent the real world completely (i.e. by being "*wrong*"), a model can still be "*useful*" when it is able to serve its purpose. We agree that a model is still relevant as long as the part of the real world that is not represented in the model does not affect the purpose of the model.

Box was using strong language to communicate the idea that a model is an abstraction of the real world. Model abstraction is important as the intent of a model is to represent and simplify the real world for analysis. The importance of model abstraction was also emphasised by Joan Robinson in her famous saying in 1965 [129], "*A model which took account of all the variegation of reality would be of no more use than a map at the scale of one to one*."

One of the key challenges of model abstraction is that there is no single best way to simplify a model as it depends on the purpose of the model. This is analogous to the example whereby there is no single best way to simplify a map away from the scale of one to one, as it depends on the purpose of the map. This is illustrated by Kolmar in his book: Principles of Microeconomics [130]. He illustrated with an example: a map that represents the actual altitude with multiple contour lines may not be useful for a driver; however, these contour lines become very important to a hiker in determining the gradient of his hiking track.

Hence, in the case of a model, the critical question is really what we are trying to accomplish with the model and if the right abstraction has been made so that the model is useful to accomplish its purpose. Since models are necessarily abstract, we must elide detail from our models to make them usable for its purpose. However, every elision will most likely accompanied with either an assumption that the detail we have elided doesn't matter to our purpose or a ram-uncertainty that the elided detail doesn't 'significantly' impact us in achieving our intended purpose (the comparison between assumption and ram-uncertainty is further elaborated in section 3.2.2.1). For safety assessment, the purpose of having models is for causal modelling as explained in section 1.2.3. To conclude, we are faced with the challenge of having epistemic uncertainty (or ram-uncertainty) in safety assessment because the model we use is naturally an abstraction of the real world.

Example. For example, a system flight trial is often conducted using a scaled down model of the operating environment when a system is tested for its air-worthiness. The result is often extrapolated during theoretical analysis to predict the safety of the system in the actual operating environment.

In chapter 2.3.1, we have described that the adaptive and emergent behaviours as key causes of epistemic uncertainties in safety-critical systems. For the modelling of such systems, as a result of both adaptive and a lack of appreciation of the emergent behaviours, we have narrowed down to two issues that make epistemic uncertainties inherent when modelling such system: temporal and model capabilities.

Temporal. During safety assessment, stakeholders may not have the time to wait for all relevant information or data to be available before making a safety decision. When there are epistemic uncertainties, stakeholders have to make decision based on past experiences and collective wisdoms among those who are present.

Example. For example, in the final hours where the management team had to decide if they should go ahead with the launch of the CHALLENGER space shuttle, some of the more experienced engineers were absent and not able to provide technical advice. Nevertheless, the team had to decide even if expert or important information was not available.

Also, in a system lifecycle with multiple milestones, certain information about the system may only be available at the later stage of the lifecycle.

Example. For example, the actual concept of operation may not be finalised during the initial design phase. Hence, stakeholders can expect epistemic uncertainties especially during early design phase where information may only be available later, such as during testing or integration with other systems.

Modelling Capability. For a single system, such as an aircraft platform, safety assessment largely focuses on simple and deterministic events such as component failure within the system. Interactions among subsystems of a social-technical system are more complicated and much less predictable than a simple individual system. Compared to simple systems, complicated safety-critical systems can behave in many more ways or scenarios, some may not even be modelled. Without the capability to build models that can comprehensively represent the complicated safety-critical systems, stakeholders can expect epistemic uncertainties during safety assessment.

Example. For example, when a new helicopter is integrated with other types of air platforms to provide a network of fighting capabilities, not all the possible operational scenarios have been conceptualised. This could be an iterative process whereby the operational concept becomes more mature as the helicopter is being developed. Hence, the requirement models that are used to represent the helicopter and its interactions with other systems to provide the fighting capabilities would have to evolve with time. Naturally, there will be epistemic uncertainties in these models since their capabilities to represent the safety-critical system depend on the maturity in understanding the eventual operational scenarios.

3.2.1.2. There is no commonly practised approach of identifying epistemic uncertainty in safety-critical system

We have mentioned earlier that epistemic uncertainty can be due to things we know we do not know (known uncertainties), or things we do not know we do not know (unknown uncertainties). There are two key shortfalls about the way epistemic uncertainty is being managed now in system safety lifecycle.

First, safety stakeholders tend to ignore unknown uncertainty. During safety assessment, stakeholders must make assessments under a myriad of challenging conditions due to a lack of time, expertise and information. While stakeholders acknowledge the existence of unknown uncertainties, it is reasonable to expect that they would prefer to focus on what they are aware of based on their collective wisdom and experiences regarding uncertainties, that is, the known uncertainties. In other words, given the limited resources, the assessment would tend to focus more on known uncertainties in safety assessment.

Example. For example, in the RSAF safety assessment for new system, what could be uncertain to the safety analysts is the detailed specifications of an existing equipment that is unavailable at the point of assessment. As the equipment has been operating for many years and the specification may not be easily retrievable, there may be tendency to assume the equipment is safe since it has been operating in the existing environment for many years. This uncertainty can be safety-critical if the equipment cannot safely interoperate with the new system.

In summary, there isn't a common and systematic approach to identify unknown uncertainty associated with a system that can be safety critical in current safety assessment.

Secondly, safety stakeholders tend to ignore the evolution of known uncertainty. As for known uncertainty, there is a possibility to ignore and not track it. Such information may be disregarded because it could be deemed not safety-critical at the time it was acquired. However, uncertainty can vary over time as the stakeholders' knowledge about the system and its environment changes throughout the lifecycle. A system that is deemed simple and predictable during design phase may become complicated and uncertain when it starts to interact more with other systems.

Example. For example, in the RSAF safety assessment for new system, a subject matter expert (SME) in air traffic control of an airbase may not be present during a safety assessment. While this may be acknowledged, the absence of the SME is usually not explicitly documented and tracked after the safety assessment session. As a result, there may not be enough attention on the uncertainty of operating the new system in the airbase under various air traffic conditions. This known uncertainty may become safety-critical downstream as the new system starts to operate in the airbase with heavy air-traffic condition (which demands a safety assessment on the safety of flight in such condition).

Also, some uncertainties need time before stakeholders can determine if they are safety-critical. For example, the safety assessment may involve preliminary documents with uncertainties about requirements and design features that can only be validated later as part of system development.

In summary, there isn't a common and systematic approach to identify changes to known uncertainty that may become safety critical in current safety assessment.

3.2.1.3. Safety assessment is more complete by recognising and identifying plausible uncertainty than assuming that there is no uncertainty

Safety assessment is about making predictions. Prediction is best when there is as much knowledge about the system as possible. One way of building up knowledge is to gather as much information related to the system as possible. Acknowledging the presence of uncertainty in the information that is used in the safety assessment is also a form of knowledge. Rather than having false confidence that there is always perfect information during safety assessment, it is better to create the awareness and build in the necessary risk tolerance for the uncertainty.

Overconfidence in a model used during safety assessment can be catastrophic. The worst scenario could be a misplaced sense of confidence that the underlying models in system safety are adequately representing the real world where the system is operating, which may lead to unsafe situation down the acquisition cycle.

Example. For example, in her research on the CHALLENGER space shuttle accident [1] (section 1.1.1), Vaughan highlighted the "overconfident belief in the assumed safety of the shuttle launch through the years". Unfortunately, such overconfidence of the existing model of operating the space shuttle became one of the reasons that desensitised the management from identifying the abnormal O-ring behaviour.

A useful model is one that is explicit about its limitation and associated uncertainty. While the statistician George Box acknowledged that models are useful even if they are "wrong", he cautioned that users need to be aware of the extent where a model can become 'not useful'. The capability of a model to be useful can be limited by the presence of epistemic uncertainty in the model. As stated in chapter 2.2.2, McDermid emphasises that professional safety engineers and the research community involved in developing a system have the duty to ensure uncertainties about the system are explicitly identified and managed. By being honest and explicit about the uncertainties within a model (e.g. what is included/not included in the model, what is known about the model), stakeholders can better apply the model during safety assessment.

3.2.2. Principle 2: Epistemic uncertainties should be documented.

We have established three reasons why it is important to **document** the epistemic uncertainties in the underlying models of safety assessment. These reasons are summarised here and will be elaborated in the following section.

- There is no established approach of documenting the nature of the uncertainties within a model in safety assessment
- Documentation of uncertainties in a model reduces the danger of ignoring safety-critical uncertainty
- Documentation of uncertainties reduces confirmation bias by being open about the reasons to believe and the reasons to doubt an assessment

3.2.2.1. There is no established approach of documenting the nature of the uncertainties within a model in safety assessment

In safety assessment, we are concerned with uncertainty associated with risk assessment, or ra-uncertainty for short. And the specific ra-uncertainty that we are focusing on is the recognised absence of knowledge associated with the models that are used to predict safety risk. We call this risk assessment models uncertainty, or ram-uncertainty for short. Before we proceed to consider uncertainty, it is important to distinguish between uncertainties and assumptions.

An assumption is usually accepted by a safety analyst as long as it is reasonable, even though it may not be certain. Using ISO 15026-2 Systems and Software Engineering—Systems and Software Assurance standard [131] as a reference, assumption is considered as a claim that “appears in an assurance case as evidence” that is provided “without any reason why it is true”. OMG Structured Assurance Case Metamodel (SACM) [132] states that an assumption is considered as a claim that is “assumed to be true” and “is intentionally declared without any supporting evidence or argumentation”.

Assumptions are commonly used in safety assessment when there is a lack of knowledge in certain areas. For example, Weaver [133] mentioned in his research on safety case that assumptions “can be considered legitimate and acceptable where there is a genuine lack of information or lack of understanding that cannot easily be resolved at the time the safety case is presented”. In summary, assumptions are always associated with an absence of knowledge - i.e. for something to classify as an assumption there must be an uncertainty.

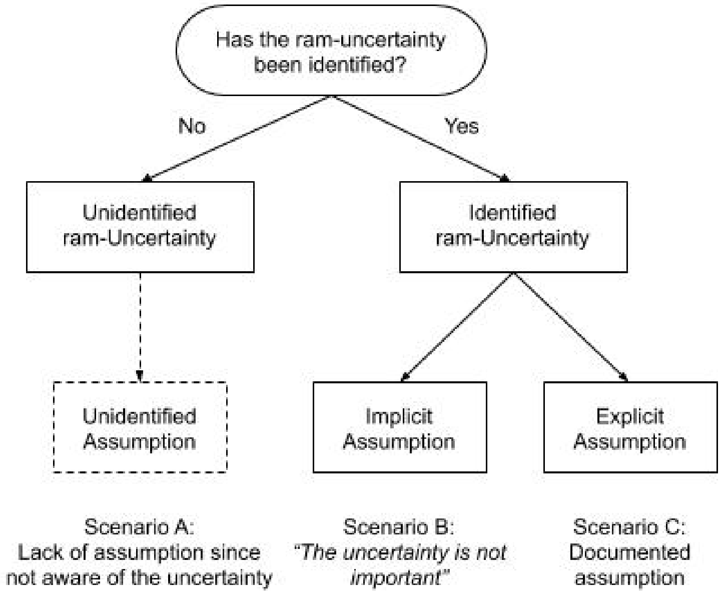

Has the ram-uncertainty been identified?
No
Yes
Unidentified ram-Uncertainty
Identified ram-Uncertainty
Unidentified Assumption
Implicit Assumption
Explicit Assumption
Scenario A: Lack of assumption since not aware of the uncertainty
Scenario B: "The uncertainty is not important"
Scenario C: Documented assumption

Figure 19. Relations between Uncertainty and Assumption

In the context of risk assessment, the presence of a ram-uncertainty does not automatically imply the existence of an explicitly stated assumption (see Figure 19). The figure shows that there are at least three possible scenarios with regards to ram-uncertainty and the associated assumptions. Each of the scenarios has its issues as explained below:

- <u>Scenario A: Unidentified ram-Uncertainty</u>. A perhaps obvious but less considered scenario is when a ram-uncertainty is not being identified. An unidentified ram-uncertainty can mean that one may not even realise that he is operating with potential assumption. As a result, there lies an unidentified assumption. In our research, we aim to develop an approach to discover such uncertainties that the safety analysts are initially unaware of, so as to transform them into identified uncertainties. Once these uncertainties become identified, they would naturally fall under either of the next two scenarios (i.e. uncertainties with either implicit or explicit assumptions).
- <u>Scenario B: Identified ram-Uncertainty dismissed with a questionable assumption of unimportance</u>. Another scenario is when an identified ram-uncertainty does not lead to an explicit assumption. A common explanation is that the ram-uncertainty may be assessed not to pose a substantial safety risk to warrant an explicit assumption

regarding the uncertainty. For example, a stationary radar system that is planned to operate in Singapore may not have been tested in winter (i.e. low operating temperature with potential of snow). This uncertainty will usually be discarded as unimportant since Singapore is a tropical country that does not experience winter. Hence, the performance of the radar in winter needs not be proven since it is currently assessed to not be a credible scenario. However, while an explicit assumption may appear not to be necessary, it does not mean that there is no assumption being made. Minimally, there is at least the default, implicit assumption that "this known uncertainty does not impact safety and need not be tracked as an explicit assumption". In our research, we are concerned with the validity of such an implicit assumption that the identified ram-uncertainty is unimportant, especially when there could be a lack of knowledge at the point of making the assessment.

- <u>Scenario C: Identified ram-Uncertainty managed with a questionable explicit assumption</u>. The most obvious scenario is when explicit assumption is introduced when a ram-uncertainty is identified during risk assessment. Usually, such explicit assumption will be systematically documented and tracked. For example, to manage the uncertainty about whether a new system can perform in all operating environment, an explicit assumption can be introduced to commit that this new system will only be operating in situations that have been tested positively during system trial. No one knows if the new system will be used in other situations in the future, but the explicit assumption helps to commit the current allowable operating environment. In our research, we are also concerned with the validity of such explicit assumptions associated with the identified ram-uncertainty throughout the system lifecycle, especially when there could be a lack of knowledge at the point of making the assessment.

For example, an established way to manage assumption is the Master Data Assumption List (MDAL) in project management. The MDAL is being used by NATO to manage life cycle costing. According to the NATO technical report [134], the MDAL aims to capture data and assumptions that are used to make estimation. While the intent is good, the description in the technical report does not provide details or any structured approach to consider uncertainty across the lifecycle of a system.

If both the assumption and the identified ram-uncertainty are tracked together, it will give the safety analysts the opportunity to reflect on questions such as "*how much doubt do we have about the assumption regarding this uncertainty?*" and "*how unsafe can it be if our assumption about this uncertainty is wrong?*".

Example. For example, back to the example in scenario C. Hypothetically, there could be (1) uncertainty about the duration of which the system will be operated daily, and (2) uncertainty about the manpower requirement to operate the system. If the system is only tracked by the assumption "*the system will only operate in a situation that has been tested in the system trial*", without explicitly stating the two associated uncertainties concerning duration of daily operation and manpower demand; we lose the wider potential of tracking identified uncertainties that may affect the validity of the system trial. When one gathers more information later to clarify these uncertainties, the system trial may need to be revisited if either (1) the daily operation is longer than the trial period or (2) the number of persons needed to operate the system is more than that during the trial.

Lack of detail regarding the cause of uncertainties. Even when uncertainties are being considered during safety assessment, there isn't a model to consider the cause of the uncertainty. Understanding the cause of the uncertainty will help in the subsequent tracking of the uncertainty. In our research, we focus on two possible causes where the epistemic uncertainty can reside: condition of interest (which represent nodes) and causal path (which represent linkages). We will discuss more about these two causes of uncertainty when we introduced one of the underlying models to manage epistemic uncertainties in section 3.3.2.

3.2.2.2. Documentation of uncertainties in a model reduces the danger of ignoring safety-critical uncertainty

Danger of fixing assessed risk in models. Stakeholders may be "lock-in" to a model and ignore safety-critical uncertainty as they cannot adapt in time when the uncertainty poses a safety risk further down the lifecycle. Models are used to identify and represent system structures, behaviours and changes over time. For example, system structural diagram, concept of operation document and hazard analysis techniques are models that are commonly used during safety assessment. Stakeholders would have to conform to the processes defined in these models, as well as accept the assumptions, justifications and epistemic uncertainties that often accompanied the models. As a system evolves throughout the lifecycle, model uncertainties may become safety-critical. However, stakeholders may not be able to adapt in time because either they are not aware of the uncertainties or they choose not to track the uncertainties throughout the lifecycle as they are not aware of the hazard posed by the uncertainties.

Example. Considering the UK Ministry of Defence (MoD) that follows the Concept, Assessment, Demonstration, Manufacture, In-service, and Disposal (CADMID) acquisition life cycle [4] (see also chapter 2.3.4.1), a safety report comprising multiple safety requirements is usually produced at each milestone of the lifecycle (see Figure 20).

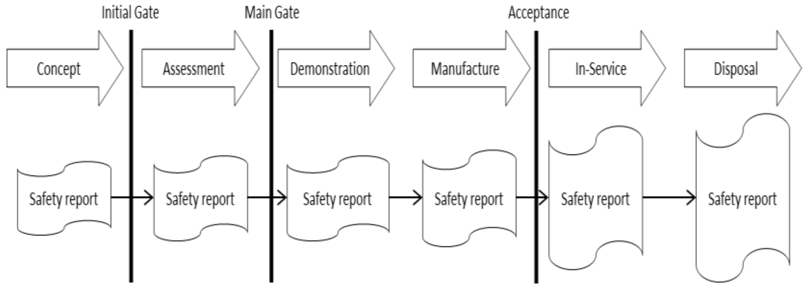


Figure 20. MoD CADMID Cycle

While the safety report at each gate provides important safety requirements, it has an unintended danger of ignoring the epistemic uncertainties in the models whenever decision is being made at each of the acquisition gates. For example, in the research posted by Barlow [122], he warns that '*when a MoD project goes through Initial or Main Gate, the IP (industrial partner) is nominally ahead of the MoD's needs in terms of their understanding of the design and will consequently seek to tie down the customer's expectation of what is achievable. This disconnect can introduce 'locked in' risk if not fully understood*.' According to Barlow, this 'locked in' refers to the finalised design requirements that are to be manufactured and the risk is about the 'disconnect' between the evolving operating requirements and these fixed design requirements that are going to be implemented. Such operating requirements may evolve multiple times throughout the system lifecycle, such as when either new tactics are discovered during demonstration or new system is added into the operation. In his research paper, Barlow provides an example of such a disconnect between an indicative engineering design lifecycle and the MoD CADMID cycle as shown in Figure 21.

Take for example at the Main Gate, the system to subsystem design requirements may have already been finalised for manufacturing. However, as can be seen from the MOD's CADMID lifecycle, the demonstration of the operating concepts has yet to begin. Very often, for military system, operating concepts can change or new ones can be discovered during the demonstration phase. This can change the original operating concepts. In such a situation, if uncertainties about the original operating concepts are not being tracked systematically, they may get ignored when the same operating concepts get modified subsequently, such as during the demonstration phase. This may have safety concern as the system has already been designed according to the original operating concepts.

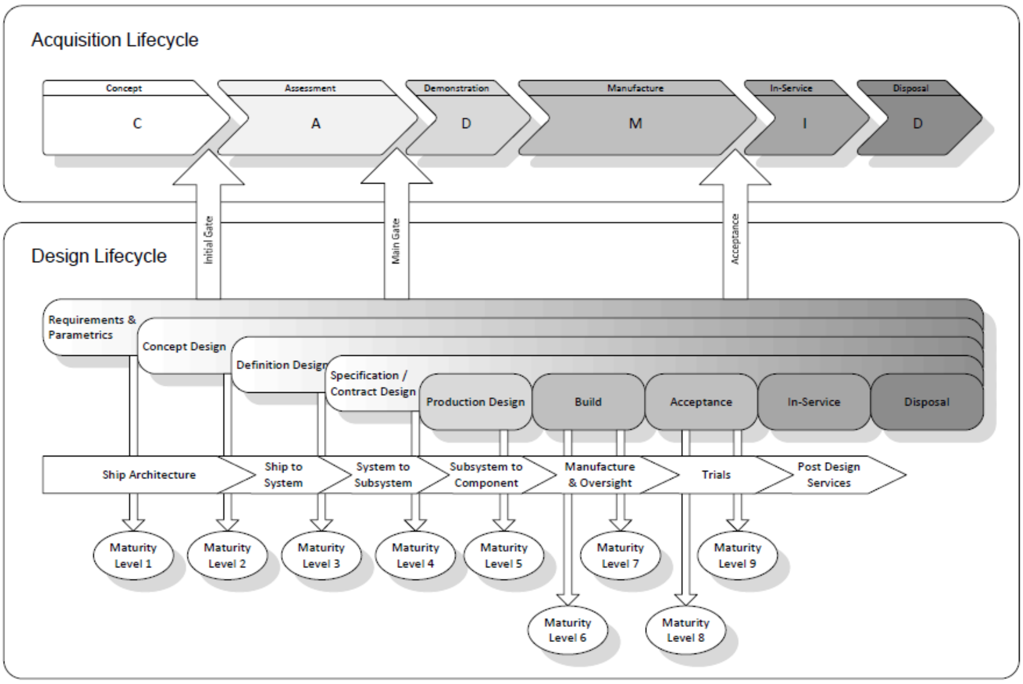


Figure 21. MOD's CADMID vs Design Lifecycle

Such asynchronous processes to develop systems are common in the military and uncertainty can be expected. We can imagine a military example whereby at the Main Gate phase, a system may have been designed for the navy according to a datalink message format recognised by the navy. At this point, there could be uncertainty about the possibility that the system may interoperate with the airforce in the future that may have a different message format. Currently, such uncertainty may not be systematically documented and tracked at the Main Gate as part of uncertainty management. This may become a safety-critical concern down the acquisition lifecycle when the operating concept includes interoperation with airforce system (with its unique message format). Such uncertainty has a high chance of being overlooked as multiple stakeholders (navy and airforce) are involved without an established documentation of the uncertainty. One potential safety concern is that the navy message format may define a target as friendly, while the airforce message format may consider the same target as hostile, leading to possible fratricide.

To reduce the 'lock-in' risk at each milestone, it is important to document as much as practically possible what are the known uncertainties, as well as what are considered or not considered during the safety analysis. Every decision or claim made during the analysis is usually justified by certain arguments. Stakeholders should regularly question if the arguments remain valid throughout the system lifecycle.

Stakeholders need to actively acknowledge such lack of information and manage the corresponding uncertainty accordingly. Creating such awareness of 'lock-in' risk would help to reduce potential false sense of confidence throughout the lifecycle.

By documenting epistemic uncertainty in a model, we create the flexibility not to commit (or “lock-in”) to the details about the uncertainty (such as the context and impact relating to the uncertainty) by explicitly acknowledging that there are some details about the uncertainty that is not available during the instance of analysis. Such information may only be available at a later phase of the system lifecycle. This is like the concept of underspecification in system modelling. For an underspecified model, important variables may be intentionally omitted (see section 2.1.2.3 on underspecification) either to create flexibility or due to uncertainty. For our context, such variables are omitted because the safety analysts do not have enough knowledge at the point of assessment. However, variables that have been omitted should be recorded in the form of uncertainty. In addition, this is useful only when stakeholders create a traceability record that allows a stakeholder to revisit the uncertainties throughout the lifecycle by systematically documenting the uncertainty and the analysis to manage the uncertainty.

3.2.2.3. Documentation of uncertainties reduces confirmation bias by being open about the reasons to believe and the reasons to doubt an assessment

In safety, making a decision in the presence of model uncertainty is a risk-based judgement against the likelihood of finding something either negative or worst-case scenario. However, finding such negative or worst-case scenarios are based on subjective judgement from the stakeholders regarding what is known and what is included; which may not correspond to the worst-case situation. With subjective knowledge and experience, there is a tendency to search for or interpret information in a way that confirms one's preconceptions, which is commonly known as confirmation bias. Whenever stakeholders make decisions in the presence of uncertainties, there is a danger that confirmation bias would influence their decision.

Example. For example, during early design phase where information may not be complete, and developers subconsciously assume certain concept of operations based on legacy systems or past experiences; they may start collecting data to confirm their beliefs and assumptions. Such subjective judgement may cause the stakeholders to miss other safety-critical conditions, which get even more obscured due to uncertainty. Such confirmation bias happened in the CHALLENGER space shuttle accident, whereby the management ended up discarding danger signs (e.g. low external temperature) that could had pointed to safety concern on the O-ring.

Natural inclination in safety assessment is to focus on areas with greater certainties and known uncertainties, which is susceptible to bias. To counter confirmation bias, it is important to consider both the reasons to believe, as well as the reasons to doubt, an assessment. This is analogous to having a 'red team' that challenges the claims, norms, assumptions and decisions made during a safety assessment. Being explicit in documenting epistemic uncertainties is one of such 'red team' approaches to challenge the findings from current safety assessment techniques. By documenting the areas where there is a lack of information, stakeholders can provide the safety assessment with a more holistic and realistic analysis to discover more hazards that could plausibly be safety-critical.

3.2.3. Principle 3: Epistemic uncertainties should be tracked and addressed in a systematic manner through-life.

We have established three reasons why it is important to **track and address** the epistemic uncertainties in the underlying models of safety assessment in a systematic manner through-life. These reasons are summarised here and will be elaborated in the following section.

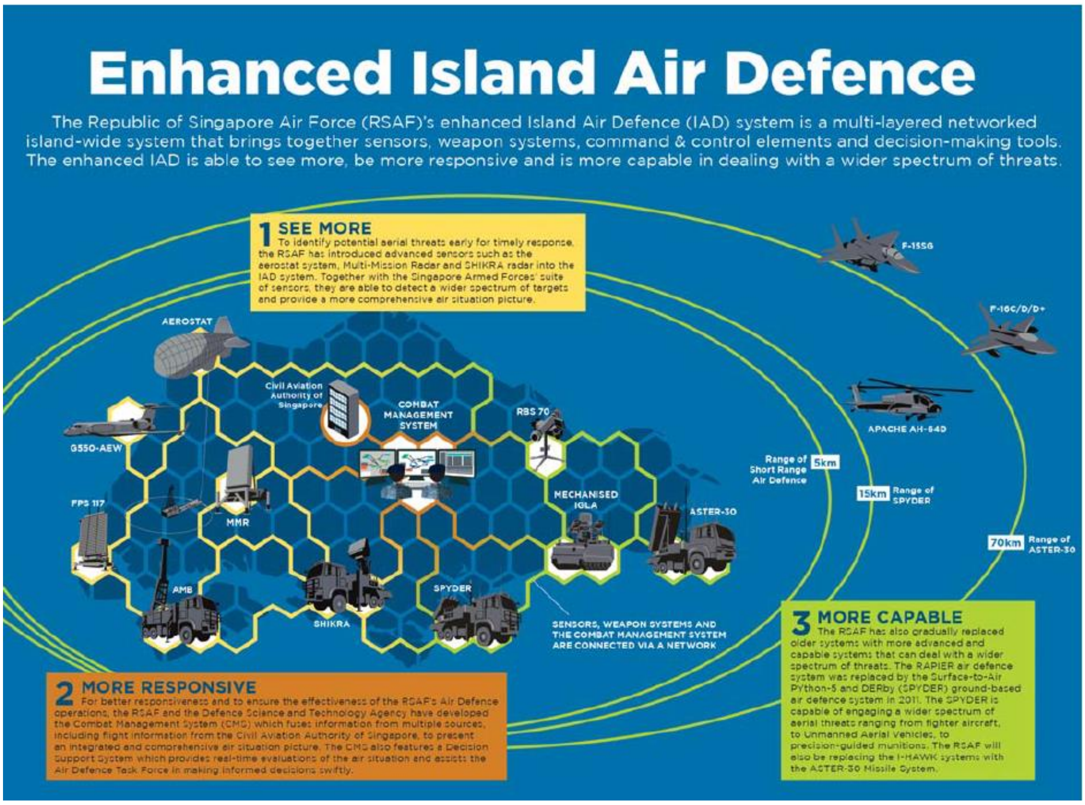
- The nature (and responsibility) of system safety is that it demands a systematic and through-life management of knowledge
- Epistemic uncertainties in system and safety causal models can change with time
- The capability to address uncertainty when it is safety-critical has the potential to increase the confidence in the safety assessment

3.2.3.1. The nature (and responsibility) of system safety is that it demands a systematic and through-life management of knowledge

In the previous section, we have highlighted the importance of documenting the uncertainties as a form of traceability and accountability. However, it is not enough to conduct a one-time documentation. We need to use this information as a leverage to improve the management of knowledge as part of the system safety process. In chapter 2, we have briefly described the through-life system safety process for system acquisition and the responsibility of safety analysts and system engineers to reduce uncertainty as much as possible. System safety for complicated system acquisition is not about making a one-time binary decision about safety (i.e. safe or not safe) but rather a continual stream of safety assessments that identify hazards and formulate mitigating actions to manage the residual risk.

In such safety assessment, safety analysts would have to make safety decisions within the available time and resources (e.g. manpower and available information) before the acquisition process can continue down the system life-cycle. Even without complete knowledge of the system models, especially of complicated systems, safety analysts are expected to still assess if a complicated system is safe to operate.

This is echoed by Sommerville et al in their research on large-scale complex IT systems [94]. Sommerville et al consider such systems as a coalition of systems whereby "*the systems in the coalition may change unpredictably, may be completely replaced and the organizations running these systems may themselves go out of existence. Coalition 'design' involves designing the protocols for communications and each organization using the coalition orchestrate the constituent systems in their own way. However, the designers and managers of each individual system have to consider how to make their systems robust enough to ensure that their organizations are not threatened by failures or any undesirable behaviour elsewhere in the coalition.*" In our context, complicated systems can be considered as a coalition of systems and our safety assessments throughout the systems life-cycle are the continual attempts to sieve out failures or undesirable behaviour in the coalition.



Enhanced Island Air Defence

The Republic of Singapore Air Force (RSAF)'s enhanced Island Air Defence (IAD) system is a multi-layered networked island-wide system that brings together sensors, weapon systems, command & control elements and decision-making tools. The enhanced IAD is able to see more, be more responsive and is more capable in dealing with a wider spectrum of threats.

1 SEE MORE
To identify potential aerial threats early for timely response, the RSAF has introduced advanced sensors such as the aerostat system, Multi-Mission Radar and SHIKRA radar into the IAD system. Together with the Singapore Armed Forces' suite of sensors, they are able to detect a wider spectrum of targets and provide a more comprehensive air situation picture.

2 MORE RESPONSIVE
For better responsiveness and to ensure the effectiveness of the RSAF's Air Defence operations, the RSAF and the Defence Science and Technology Agency have developed the Combat Management System (CMS) which fuses information from multiple sources, including flight information from the Civil Aviation Authority of Singapore, to present an integrated and comprehensive air situation picture. The CMS also features a Decision Support System which provides real-time evaluations of the air situation and assists the Air Defence Task Force in making informed decisions swiftly.

3 MORE CAPABLE
The RSAF has also gradually replaced older systems with more advanced and capable systems that can deal with a wider spectrum of threats. The RAPIER air defence system was replaced by the Surface-to-Air PYthon-5 and DERby (SPYDER) ground-based air defence system in 2011. The SPYDER is capable of engaging a wider spectrum of aerial threats ranging from fighter aircraft, to Unmanned Aerial Vehicles, to precision-guided munitions. The RSAF will also be replacing the I-HAWK systems with the ASTER-30 Missile System.

Figure 22. The RSAF Island Air Defense System [135]

For example, in the RSAF Island Air Defense system of systems (IAD SoS) [135] there are multiple sensors, weapons, command and control and decision making systems interoperating with each other to provide the enhanced IAD SoS capability (refer to the pictorial view of the IAD SoS in Figure 22). The IAD SoS can be considered as a coalition of systems. Each system in the IAD SoS is developed and delivered independently according to its own system lifecycle. When safety assessment is carried out at a specific moment, different systems would be at different phases of their lifecycle. For example, when a safety assessment was conducted for the IAD SoS in 2017, the RBS 70 system was already a legacy system with a wealth of data and knowledge; the SPYDER weapon system had just been delivered into operation; and the AEROSTAT sensor balloon had yet to be delivered.

For such safety assessment to be effective across multiple systems at different levels of development and maturity, safety analysts need to reduce the lack of knowledge or epistemic uncertainties about the system models, so as to make a confident safety assessment. The decision to investigate an unknown situation should also be made systematically by focusing on areas that can help to increase the confidence in the safety assessment by soliciting more information. With the push towards agile development to shorten software development cycle, it is even more important that epistemic uncertainty is managed more effectively and efficiently so that safety analysts can make decisions more confidently.

3.2.3.2. Epistemic uncertainties in a system and safety causal model can change with time

Uncertainty regarding any given system element can vary over time as the stakeholder's knowledge about the system and its environment changes throughout the system engineering lifecycle. For example, we present here two of the ways that epistemic uncertainties may change.

Uncertainties can intensify or diminish with time. Some uncertainties may become lesser with time, e.g. as testers produce test results, operators clarify the expected operating environment and new component developers lock down their interface specifications. Such relevant information can allow stakeholders to make better judgement about safety risk, albeit at a later phase of the system lifecycle. Alternatively, a system that is deemed simple and predictable during design phase may become more complicated and uncertain when it starts to interact more with other systems. Also, some uncertainties need time before stakeholders can determine if they are safety-critical.

Example. For example, the safety assessment may involve preliminary documents with uncertainties about operational concepts, requirements and design features that can only be validated later as part of system development.

Addressing an uncertainty can lead to new uncertainty. The uncertainty that remains after a safety analysis has been undertaken can be considered as residual uncertainty and it may evolve. When mitigating argument or evidence is presented with regards to an uncertainty, a new residual uncertainty may be created.

Example. For example, imagine that there is uncertainty about the operating profiles of an aircraft. After performing the flight trial, it has been ascertained that it is indeed possible to fly the aircraft under the predefined operating profiles. However, during the safety assessment, it was shared that the profile was flown by an experienced pilot and no one was sure if a junior pilot could carry out the same flight profiles. As a result, this leads to a new uncertainty about the ability of less experienced pilot to fly the stipulated operating profiles. This is a scenario whereby addressing the original uncertainty eventually leads to the discovery of a new uncertainty. This new uncertainty must be tracked and addressed until there is more information regarding the experience level needed to fly the operating profiles.

Tracking the uncertainties will enable stakeholders to provide timely intervention to address potential impacts in a systematic manner. If there is no tracking, one may lose information that eventually becomes safety-critical.

3.2.3.3. The capability to address uncertainty when it is safety-critical has the potential to increase the confidence in the safety assessment

To be confident with the safety assessment, we need to have sufficient knowledge to exploit the system models and safety causal model (as explained in section 1.2.3), besides having the knowledge to apply established safety techniques. For example, in the case of FTA safety technique, the safety causal model that is used will determine the cause and effect relationships in the FTA. If there is epistemic uncertainty about the safety causal model, it will directly affect the confidence in applying the FTA safety technique.

Confidence in a safety assessment can be increased if we explicitly acknowledge and reason about the presence of uncertainty within that assessment, rather than allow the variable level of confidence in different aspects of the assessment (e.g. certain causal influences) to remain implicit. Without making uncertainty explicit, the reader has a significant challenge (relying effectively on their own ability to regenerate the result) in determining the confidence they should place in the assessment.

One way to raise the confidence in the safety assessment result is to be aware of what is certain, what is uncertain and what needs to be done about the uncertainty about the safety causal model. This allows the safety analysts to intervene and address a situation when uncertainty is risky enough to be safety-critical. This provides the additional assurance beyond the conventional safety assessment that focuses on evaluating the residual risk and developing mitigation actions for the identified hazards.

In this thesis, we are looking into methods to encourage thinking ahead of time, observing control and influence dependency and deriving taxonomy of how to consider the dependency. However, it is important to note that having more information about the uncertainties in the underlying models of system safety increases confidence but will not directly make a system safer. Whether the additional information makes the system safe or unsafe depends on the follow-on hazard analysis and the safety actions carried out using the results of the analysis.

3.3. Underlying Models to Manage Epistemic Uncertainties

In this section, we focus on representing the "*epistemic uncertainties in the underlying models of safety assessment*", which has been a common theme in all the three principles of managing epistemic uncertainties. This has been briefly explained in Table 3 of section 1.3 when we introduced our research strategy. We introduce two underlying models that are important but not commonly expressed explicitly in safety assessment (see Figure 23). The CMSS concerns the high-level system domain, while the CRMSA is specifically focusing on the safety domain. We will use these two models to recognise critical relationships in safety assessment where epistemic uncertainties may reside.

- **Conceptual Model of System Safety (CMSS)**. In the system domain, we use the CMSS to represent critical relationships between system and safety-related elements. This is adopted from the IEEE 42010 standard that models architectural description for system elements. We have modified the description to better represent the influence of safety elements and the associated epistemic uncertainties. To do that, the model includes three types of elements: system, safety and uncertainty elements, as shown in Figure 23. The model will be described in greater detail under section 3.3.1.
- **Causal Relationship Model of Safety Assessment (CRMSA)**. In the safety domain, safety assessment is carried out with the aim of identifying causal relationships that could be hazardous. Our CRMSA consists of causal conditions that represent the causal relationships when identifying hazards. A causal condition can exhibit causal relationship with potentially many other conditions. Examples of causal conditions in the safety domain include cause, effect, hazard and accident (see Figure 23). The model will be described in greater details under section 3.3.2.

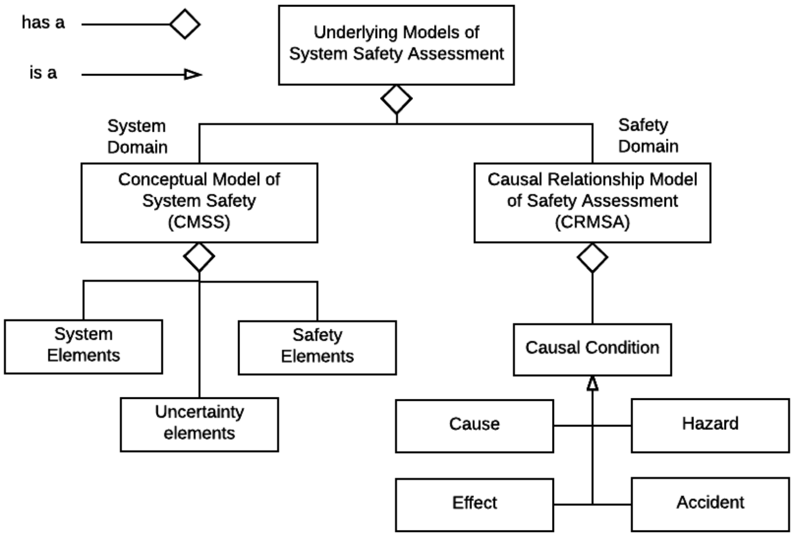

has a
is a
Underlying Models of System Safety Assessment
System Domain
Safety Domain
Conceptual Model of System Safety (CMSS)
Causal Relationship Model of Safety Assessment (CRMSA)
System Elements
Safety Elements
Uncertainty elements
Causal Condition
Cause
Hazard
Effect
Accident

Figure 23. The two Underlying Models of Safety Assessment

3.3.1. Conceptual Model of System Safety

As mentioned earlier, the CMSS is adopted from the IEEE 42010 standard that is used to model architectural description for system elements. For system safety, it is not sufficient to consider system or safety causal models in isolation. We have modified the IEEE 42010 architectural description to better represent safety elements and the associated epistemic uncertainties; and their relationships with system elements. The CMSS is constructed progressively in three steps as follows:

- Considering the relationships among system elements (see section 3.3.1.1).
- Considering the relationships between system and safety elements (see section 3.3.1.2).
- Considering the relationships between system, safety and uncertainty elements (see section 3.3.1.3).

3.3.1.1. Relationships among System Elements

We begin by modelling the relationships between system elements in a typical system acquisition environment. We choose to adopt the IEEE 42010 standard that is used to represent system description. The standard IEEE 42010: "Systems and software engineering – Architecture description" describes "*the manner in which architecture descriptions of systems are organised and expressed*" [20]. It provides a common way to express a system and its relationships with other elements generically, such that it can be applied in different domains. This is useful for our research as the generic model in IEEE 42010 is easily accessible and can serve as a baseline to describe a system. The model can subsequently be expanded to include safety and uncertainty.

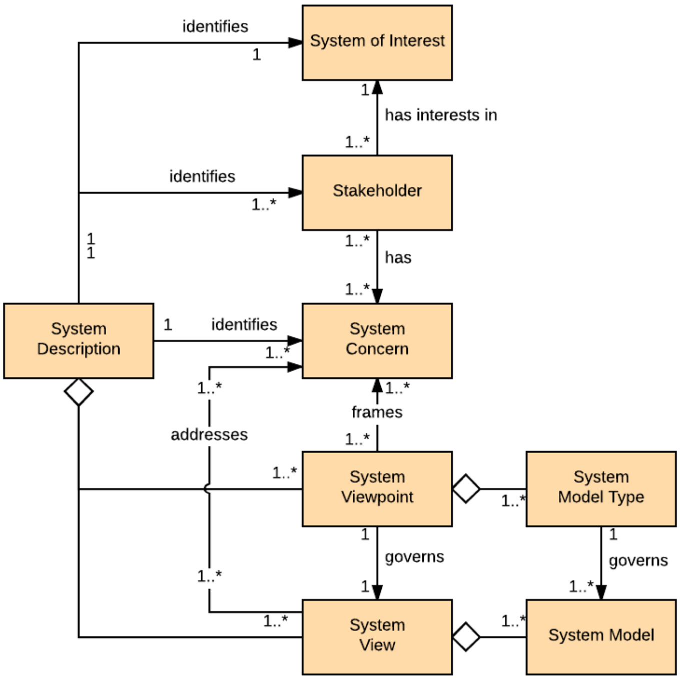


Figure 24. Conceptual Model of System Description according to IEEE 42010

The diagram in Figure 24 shows the system elements and their relationships extracted from IEEE 42010 standard. In the diagram, the system of interest (or ‘system’ for short) is modelled using a system description. The stakeholders and their concerns influence how a system and its environment will be modelled by the system description. The stakeholders are parties interested in the system and their interests are known as system concerns. Per the standard, system concerns can surface throughout the system engineering life cycle “*from system needs and requirements, from design choices and from implementation and operating considerations*”. A system concern could result from “*stakeholder needs, goals, expectations, responsibilities, requirements, design constraints, assumptions, dependencies, quality attributes, architecture decisions, risks or other issues pertaining to the system*”.

For our research, we focus on two key elements in the system description: system view and system viewpoint.

- **System View**. The system view represents how the system behaves in a descriptive manner. Per the standard, a view is “*a representation of one or more structural aspects of an architecture that illustrates how the architecture addresses one or more concerns held by one or more of its stakeholders*.” The system description represents the system using one or more system views. System model is the most common example of a system view used in system engineering.

- **System Viewpoint**. A system view addresses one or more concerns from the stakeholders based on a system viewpoint. A viewpoint is defined in the standard as *"a collection of patterns, templates, and conventions for constructing one type of view*." The system viewpoint exhibits two important associations: it frames the system concerns for the stakeholders and governs the conventions for the system views. In other words, the viewpoint "*establishes the conventions for constructing, interpreting and analysing the view to address concerns framed by that viewpoint.*" Examples of viewpoint conventions include languages, notations, model kind, modelling methods and analysis techniques. For example, a system model (which is an example of a system view) uses conventions that are governed by a system model type (which is a kind of system viewpoint). Such system model type determines the expressive power of the system model.

Example. As an illustration, we consider an example of integrating a military weapon with other complicated command and control (C2) systems. We use the IEEE 42010 template to describe the system elements related to the weapon system (see Figure 25).

As part of system engineering, the weapon would have gone through various milestones in its lifecycle from development to operation. Throughout the lifecycle, multiple stakeholders would have been involved to address various system concerns such as operational performance, availability, security, reliability; as well as safety concern that we are most interested with. System models such as weapon system structure, weapon system specifications and initial concept of operation are used to represent the real world when addressing system concerns. These system models are constructed based on patterns, templates or conventions from system model types such as flow diagram, organisation hierarchy chart and data reporting format.

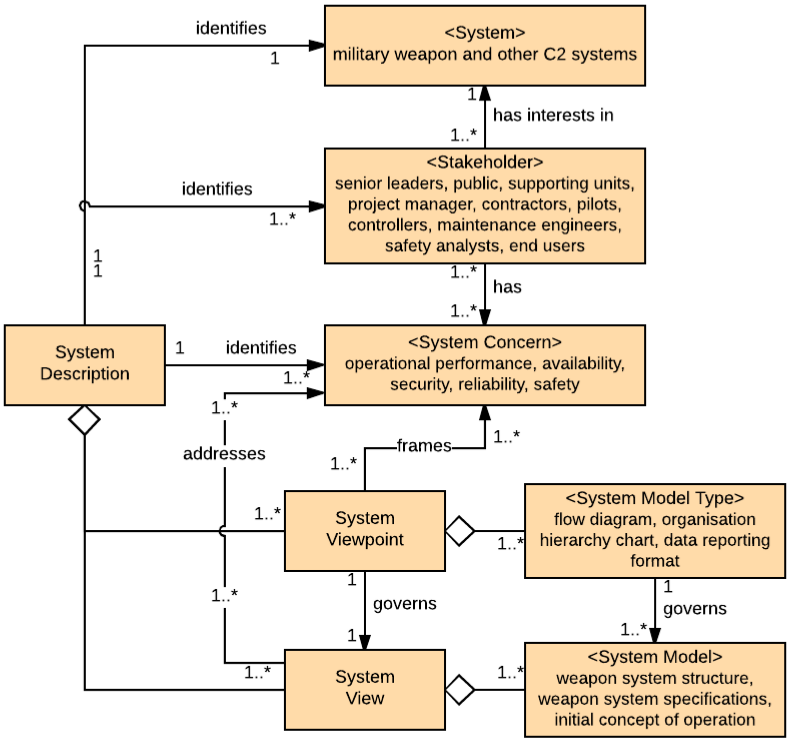

identifies
1
<System>
military weapon and other C2 systems
1
has interests in
1..*
<Stakeholder>
senior leaders, public, supporting units, project manager, contractors, pilots, controllers, maintenance engineers, safety analysts, end users
identifies
1..*
1
1
1..*
has
1..*
System Description
1
identifies
1..*
<System Concern>
operational performance, availability, security, reliability, safety
1..*
addresses
frames
1..*
1..*
1..*
System Viewpoint
<System Model Type>
flow diagram, organisation hierarchy chart, data reporting format
1..*
1
governs
1
1
governs
1..*
1..*
System View
1..*
<System Model>
weapon system structure, weapon system specifications, initial concept of operation
1..*

Figure 25. Example of Conceptual Model of System Description

3.3.1.2. Relationships between System and Safety Elements

While the IEEE 42010 standards provide a conceptual model to represent system elements, it does not explicitly consider safety. Hence, for safety assessment, we need to extend the standard to include elements that are important for the domain of system safety. We need to consider how these safety elements can either influence or be influenced by system elements. To do that, we augment the system elements described in the previous section with their corresponding safety counterparts (see Figure 26). We have also introduced two distinct relationships between the system and safety elements: inheritance and association[9].

Inheritance relationships. The structure, properties and behaviours of safety elements are inherited from their corresponding system elements. This implies that a safety element is an instance of system element that focuses on the issue of safety. Specifically, we have the following five “is a” inheritance relationships:

- Safety system of interest is a system of interest,
- Safety stakeholder is a stakeholder,
- Safety concern is a system concern,
- Safety causal model type is a system model type, and

[9] These relationships have been explained in section 3.1.

- Safety causal model is a system model

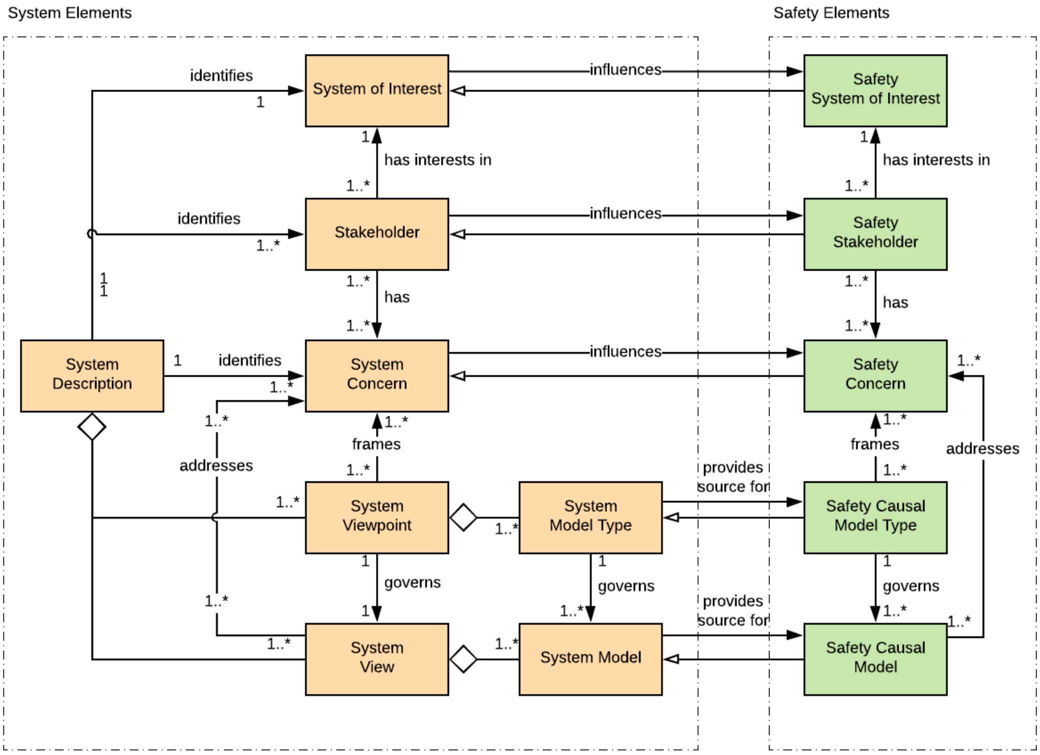


Figure 26. Interactions between System and Safety Elements

Since safety elements are inherited from system elements, we can expect the interactions among the safety elements to mirror that of the system elements. Hence, we can make the following conclusions regarding the safety elements.

- Safety system of interest (or “safety system” for short) refers to the part of the system that is considered during a safety assessment.
- The safety stakeholders are parties interested in the safety system during the safety assessment. They can include contractors, safety analysts, system managers, operators, maintenance engineers and end users.
- The safety stakeholders’ main safety concern is to prevent any harm that may result in death, injury or damage to property under the safety system. This usually involves assessing the risk of potential hazards and coming out with mitigation measures.
- Like system models, safety causal models are constructed to address safety concerns.
- Safety causal models are governed by conventions from specific safety causal model types. Such model types would also frame the safety concerns that can be discovered by safety stakeholders.

Association relationships. We can observe from Figure 26 that safety elements can be influenced by other system elements of the same type as the safety elements. We have identified five of such associations in safety-critical system and provided examples for each of them in Table 13.

Table 13. Association between Safety Elements and System Elements of the same Type

No.	Element Type	Association	Example
1	System	Other system of interest can influence a safety system of interest	Legacy systems using the same resources (such as bandwidth and manpower) as the safety system of interest may have safety impact even if these legacy systems are not directly considered in the safety assessment.
2	Stakeholder	Other stakeholder can influence a safety stakeholder	Safety stakeholders can be influenced by other system stakeholders (such as external contractors, senior management and other company agencies) not directly involved in the safety assessment.
3	Concern	Other system concern can influence a safety concern	Operational concern (especially in the military) can influence the acceptable level of risk in safety.
4	Viewpoint	Other system model type can influence a safety causal model type	The STAMP/STPA safety causal model type is based on system dynamic and control theory, which can be considered as distinct system model types. During initial system development, the presence of system function diagrams (system model type) makes it suitable to use Functional Hazard Analysis (safety causal model type) to identify the preliminary hazards list.
5	View	Other system model can influence a safety causal model	A fault tree analysis (safety causal model) may depend on an Interface Control Diagram (system model) to establish the independency between components. A sneak circuit analysis (safety causal model) may not be possible when the system circuit diagram (system model) is not available.

From the above examples, it can be concluded that safety assessment for safety-critical systems is incomplete if the safety analysts only consider the safety elements. The safety of a system can be affected by other non-safety elements, such as other systems, stakeholders, system concerns, system viewpoints and system views. These elements may not be considered or readily available during safety assessment. Our model helps to create this awareness among safety analysts to identify not only safety elements in safety assessment, but also the system elements that influence these safety elements.

Example. To illustrate, we will continue with our running example of integrating a military weapon with other C2 systems. In this step, we populate Figure 25 with safety elements to show the interactions between system and safety elements (see Figure 27).

As part of system safety engineering, safety assessments are conducted at various milestones across the weapon system lifecycle from development to operation. Multiple safety stakeholders are involved in assessing safety concern (which is also one of the system concerns). These stakeholders include project manager, contractors, pilots, controllers, maintenance engineers, safety analysts and end users. They are also part of the system stakeholders.

One of the common safety concerns is the extent of residual risk identified from safety causal models. This is addressed by using safety causal models to analyse the safety of the weapon. An example of a safety causal model is the causal relationships during Failure Modes and Effect Analysis (FMEA) conducted to identify hazards from operating the weapon. It is based on the FMEA methodology which is a kind of safety causal model type. Examples of hazards include components failure, unsafe operator behaviour, unexpected software interaction, incorrect or insufficient operator practice and undesired change in external environment. Such hazards can be identified from causal relationships between entities, states, behaviours and events that are related to the weapon and its environment. Such causal relationships would often take reference from a set of causal mechanisms that the safety stakeholders are aware of, which would serve as another kind of safety causal model type. We will discuss more about causal relationships in section 3.3.2.

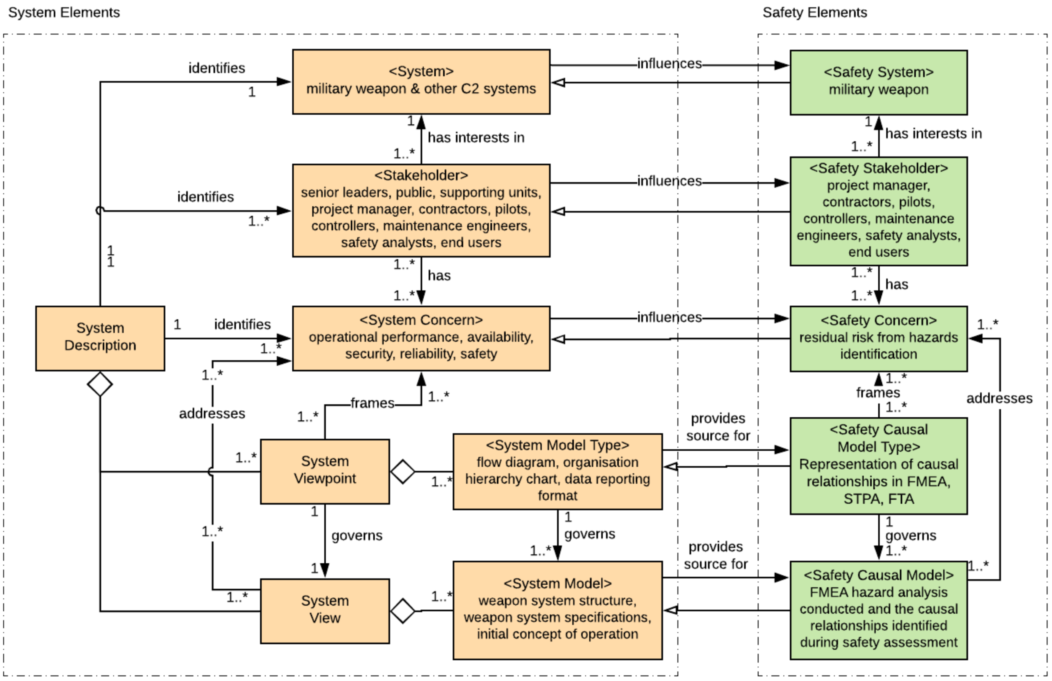

System Elements
Safety Elements
<System>
military weapon & other C2 systems
identifies
1
influences
<Safety System>
military weapon
1
1..*
has interests in
1
1..*
has interests in
<Stakeholder>
senior leaders, public, supporting units, project manager, contractors, pilots, controllers, maintenance engineers, safety analysts, end users
identifies
1..*
influences
<Safety Stakeholder>
project manager, contractors, pilots, controllers, maintenance engineers, safety analysts, end users
1
1
1..*
has
1..*
1..*
has
1..*
System Description
1
identifies
1..*
<System Concern>
operational performance, availability, security, reliability, safety
influences
<Safety Concern>
residual risk from hazards identification
1..*
1..*
addresses
1..*
frames
1..*
frames
1..*
1..*
addresses
provides source for
1..*
System Viewpoint
1..*
<System Model Type>
flow diagram, organisation hierarchy chart, data reporting format
<Safety Causal Model Type>
Representation of causal relationships in FMEA, STPA, FTA
1..*
1
governs
1
1
governs
1..*
1
governs
1..*
provides source for
1..*
System View
1..*
<System Model>
weapon system structure, weapon system specifications, initial concept of operation
<Safety Causal Model>
FMEA hazard analysis conducted and the causal relationships identified during safety assessment
1..*

Figure 27. Example of Interactions between System and Safety Elements

3.3.1.3. *Relationships between System, Safety and Uncertainty Elements*

A safety assessment can cease to be useful when it doesn’t address the safety concerns of the stakeholders. This can happen when the system and safety causal models do not accurately or adequately represent the real world and yet the stakeholders depend on the models to make decision during safety assessment. This can happen when stakeholders are not aware of such model inadequacies due to a lack of knowledge, i.e. the presence of epistemic uncertainty. This is often unavoidable as epistemic uncertainty is inherent in safety assessment.

In our research, we consider that when a model has epistemic uncertainties, it will make the model less adequate to represent the real world for a specific purpose. For example, uncertainty in the inputs to a specific technique (e.g. Fault Tree Analysis) that is based on a certain safety assessment model will make the analysis less capable to assess the actual risk in a system. This would make the FTA less adequate to serve its purpose as a tool to assess the safety of the system. The capability of a model to be adequate is often a subjective judgement [136] that is based on many concerns. In our work, we focus on two concerns related to uncertainty: (1) uncertainty in the model coverage (or type) and (2) uncertainty in the model instance. Uncertainty in model coverage focuses on 'creating the right model' and uncertainty in model instance focuses on 'creating the model right'. We describe both uncertainties in details here.

Uncertainty in the Capability of Model Types. Capability of model types concerns the system viewpoints that bound the knowledge available to create the model. For example, when analysing the operational safety of a system, the concepts of operation will encapsulate the operational viewpoint from where different operational models (e.g. missions) can be created. Two important questions to ask about the capability of the model types during safety assessment are:

> Q1: "Do we know enough of the system and the associated safety concerns which we want to address?"
>
> Q2: "What aspects do we have to consider in the model to address our safety concerns?"

When there is uncertainty in the capability of model types to represent the real world, we call this the "**model type uncertainty**". The model type uncertainty focuses on the uncertainty in the collection of patterns, templates, and conventions for constructing system model. In addition, the model type uncertainty can also exist in the conventions for constructing, interpreting and analysing uncertainty. Hence, model type uncertainty can affect the expressive power of both the system and its corresponding uncertainties.

Uncertainty in the Capability of Model Instances: Capability of model instances, or model view, concerns the appreciation of the real world at the point of conducting the safety assessment, regardless if the model viewpoint is adequate. Having a comprehensive viewpoint does not assure that a model represents all important aspects of the real world at the point of assessment. For this case, it is common for safety analysts to ask the following:

> Q3: "How do we know if it is worth modelling aspect X of the system right now (rather than at some later date where we may know more about it)?"

For safety assessment, such uncertainties can reside either in the system or safety causal model. One of the most common cause of epistemic uncertainties for safety assessment is in the system models. We call this "**system model uncertainty**". This is especially so for complicated safety-critical systems where it is challenging for system models to accurately and adequately represent the real world. Hence, it is common for system models to contain system model uncertainty. Since safety causal model is also a kind of system model, it follows that safety causal model would logically contain uncertainty. We label such uncertainty in safety causal model as "**safety causal model uncertainty**".

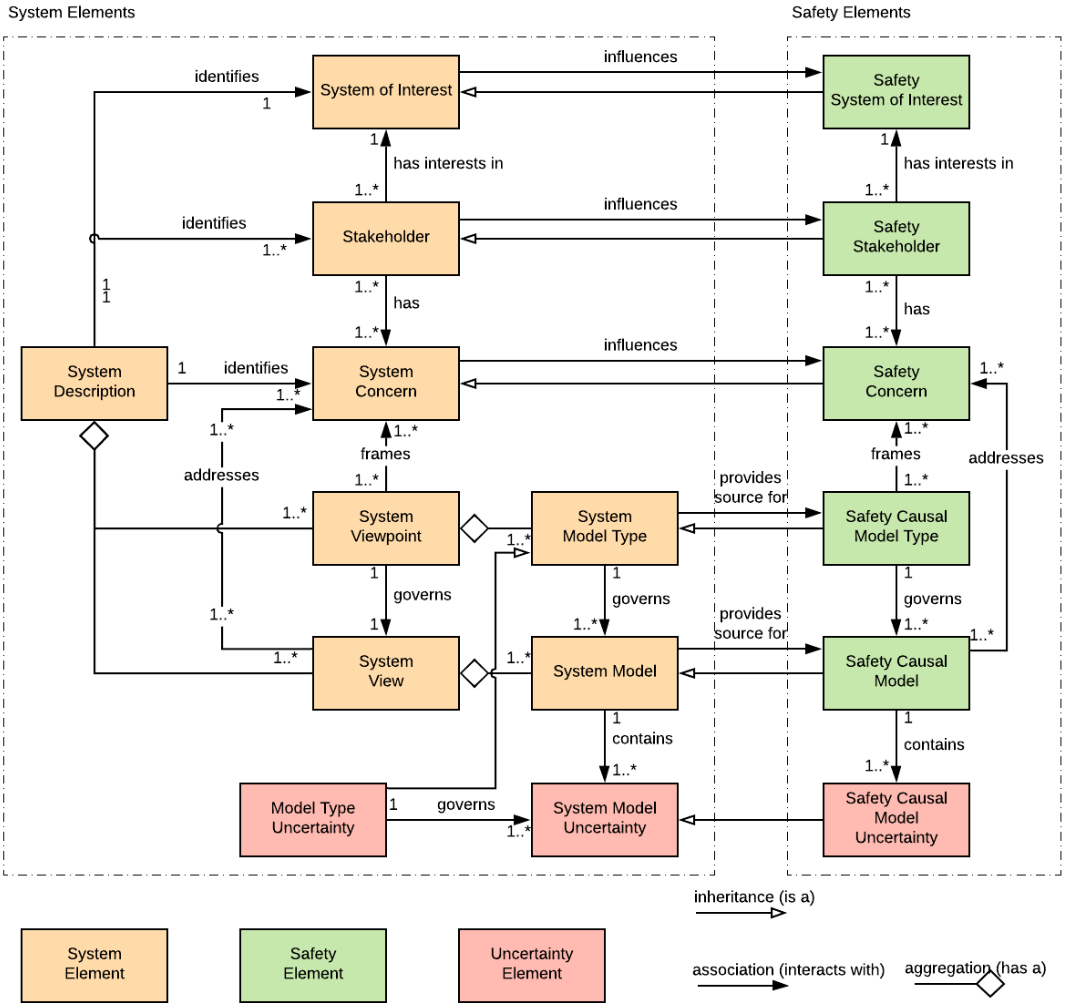


Figure 28. Conceptual Model of System Safety

To summarise, three types of uncertainty elements have been defined, namely model type uncertainty, system model uncertainty and safety causal model uncertainty. We have inserted the three uncertainty elements into the interaction between system and safety elements to form the final CMSS model (see Figure 28). These three uncertainty elements are important as they would guide the safety analysts where to look for uncertainties during systems safety assessment. This would be considered as the cause of uncertainties when we implement our approach in the next chapter.

Example. Back to the running example for the weapon integration, the following are examples of possible uncertainties in the safety assessment that fall into the three types of uncertainty elements (see Figure 29).

Model Type Uncertainty. Uncertain about the types of interface to be considered between systems, asking question like "*are there other interfaces yet to be considered between the weapon and C2 system A besides the known ones, such as communication, physical and electrical interfaces?*".

System Model Uncertainty. Uncertain about the functional linkages between system elements, asking question like "*does weapon have other electrical linkages with C2 system A that have not been captured by the system circuit diagram?*".

Safety Causal Model Uncertainty. Uncertain about the causal relationship with other legacy systems when identifying failure effect using FMEA, asking question like "*will operating the weapon with legacy system B in extreme weather condition lead to an electrical overload hazard?*".

In summary, the CMSS can be used by safety analysts to visualise the relationships between different elements involved in the safety assessment for the weapon integration. Most importantly, it provides a means to describe the possible sources where in the model epistemic uncertainties can reside when system and safety causal models are applied during the safety assessment. It is also useful to appreciate the relationships between these epistemic uncertainties, system elements and safety elements.

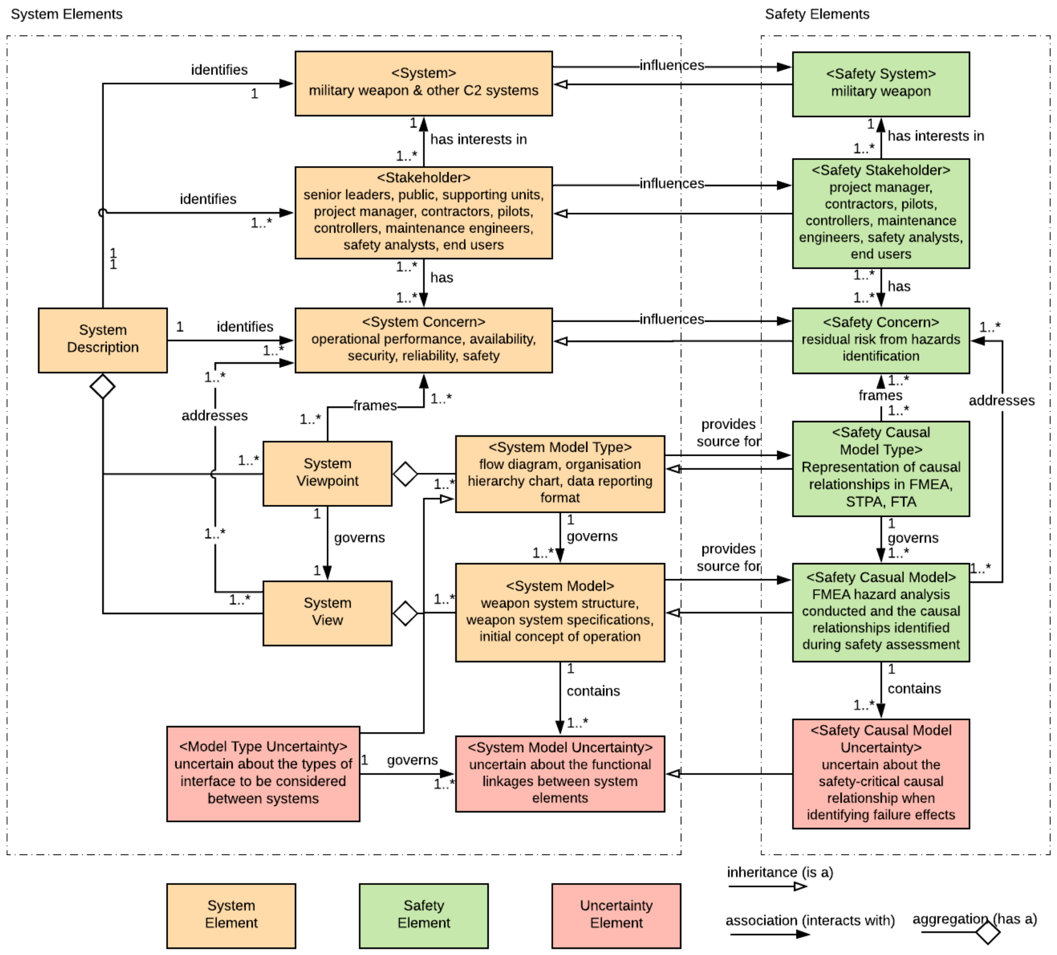

System Elements
Safety Elements
identifies
1
<System>
military weapon & other C2 systems
influences
<Safety System>
military weapon
1
has interests in
1..*
1
1..*
has interests in
identifies
1..*
<Stakeholder>
senior leaders, public, supporting units, project manager, contractors, pilots, controllers, maintenance engineers, safety analysts, end users
influences
<Safety Stakeholder>
project manager, contractors, pilots, controllers, maintenance engineers, safety analysts, end users
1
1
1..*
has
1..*
1..*
has
1..*
System Description
1
identifies
1..*
<System Concern>
operational performance, availability, security, reliability, safety
influences
<Safety Concern>
residual risk from hazards identification
1..*
1..*
addresses
1..*
frames
1..*
1..*
frames
1..*
addresses
provides source for
1..*
System Viewpoint
1..*
<System Model Type>
flow diagram, organisation hierarchy chart, data reporting format
<Safety Causal Model Type>
Representation of causal relationships in FMEA, STPA, FTA
1
governs
1
1
governs
1..*
1..*
governs
1..*
1
provides source for
1..*
System View
1..*
<System Model>
weapon system structure, weapon system specifications, initial concept of operation
<Safety Casual Model>
FMEA hazard analysis conducted and the causal relationships identified during safety assessment
1..*
1
contains
1..*
1
contains
1..*
<Model Type Uncertainty>
uncertain about the types of interface to be considered between systems
1
governs
1..*
<System Model Uncertainty>
uncertain about the functional linkages between system elements
<Safety Causal Model Uncertainty>
uncertain about the safety-critical causal relationship when identifying failure effects
System Element
Safety Element
Uncertainty Element
inheritance (is a)
association (interacts with)
aggregation (has a)

Figure 29. Example of CMSS applied on Weapon System Safety Assessment

3.3.2. Causal Relationship Model of Safety Assessment

In the previous section, we have developed the CMSS to identify where in the model that epistemic uncertainty can affect safety assessment. This is important but not enough to describe the nature of epistemic uncertainties during hazard identification. We need to be more specific in locating where the epistemic uncertainties can reside when identifying hazards.

To do that, we use the 'Condition' concept to develop a model that focuses on the causal relationship in safety assessment. The 'Condition' concept is adopted from the Wilson's safety data model [21] and Habli's product-line functional failure model [79]. Next, we use the Coleman's layered concept of having a boat of causal pathways [22] to tag the safety elements in a causal relationship to the corresponding system elements. Finally, we combine both concepts to form our Causal Relationship Model of Safety Assessment (CRMSA). Both concepts and the CRMSA are described in detail in this section.

3.3.2.1. The "Condition" Concept

The "Condition" concept is based on the safety data model introduced by Wilson. In his model, conditions have causes and consequences. A condition is defined by him as "*an abstraction for capturing some 'state of affairs' – event or state, in the system or its environment".* Hazards, failure conditions and faults are examples of safety conditions. Wilson considers the modelling of causes and consequences as entities rather than as relationships or attributes to cater for various causes and consequences generated by different safety assessment techniques. He also rationalised that it is a more "convenient' and 'clean' approach to consider the causes and consequences as standalone entities. Each condition also has its unique safety properties like likelihood, severity and risk.

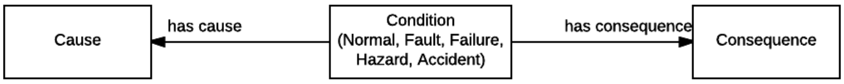


Figure 30. The 'Condition' Concept adopted from Wilson's Safety Data Model

This concept is similar to the Event-Condition-Action (ECA) rule that is used in domains that represent system by events. For example, Almeida et al [137] highlight that such ECA rule can be found in "business workflow [138], manufacturing control [139] and web applications [140]". The rule has three parts: an event, a condition, and an action. It states that when an event is detected, proceed to evaluate the condition; and the action shall be executed if the condition is satisfied. According to Almeida, such an event can be "*database operation, an external incoming event or a timed or untimed temporal event*". In the safety domain, the cause is the event, the causal relationship is the condition and the consequence is the action.

The "Condition" concept was further developed by Habli when he defined his product-line functional failure model (see section 2.3.3.2 for a detailed description of the product-line model). However, unlike Wilson, Habli defines causes and effects as "as types of relationships with which a condition can be associated". He explains that "modelling of causes and effects as relationships rather than as separate elements reduces the complexity involved in representing instances in which a condition within a product-line is both a cause of one condition and an effect of another". We agree with Habli that modelling causes and effects as relationships are more appropriate than Wilson's model, especially when we monitor complicated causal relationships during a safety assessment.

Habli also provided specific examples for safety conditions by referring to the Aerospace Recommended Practice ARP 4754 [141]. The list of safety conditions includes failure condition, failure mode, failure, malfunction, common mode failure, hazard, cascading failure and fault. Each safety condition has causal relationships with other conditions (which can either be the cause or the consequence associated with the safety condition). In addition, a causal relationship occurs only when certain assumptions regarding the environment within which this causality occurs hold, i.e. the causal relationship must be "in context of" a "condition" regarding the environment.

It is important to monitor safety conditions as exemplified in the ARP 4754. However, it is not sufficient to only track safety conditions when we manage epistemic uncertainties in safety assessment. Many of such uncertainties are related to conditions that may not or have yet to transit to some states of affair that are risky enough to be considered as safety conditions. We use a more encompassing term to describe such conditions as 'causal mechanisms'. Causal mechanisms include known safety conditions and uncertain conditions that may be safety-critical but cannot be determined yet due to a lack of knowledge. In addition, even known safety conditions (either causes or consequences) can evolve with the system lifecycle and become uncertain. With these considerations, we have modified Habli's functional failure model described in chapter 2 to include the potential presence of uncertainties in the context of through life safety assessment. In fact, this is the same uncertainty elements that we have defined in section 3.3.1.3.

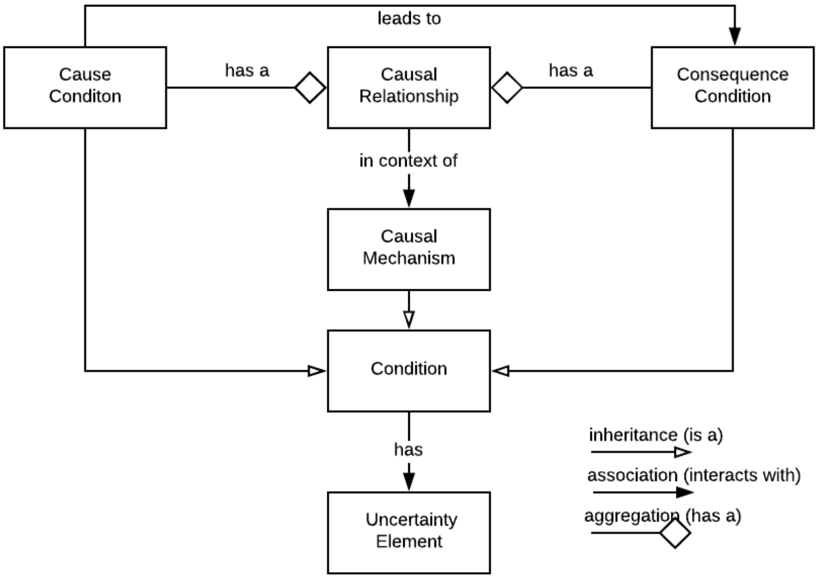


Figure 31. Causal Relationship in Safety Causal Model

Our approach to analysing causal relationships in a safety causal model can be summarised in Figure 31. From the diagram, a 'Causal Relationship' has a 'Cause Condition' (a subtype of a 'Condition') which can lead to one or more 'Consequence Condition' (also a subtype of a 'Condition'). Such causal relationship holds only in the context of certain 'Causal Mechanism', which is also a subtype of a 'Condition'. Due to a lack of knowledge during safety assessment, such ‘Condition’ can consist of ‘Uncertainty element’ that represents the epistemic uncertainties when considering the state of affairs in the system or its environment.

To summarise, we have focused the attention on uncertainty in the causal relationship that is identified during safety assessment. This uncertainty can manifest either in the safety causal model (known as safety causal model uncertainty), system model (known as system model uncertainty) or the system viewpoint that governs the uncertainty (known as model type uncertainty). We have also specified that in such causal relationship, it is the cause condition, consequence condition and the causal mechanism that can potentially be uncertain.

3.3.2.2. *Extending Causal Relationship to System Elements*

As discussed in section 3.3.1, safety assessment involves both safety and system elements. Hence, we need to extend the causal relationship in a safety assessment to represent the system elements that influence or are influenced by the conditions in the safety assessment. To do that, we have adopted the Coleman’s “boat” model of causal pathways [142, 143] to represent such multi-tiered relationships between system and safety elements (see Figure 32).

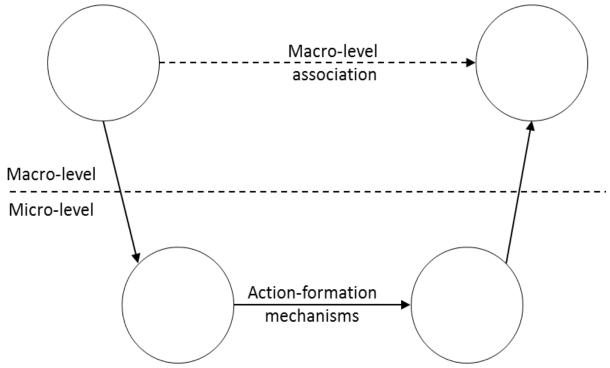


Figure 32. Coleman’s Boat of Causal Pathways

The Coleman model, which is commonly applied in the social and biological domains, comprises two levels of association between elements - the macro and micro level.

- **Macro-level Association**. The macro-level association is a high-level concept that shows the relationship between system elements. For example, in biology, some scientists may work at the macro-ecosystem level (e.g. between human and animals) which can be highly abstract. One such association could be “*human eats animal*”.
- **Micro-Level – Action-Formation Mechanisms**. Other scientists may work at the micro-organism level such as investigating organs and cells in the circulatory system. At this level, the action-formation mechanisms provide a means to describe specific causal relationships that are present where an element’s ‘*action’* leads to another element’s ‘*formation’*. This is similar to our concept of representing causal

relationships between conditions in safety causal model. For example, a causal relationship "*fire cooks food*" could be a specific micro-level causal relationship for the macro-level "*human eats animal*" relationship (see Figure 33).

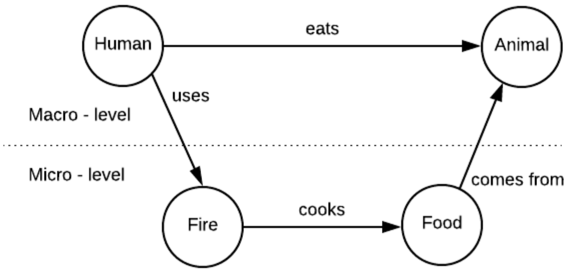


Figure 33. Example of a Coleman's Boat of Causal Pathway

3.3.2.3. Applications

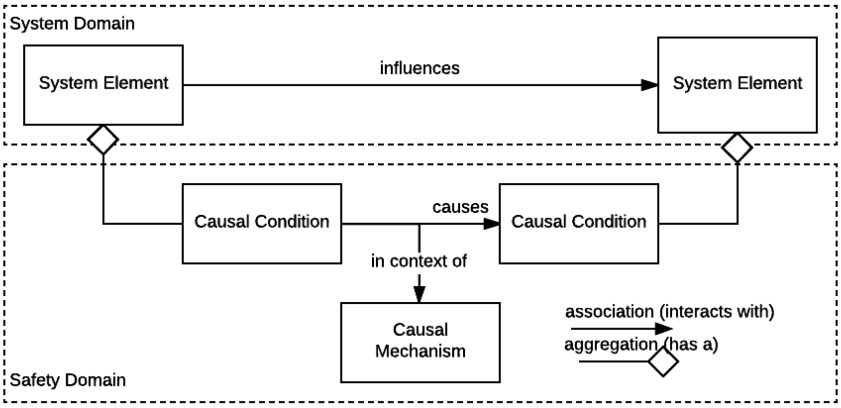


Figure 34. Causal Relationship Model of Safety Assessment

We apply a similar layered concept like the Coleman's model in our Causal Relationship Model of Safety Assessment (CRMSA) (see Figure 34). To appreciate the model, we apply the CRMSA on a simple safety assessment example about an engineer performing a runway repair (see Figure 35).

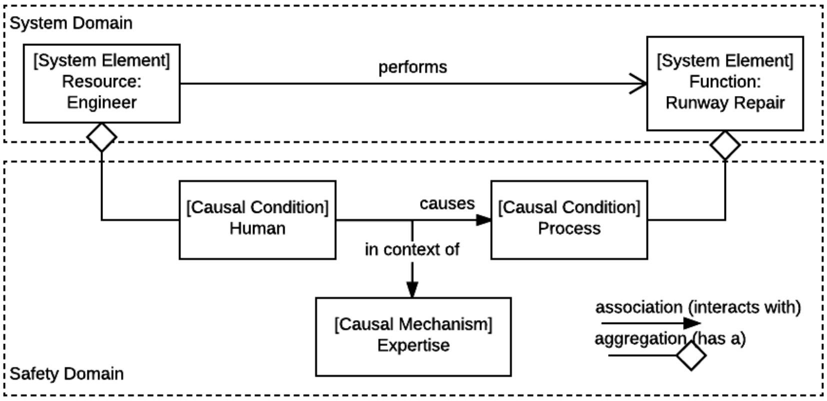

System Domain
[System Element] Resource: Engineer
performs
[System Element] Function: Runway Repair
[Causal Condition] Human
causes
[Causal Condition] Process
in context of
[Causal Mechanism] Expertise
association (interacts with)
aggregation (has a)
Safety Domain

Figure 35. Illustration of CRMSA

Example. We break the example into two aspects: system and safety domains.

System Domain. In safety assessment, causal relationships are derived from some macro-level association at the system perspective. This is like the relationships between system and safety elements that we have introduced in the CMSS. In the runway repair example, the two system elements are "the engineer" and "the runway repair", where the former (which is a type of system resource) is expected to perform the latter (which is a type of system function) at the system level.

Safety Domain. As explained in section 3.3.2.1, a causal relationship between two causal conditions exists in the context of some causal mechanisms. This causal relationship is like the "action-formation mechanisms" under the micro-level perspective of the Coleman's model. The causal mechanism provides the narrative or instantiation of how causal conditions can influence each other. For the runway repair example, a potentially safety-critical causal mechanism between the engineer (being considered as a human causal condition) and the runway repair function (being considered as a process causal condition) is the level of expertise. In other words, there is safety concern that the uncertainty may result in a lack of expertise by the engineer in carrying out the process of runway repair. This situation may potentially become a hazard.

3.4. **Summary**

In this chapter, we have presented the theoretical foundation to manage epistemic uncertainties. The chapter answers two questions that we have defined in the beginning, i.e. 1) "*WHAT aspects of the epistemic uncertainties do we want to manage?*" and 2) "*WHERE do these epistemic uncertainties reside in safety assessment model?".*

To answer the "WHAT", we have explained the need to identify, document, track and address epistemic uncertainties.

We need to identify epistemic uncertainties because

- Epistemic uncertainty is inherent in safety assessment
- There is no commonly practised approach of identifying epistemic uncertainty
- Safety assessment is more complete by recognising and identifying plausible uncertainty than assuming that there is no uncertainty

We need to document epistemic uncertainties because

- There is no established approach of documenting the nature of the uncertainties within a model
- Documentation of uncertainties in a model reduces the danger of ignoring safety-critical uncertainty
- Documentation of uncertainties reduces confirmation bias by being open about the reasons to believe and the reasons to doubt an assessment

We need to track and address epistemic uncertainties because

- The nature (and responsibility) of system safety is that it demands a systematic and through-life management of knowledge
- Epistemic uncertainties in system and safety causal models can change with time
- The capability to address uncertainty when it is safety-critical increases the confidence in the safety assessment

To answer the "WHERE", the CMSS and CRMSA are constructed to help users locate epistemic uncertainties during safety assessment. Most importantly, the CMSS and CRMSA are used to emphasise that epistemic uncertainties in safety assessment reside in the underlying system and safety causal models. The system models are necessarily to represent the real world. When there is insufficient information regarding the real world, the system models will have uncertainty. The safety causal models are used during the conduct of safety assessment. Similarly, epistemic uncertainty can be exacerbated if there is insufficient information associated with the safety causal models.

Chapter 4 - An Approach to Identify, Document, Track and Address Epistemic Uncertainties

4.1. Introduction

In the previous chapter, we have established the principles that are important to manage epistemic uncertainties: identify, document, track and address. Following that, we shall make use of these principles in this chapter to develop a concrete approach to manage epistemic uncertainties in safety assessment. For the approach to be effective, it is developed based on three desired outcomes: comprehensive, targeted and feasible.

Outcome 1: Be Comprehensive. To be comprehensive, the approach should be able to manage both known and unknown uncertainties. As explained in section 2.1.1, known uncertainties are uncertainties that stakeholders are aware of and unknown uncertainties are uncertainties that stakeholders are not even aware of. In our approach, we will proactively manage the risk due to uncertainties associated with safety assessment. In traditional safety assessments, the focus is usually on known uncertainties where stakeholders have more confident with their safety impacts. When there is pressure to perform within limited resources, stakeholders would most likely be more confident to analyse safety concerns based on known uncertainties than to analyse unknown uncertainties, such as worrying about an emerging technology that none of the stakeholders is aware of. However, not having the capacity or confidence to focus on unknown uncertainties does not mean unknown uncertainties are not safety-critical.

Outcome 2: Be Targeted. To be targeted, the approach should differentiate the management of certainties and uncertainties. In other words, we advocate that stakeholders should differentiate causal relationships in safety assessment that are plausible-but-uncertain from those that are more certain, instead of treating all the same in traditional safety assessment approach. The following provides a contrast between the traditional approach and our approach of managing epistemic uncertainty:

- **Traditional Approach**. In traditional safety assessment, some safety analysts would carry out risk assessment on issues and conditions without putting a lot of focus on ram-uncertainty. This was echoed by McDermid [68], when he commented that '*the fundamental assumptions which underpin the classical approaches to safety analysis do not reflect the importance of model uncertainty*.' He further suggests that safety engineers should "*recognise the need to analyse the uncertainty in the environment (the demands to which the system is subjected) and to employ a proper balance* *between analysing the system and its environment*", as well as "*investigate and refine the risk assessment methods to include model (epistemic) uncertainty, especially as this is likely to be the dominant area of risk in many (e.g. software intensive) applications*."

There are also examples where ram-uncertainty has been recognised in existing safety assessment techniques but are not always fully documented and managed. For example, in a research by Ferdous et al [144], they raise the issues of uncertainty in Quantitative risk analysis (QRA), a popular technique in safety assessment to evaluate the likelihood, consequences and risk of adverse events. QRA can be based on event (ETA) and fault tree (FTA) analyses and Ferdous et al describe two known uncertainties regarding the QRA. The first known uncertainty is that the likelihood values of input events are based on "*probability distributions that are often hard to come by and even if available, they are subject to incompleteness (partial ignorance) and imprecision*". The second known uncertainty is that the interdependence among the events (for ETA) or basic events (for FTA) may not be guaranteed, which is an assumption for the method to be valid. While both uncertainties are known, Ferdous et al observe that both uncertainties may not always be tracked and often assumed not to affect the safety assessment "*for the purpose of simplicity*".

- **Our Approach**. In our approach, we advocate that epistemic uncertainty, or what we refer to as plausible-but-uncertain causal relationships in safety, should not be discarded. This is important as the level of certainty may not have direct correlation with the level of risk related to the causal relationships. Instead of discarding it, plausible-but-uncertain causal relationships should be managed differently from more certain causal relationships. For example, there could be plausible-but-uncertainty causal relationship about the effect of a weapon system on another legacy system which is not established in the current concepts of operation. Instead of ignoring this uncertainty, it should be captured as a plausible-but-uncertain causal relationship and managed through life. This calls for the commitment to track uncertainties in existing safety techniques till the relevant information is available down the lifecycle. The aim is to encourage stakeholders to defer judgement on potentially safety-critical causal relationships, rather than discard them early. This is important as the stakeholders may not fully appreciate the risk of the epistemic uncertainty in a complicated system at the point where it was first being considered.

Outcome 3: Be Feasible. To be feasible, the approach should be integrated with existing safety assessment techniques. As part of the system safety lifecycle, it is less disruptive if the management of epistemic uncertainties is integrated into existing processes as much as possible. Furthermore, the efforts to manage the uncertainties should be reasonably efficient such that the approach is still practical to be applied by industry. By being less disruptive, stakeholders can be more receptive to the approach and be more supportive in managing such uncertainties. This enables stakeholders to easily recognise the core of the new approach and apply it in their safety assessments. In our approach, we aim to simplify as much of the core concepts and consolidate them into a single user guide so that it is convenient for the stakeholders to refer to during safety analyses. To summarise, by being feasible, we introduce two requirements that the approach must achieve:

- The approach needs to be useful to complement existing processes and still provide its intended purpose, and
- The approach needs to be practical whereby the amount of effort to implement it is reasonable for the user, given the benefit gained.

4.1.1. Overview of the Target-Analysis-Goal (T.A.G.) Approach

In the previous chapter, it has been determined that epistemic uncertainties need to be identified, documented, tracked and addressed. In this chapter, we shall introduce the **Target-Analysis-Goal** (T.A.G.) approach to help stakeholders accomplish these principles. We name it the T.A.G. approach as it comprises three steps according to the three principles in chapter 3: identify **Target**, document **Analysis**, as well as track and address **Goal**.

The T.A.G. approach is summarised in Table 14. On the left are the principles to be achieved according to chapter 3. In the middle, we list our initial requirements as part of the T.A.G. approach to help achieve the principles. On the right is the expected actions to be carried out by stakeholders to make use of the initiatives to achieve the desired principles.

For example, to better identify epistemic uncertainties (the first principle), we have developed a taxonomy of causal mechanisms to help recognised uncertainties that may not be covered in the original safety assessment. The users would have to use the taxonomy as reference to conduct the physical identification.

Table 14. The T.A.G Approach to Proactively Manage Epistemic Uncertainty

STEP	**Principles**	**Target-Analysis-Goal (T.A.G.) approach**		**Elaborated in**
		Initial Requirements	**Stakeholders' Actions**	
1	**Identify**	**Target**		
	Epistemic uncertainties in the underlying models of safety assessment should be identified	**Taxonomy** of causal mechanisms for stakeholders to recognise epistemic uncertainties not covered in original assessment	Use the taxonomy of causal mechanisms to identify plausible-but-uncertain causal relationships in safety assessment	Section 4.2
2	**Document**	**Analysis**		
	Epistemic uncertainties in the underlying models of safety assessment should be documented	**Process and factors** to prioritise epistemic uncertainties analysis so as to document the findings in existing safety assessment technique	Prioritise and document epistemic uncertainties as part of existing safety assessment techniques	Section 4.3
3	**Track and Address**	**Goal**		
	Epistemic uncertainties in the underlying models of safety assessment should be tracked and addressed in a systematic manner through-life	**Method** to develop actionable goals to track epistemic uncertainties through-life and address when some thresholds are met	Implement the initiatives to fulfil the action plan to track and address epistemic uncertainties through-life	Section 4.4

4.1.2. When Is T.A.G. Approach Applicable in Safety Assessment?

Recall that in chapter 1, we introduced the concept of risk and confidence, as well as the understanding that it is possible to track known epistemic uncertainty in safety assessment only when the risk due to the uncertainty can be tolerable. With regards to the T.A.G. approach (see Table 14), the two steps of identifying and documenting epistemic uncertainties should be conducted in parallel with the ongoing risk analysis. As for the third step of tracking and addressing the epistemic uncertainties, it can happen through two scenarios:

- **Scenario 1**. Clarity in the epistemic uncertainty can reduce the assessed risk such that it becomes tolerable or acceptable.

- **Scenario 2**. While the assessed risk is tolerable, there is not enough confidence in the risk analysis to make it acceptable due to the epistemic uncertainty.

The way T.A.G. approach is integrated with existing safety assessment is shown in Figure 36.

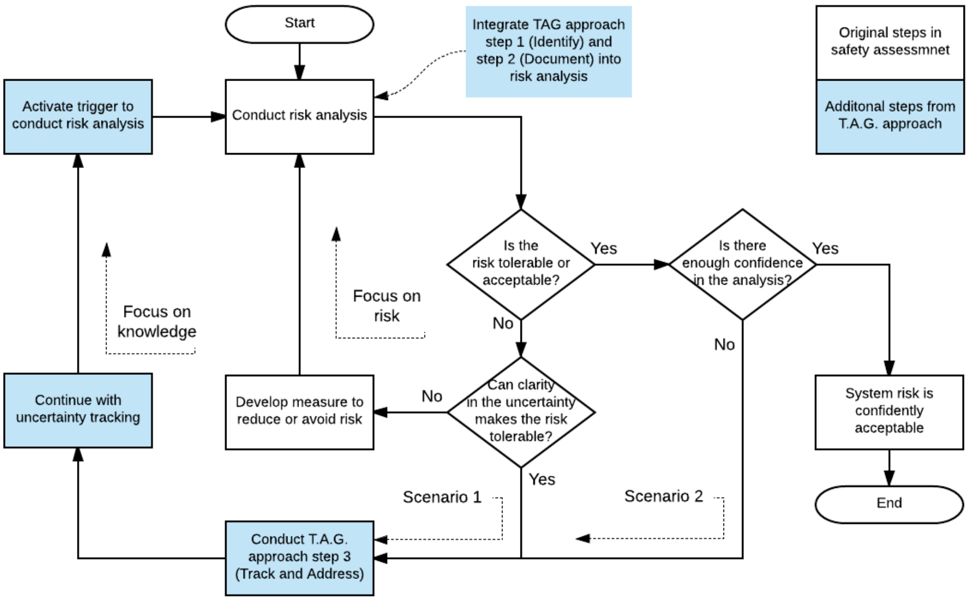

Start
Integrate TAG approach step 1 (Identify) and step 2 (Document) into risk analysis
Original steps in safety assessmnet
Additonal steps from T.A.G. approach
Activate trigger to conduct risk analysis
Conduct risk analysis
Is the risk tolerable or acceptable?
Yes
No
Is there enough confidence in the analysis?
Yes
No
Focus on knowledge
Focus on risk
Continue with uncertainty tracking
Develop measure to reduce or avoid risk
No
Can clarity in the uncertainty makes the risk tolerable?
Yes
System risk is confidently acceptable
End
Scenario 1
Scenario 2
Conduct T.A.G. approach step 3 (Track and Address)

Figure 36. Integrating T.A.G. Approach into Existing Safety Assessment

The extent of risk tolerance depends on the phase in the system lifecycle in which the uncertainty is discovered. For example, an uncertainty about the system design during the early conceptual phase of the system lifecycle tends to be tolerable since there is still time to reassess the risk further down the system lifecycle when the system design is more definite. In contrast, an uncertainty about the operational concept during deployment flight trial may need to be resolved quickly since there is not much time till the operation phase of the system. For example, we can expect a higher risk appetite for the uncertainty about the system design during conceptual phase, vis-à-vis the uncertainty about the operational concept during flight trial. As a result, it may be more applicable to use the T.A.G. approach to track uncertainty during conceptual phase than during flight trial. Having said that, the T.A.G. approach can still be used to track the uncertainty about the concept of operation by specifically annotating the urgency of the tracking, if the risk is assessed to still be tolerable.

In the remaining sections, we will define the three steps under the T.A.G. approach.

4.2. Initiative 1: A Taxonomy of Causal Mechanisms to Identify Uncertainty Target

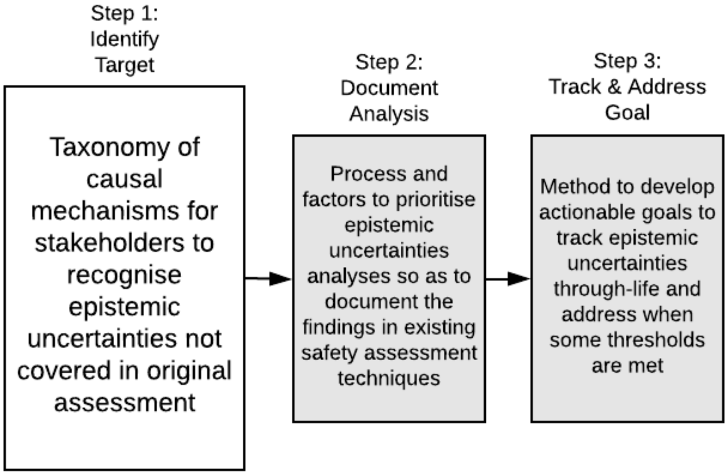

Step 1:
Identify
Target
Taxonomy of causal mechanisms for stakeholders to recognise epistemic uncertainties not covered in original assessment
Step 2:
Document
Analysis
Process and factors to prioritise epistemic uncertainties analyses so as to document the findings in existing safety assessment techniques
Step 3:
Track & Address
Goal
Method to develop actionable goals to track epistemic uncertainties through-life and address when some thresholds are met

Figure 37. T.A.G. Approach First Step: Identify Target

To support step 1 of the T.A.G. approach, we develop a taxonomy of causal mechanisms (see Figure 37) to identify plausible-but-uncertain causal relationship. Recall from Figure 36 that Step 1 of the T.A.G. approach to identify epistemic uncertainties should be integrated with risk analysis in existing safety assessment techniques. The aim is to recognise epistemic uncertainties that have not been covered in the original safety assessment.

The first principle of managing epistemic uncertainties states that the epistemic uncertainties in the underlying models of safety assessment should be identified (see section 3.2.1). To be identified, stakeholders must first be able to recognise these uncertainties, be it initially known or unknown Safety assessment involves finding causal conditions that could be hazardous. As discussed in section 3.2.1, safety assessment inherently contains uncertainties. A taxonomy can actively prompt the stakeholders to recognise those plausible-but-uncertain causal conditions in safety assessment and help stakeholders identify safety-critical epistemic uncertainties early.

To recognise these uncertainties, the stakeholders should be aware of where to search for them. There is a lot of understanding of the nature of causal relationships from collective wisdom. In the spirit of good safety engineering practice, the taxonomy of causal mechanisms that we are creating aims to harness the maximum effect of prior knowledge about credible causal mechanisms, especially those that may reveal certain types of unsafe situations related to the system of interest. This is akin to the ‘observability-in-depth’ principle under system safety [145] to identify hazards. The principle advises stakeholders to scan the horizon in depth for as many scenarios as possible that can transit a system to an increasingly hazardous state. This encourages stakeholders to shift the boundary of knowledge from not knowing to knowing about critical epistemic uncertainty, through surfacing and modelling of previously unknown uncertainty.

Besides knowing where to search, stakeholders need to do the search in an effective and efficient way in view of limited resources.

- **Searching in an Effective Way**. To be effective, the search space should be as comprehensive as possible. A taxonomy can do this by specifying a reasonable coverage of diverse issues to help stakeholders recognise a wide range of safety-critical concerns and causal relationships.

  Since most of the literature about safety are specialised in their domains, none of them can serves as an isolated and complete guide to discover all types of hazards. For example, Shappel's Human Factors Analysis and Classification System (HFACS) [146] provides a detailed review of issues related to human such as complacency, distraction and confusion; but does not consider technology issues that can also cause hazards. His work can be complemented by O'Halloran's taxonomy of Failure Mode/Mechanisms Distribution (FMD) [147] that lists the possible safety-critical issues resulting from technical properties such as kinetic, chemical and electrical. In a different study, Endsley's taxonomy of situation awareness error [80] focuses on information and decision making, which provides another dimension of causal mechanism.

  Specifically, in the military domain, the Defence Lines of Development (DLoD) is a useful reference to consider when conducting military system development. This is supported by the studies from Yue and Henshaw [148] and Burton, Paige, Poulding and Smith [149]. Both groups of author highlighted that the DLoD components (i.e. Training, Equipment, Personnel, Infrastructure, Concepts & Doctrine, Organisation, Information and Logistics) are fundamental components to guide military through life capability management for the UK Ministry of Defence. To make the search more comprehensive, we have also included similar capability management considerations from the Department of Defence in Australia (known as 'Fundamental Inputs to Capability') [150], United States (DOTLMPF) [151] and Canada (PRICIE) [152].

  To create a credible taxonomy, an extensive literature review of different subjects that are related to safety was conducted (see Annex A for the range of literature surveyed). The literature review involved identifying potential causal mechanisms that may result in unsafe situations. These causal mechanisms covered a wide range of topics such as system safety, human factor ergonomic, project uncertainty, taxonomy of safety-related subjects and situational awareness.

- **Searching in an Efficient Way**. To be efficient, stakeholders should be guided in some ways to narrow down the search for uncertainties whenever they have some ideas of where to search. Currently, most stakeholders depend on random experiences or unsystematic brainstorming among those that are present during the safety assessment to discover known uncertainties. Without a systematic way of searching, the process may end up being laborious and inefficient. Our taxonomy helps the stakeholders to narrow their search by grouping the causal mechanisms into primary and secondary causal conditions for more efficient searching.

From the initial list of causal mechanisms and the suggested classifications within the literature, six primary causal conditions were defined: Human, Organisation, Technology, Process, Information and Environment (see Table 15). This list is collectively known as the Human, Organisation, Technology – Process, Information, and Environment (HOT-PIE) taxonomy.

Table 15. Descriptions of the Six Primary Causal Conditions in HOT-PIE Taxonomy

No.	Primary Causal Condition	Description
1	Human (section 4.2.1)	People that directly or indirectly influence the system. They include managers, users, maintenance engineers, contractors, senior leaderships and operators.
2	Organisation (section 4.2.2)	A discrete, relatively stable group of individuals linked by relatively stable patterns of interaction and pursuing common objectives affecting the system [153]
3	Technology (section 4.2.3)	Application of scientific knowledge in practice, especially for system acquisition in our context.
4	Process (section 4.2.4)	A systematic series of operations that are performed to produce or service something.
5	Information (section 4.2.5)	Data that are conveyed or represented in such a way to produce knowledge in a system.
6	Environment (section 4.2.6)	The surroundings within which the system operates.

From the six primary causal conditions, we further subdivide them into fifteen secondary casual conditions (see Table 16) that comprises over three hundred of possible causal mechanisms (see the full list in Annex K). While this may not be the only way to sub-divide the primary causal condition based on the causal mechanisms consolidated, having such subdivision would help stakeholders narrow down to an area of interest when they have some ideas where to search.

In the rest of section 4.2, we shall describe these causal mechanisms in detail based on the six primary causal conditions.

Table 16. HOT-PIE Primary and Secondary Causal Conditions

Primary	Secondary	Primary	Secondary
Human	**H1: Manpower**	Process	**P1: Nature**
	H2: Mental state		**P2: Phase**
	H3: Action		
Organisation	**O1: Management**	Information	**I1: Knowledge**
	O2: Policy		**I2: Error**
	O3: Resource		

Primary	Secondary	Primary	Secondary
Technology	**T1: Machine**	Environment	**E1: Physical**
	T2: Property		**E2: Non-physical**
	T3: Support		

4.2.1. Primary Causal Condition: Human

We have subdivided the Human primary causal conditions into three secondary causal conditions: Manpower, Mental state and Action (see Table 17). The focus of each secondary condition is explained in the table.

Table 17. Causal Mechanisms under the Human Primary Causal Conditions

Secondary Conditions	**Causal Mechanisms**
H1: Manpower (capability of the individuals to work on the system)	expertise[154-156] staffing[146, 154, 156-160] mix[158] ownership[154] experience[154, 158] leadership[142, 146] skill[146, 156, 158, 159, 161-163] ability[158] characters[164] individualistic[165] demographic[165] cultural[165] obligation[166] survivable[158] stakeholders[66, 156, 167-169] user[170] turnover[156] education[156]
H2: Mental state (frame of mind of the individuals working on the system)	escalation[142] brokerage[142] free rider[142] convention[142] norm[142] selective benefit[142] morale and motivation [142, 156, 158] social[163, 171] deliberate[161] esteem[166] complacency[146] stress[146] overconfidence[146] fatigue[146] distraction[80, 146] confusion[146] health[158] comfort[158] visual limitation[146] illness[146] injury[158] disability[158] hearing limitation[146] cognitive[158] physical[158, 172] sensory[158] team dynamic[158, 159] aptitude[158] emotional[172] psychic[173] conflict of interest[156] lack purpose[156] perception[80, 174] memory fail[80] poor mental model[80] incorrect mental model[80] reliance on default[80]
H3: Action (activity conducted by an individual on the system)	operation[155] network[142] broadcast[142] rumour[142] communication[146, 154, 156, 159] open[159] interrelation[159] atmosphere[159] engagement[173] coordination[146] omission[80, 161] commission[161] extraneous act[161] observation[164] interpretation[164] overcommit[166] performance slip[175] specification slip[175] lapse-forgot[175] lapse-overlook[175] rest[146] preparation[146] intentional violation[159, 163, 176] behaviour[173] lack involvement[156] influence[174]

Example. As an illustration, Expertise is one of the causal mechanisms under the secondary casual condition of Manpower (H1). We can refer to references [154-156] for more details about how the subject of Expertise can cause uncertainties related to Human causal conditions. For example, in the RSAF safety assessment for new system, the absence of some SMEs during safety assessment can potentially means that certain important safety concerns were not surfaced. This uncertainty is captured using the CRMSA model in Figure 38 for illustration.

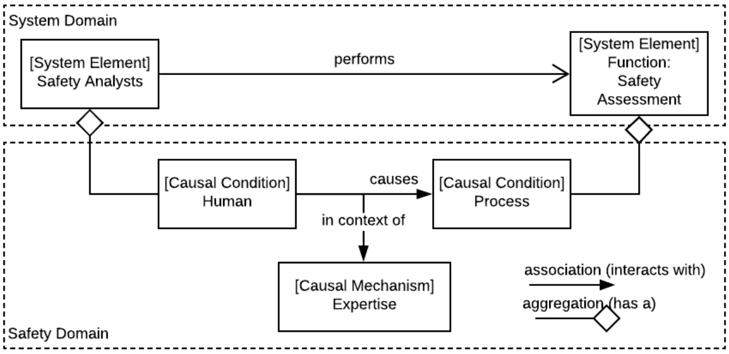

System Domain
[System Element] Safety Analysts
performs
[System Element] Function: Safety Assessment
[Causal Condition] Human
causes
in context of
[Causal Condition] Process
[Causal Mechanism] Expertise
association (interacts with)
aggregation (has a)
Safety Domain

Figure 38. Illustration of a Human Related Uncertainty represented with CRMSA

4.2.2. Primary Causal Condition: Organisation

We have subdivided the Organisation primary causal conditions into three secondary causal conditions: Management, Policy and Resource (see Table 18). The focus of each secondary condition is explained in the table.

Table 18. Causal Mechanisms under the Organisation Primary Causal Conditions

Secondary Conditions	Causal Mechanisms
O1: Management (factors affecting the control of the system)	supervision[142, 146, 155] audit[142] communication[164] structure [146, 164, 168, 174, 177] levels of domain[174] role ambiguity and conflict[165] schedule[165] demand[166] feedback and refine[146] company[160] project size[156] project uniqueness[156] project density[156]
O2: Policy (principle of actions adopted by the system)	regulation and control[142, 160, 167, 174] job future and security[165, 166], culture and climate[146, 156, 162, 165, 177, 178] reward and recognition[165, 166] incompatible goals[156, 159, 176] trade-off[159] ambiguous goal[156] narrow goal[156] expectation[156] customer satisfaction[170]
O3: Resource (supporting assets that are needed for the system to function properly)	training facility[142, 155, 164, 170, 176] material[154, 155, 162] supplier management[142, 156, 169] support facility[146, 156, 161, 170, 172] time phase[157, 161] time step[157, 161] project urgency[156] allocation[146] monetary[146, 156] instructional[158] unrealistic time frame[156] outsource management[156] infrastructure[163, 170] test equipment[178] test procedure[178]

Example. As an illustration, Supervision is one of the causal mechanisms under the secondary casual condition of Management (O1). We can refer to references [142, 146, 155] for more details about how the subject of Supervision can cause uncertainties related to Organisation causal conditions. For example, for the RSAF IAD SoS [135] mentioned in section 3.2.3.1, multiple systems may interoperate with each other to provide the enhanced IAD SoS capability. Junior operators for each system may only have knowledge for their own system, while the senior operators have additional knowledge to supervise interoperability with other systems within the IAD SoS. In certain situation, there could be uncertainty about the supervisory level needed to operate a system if the actual complexity of the operation (e.g. the type of systems involved) is not confirmed. This uncertainty is captured using the CRMSA model in Figure 39 for illustration.

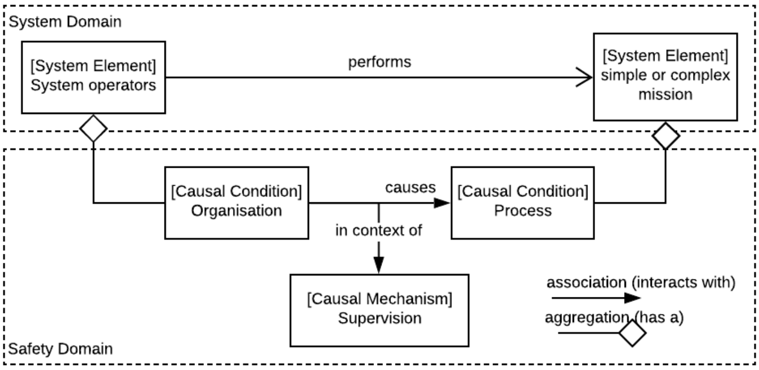

System Domain
[System Element] System operators
performs
[System Element] simple or complex mission
[Causal Condition] Organisation
causes
[Causal Condition] Process
in context of
[Causal Mechanism] Supervision
association (interacts with)
aggregation (has a)
Safety Domain

Figure 39. Illustration of an Organisation Related Uncertainty represented with CRMSA

4.2.3. Primary Causal Condition: Technology

We have subdivided the Technology primary causal conditions into three secondary causal conditions: Machine, Property and Support (see Table 19). The focus of each secondary condition is explained in the table.

Table 19. Causal Mechanisms under the Technology Primary Causal Conditions

Secondary Conditions	Causal Mechanisms
T1: Machine (a component that has a definite function and perform a task under the system)	hardware capability[155, 157, 163, 167, 169, 174, 176, 177] hardware compatibility[178] technical[168, 171, 179] equipment [146, 161, 164] interface[146, 164] link[163] node[163] display[146] construction[162] software[147, 157, 163, 167, 169, 174, 177, 178] communication[147, 170, 176] engineering[66] mobility[163] traffic[163] area coverage[163] services[170] tool[170] technique[170] abstraction[154] working range[154] tech change[154, 156] innovation[154] complexity[146] availability[159] function[159] standardisation[159] features[156] customisation[156] interdependency[163]
T2: Property (technical attribute, quality, or characteristic)	energy[157] kinetic[154] biological[154] acoustical[154] chemical[154] electrical[154] mechanical[154] electro-magnetic[154] thermal[154] radiation[147, 154] bonding[147] buckling[147] change in property[147] corrosion[147] cracking[147] deformation[147] fatigue[147] seizure[147] impact[147] rupture[147] voiding[147] wear[147, 178] breakdown[147] contamination[147] diffusion[147] degradation[147] incorrect current[147] punch through[147] leak[178] loose[178] drift[178] synchronisation[178]
T3: Support (assistance to actual work)	system design[162, 176] tool design[165] tool usability[165] work area design[165] task design[146, 176] medium[163]

Example. As an illustration, System Design is one of the causal mechanisms under the secondary casual condition of Support (T3). We can refer to references [162, 176] for more details about how the subject of System Design can cause uncertainties related to Technology causal conditions. For example, in the CHALLENGER space shuttle accident, one of the concerns was the uncertainty surrounding the behaviour of the O-ring under cold temperature. Unfortunately, this design uncertainty was not raised sufficiently to create the awareness that it might cause catastrophic failure to the space shuttle. This uncertainty is captured using the CRMSA model in Figure 40.

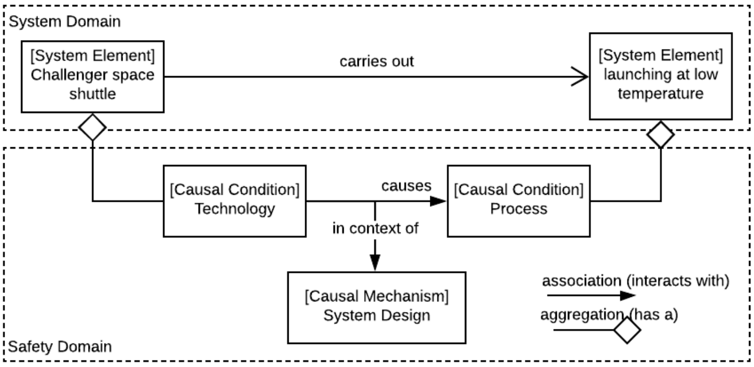

System Domain
[System Element] Challenger space shuttle
carries out
[System Element] launching at low temperature
[Causal Condition] Technology
causes
[Causal Condition] Process
in context of
[Causal Mechanism] System Design
association (interacts with)
aggregation (has a)
Safety Domain

Figure 40. Illustration of a Technology Related Uncertainty represented with CRMSA

4.2.4. Primary Causal Condition: Process

We have subdivided the Process primary causal conditions into two secondary causal conditions: Nature and Phase (see Table 20). The focus of each secondary condition is explained in the table.

Table 20. Causal Mechanisms under the Process Primary Causal Conditions

Secondary Conditions	Causal Mechanisms
P1: Nature (types of systematic operations performed in the system)	segregation[154] systematic[154] oversight[146, 154] procedure [146, 154, 157, 159, 161, 162, 164, 167, 176-178] practice[154, 167] overload[80, 165] control[157, 165] autonomy[165, 172] repetitiveness[165, 174] feedback[165, 172] ability to learn[165] input[157] output[157] lower level failure[163] cascade failure[163] delay[163]
P2: Phase (a distinct period in the system's lifecycle)	design and plan[164, 179] validation[154] verification[154] manufacturing[66] operation[66] risk management[154, 156, 158, 176] review[154] maintenance[159, 176, 178, 179] housekeeping[176] inspection[179] supervision[179] work[160, 170, 171, 177] training[159, 161] execution and operation[146, 161, 170, 178] mis-operation[161] task[165, 168, 169] sense-making[170] decision making[170] thinking[170]

Example. As an illustration, Autonomy is one of the causal mechanisms under the secondary casual condition of Nature (P1). We can refer to references [165, 172] for more details about how the subject of Autonomy can cause uncertainties related to Process causal conditions. For example, for the RSAF IAD SoS [135], we have mentioned that junior operators may only have knowledge for their own system, while the senior operators have additional knowledge to supervise interoperability with other systems. The organisation needs to recognise the limitation of the junior operators' knowledge and calibrate the level of autonomy given to them in operating the system either during standalone mode or integrated mode with other systems. Unfortunate, this calibration can only be confirmed when the concept of operation is finalised, which may only happen later in the system lifecycle. This uncertainty about the level of autonomy given to junior operators is captured using the CRMSA model in Figure 41 for illustration.

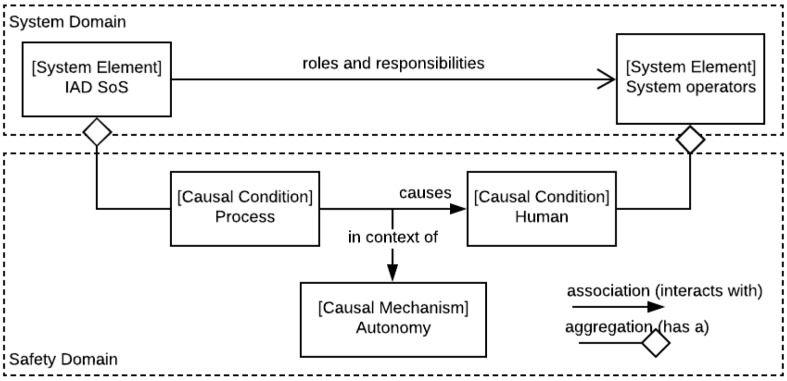


Figure 41. Illustration of a Process Related Uncertainty represented with CRMSA

4.2.5. Primary Causal Condition: Information

We have subdivided the Information primary causal conditions into two secondary causal conditions: Knowledge and Error (see Table 21). The focus of each secondary condition is explained in the table.

Table 21. Causal Mechanisms under the Information Primary Causal Conditions

Secondary Conditions	Causal Mechanisms
I1: Knowledge (facts and skills acquired to understand the system)	procedure[155] standard[155] method[155] assumption[161] policy[146, 169] rule[162, 167] guideline[157] precondition[157] type of info[164] manual and checklist[146] protocol[159, 163] roles and responsibilities[156] best practice[156] data[156] concept[156] no fault found[178] rationalities[174] evidence[174] values[174] fluctuation[174] customer requirements[170] codified information[170]
I2: Error (the state or event of being wrong)	application error[175] assumption error[175] syntax error[175] requirement error[175] lack of distinction[175] lack of awareness[175] insufficient knowledge[175] situational awareness error[159] incomplete specification[156] conflicting requirements[156] info processing problem[156, 170] data unavailable[80] data not detected[80] decisional error[174] executional error[174]

Example. As an illustration, Standard is one of the causal mechanisms under the secondary casual condition of Knowledge (I1). We can refer to references [155] for more details about how the subject of Standard can cause uncertainties related to Information causal conditions. For example, in the NIMROD aircraft example, there was concern about compliance with standards by subcontractors during modification of the aircraft [180]. In Russell's paper, he mentioned that some subcontractors might not be ISO9000 series complaint and the MOD was unable to verify that since it was depended on the next higher level of subcontractor to check the standards of those below them. Russell warned that "*although the contractor certified to manufacture parts, they did not take on the role of testing parts they had sub-contracted, relying on the subcontractors own internal testing*". This uncertainty is captured using the CRMSA model in Figure 42 for illustration.

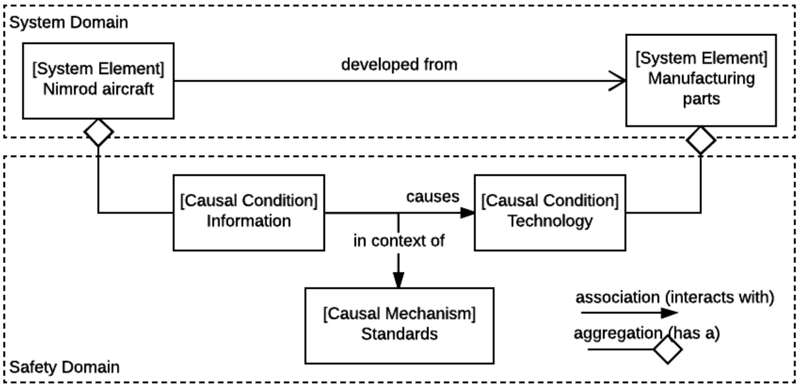

System Domain
[System Element] Nimrod aircraft
developed from
[System Element] Manufacturing parts
[Causal Condition] Information
causes
[Causal Condition] Technology
in context of
[Causal Mechanism] Standards
association (interacts with)
aggregation (has a)
Safety Domain

Figure 42. Illustration of an Information Related Uncertainty represented with CRMSA

4.2.6. Primary Causal Condition: Environment

We have subdivided the Environment primary causal conditions into two secondary causal conditions: Physical and Non-physical (see Table 22). The focus of each secondary condition is explained in the table.

Table 22. Causal Mechanisms under the Environment Primary Causal Conditions

Secondary Conditions	**Causal Mechanisms**
E1: Physical (relating to physics or the operation of natural)	network[142] ambient condition[161, 164] weather[146, 161] orientation[161] size[161] location[161] elevation[161] operating condition[158, 164] noise[146, 165] lighting[146, 165] vibration[146, 165] pollution[165] heat[146] terrestrial[163] meteorological[163] cosmological[163]
E2: Non-physical (not tangible or concrete)	cultural[155, 170, 177] social[167] attitude[155] economic[142, 156, 163, 167, 177] competitiveness[170] political[142, 156, 163, 167, 169, 177] regulatory[170, 177] legal[156, 167] contract[142] propaganda[142] duration[161] delayed[161] alternative[166] strategic interest[166] government[160] complexity[156] security[163]

Example. As an illustration, Operating Condition is one of the causal mechanisms under the secondary casual condition of Physical (E1). We can refer to references [158, 164] for more details about how the subject of Operating Condition can cause uncertainties related to Environment causal conditions. For example, in the NIMROD disaster, the air-to-air refuelling modification was introduced in 1989 to allow the aircraft to change its operating condition such that it can extend its reconnaissance capability. There could have been added vigilance regarding the uncertainty that a change in operating condition might become dangerous. It was unfortunate that the modification was, according to the Haddon report [181], "*in breach of Defence Standard 00-970"* as the blow-off valves and fuel pipes for air-to-air refuelling were fitted too close to other components on the aircraft that can pose catastrophic fire risk. This uncertainty is captured using the CRMSA model in Figure 43 for illustration.

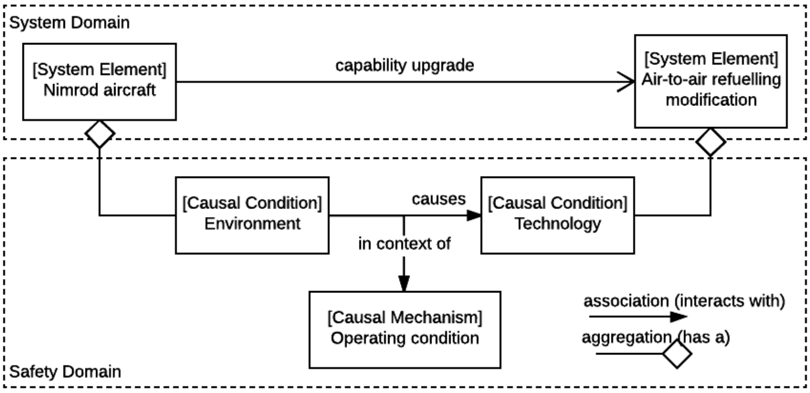

System Domain
[System Element] Nimrod aircraft
capability upgrade
[System Element] Air-to-air refuelling modification
[Causal Condition] Environment
causes
[Causal Condition] Technology
in context of
[Causal Mechanism] Operating condition
association (interacts with)
aggregation (has a)
Safety Domain

Figure 43. Illustration of an Environment Related Uncertainty represented with CRMSA

4.2.7. Apply the Taxonomy to Recognise Uncertainty

With the HOT-PIE taxonomy, stakeholders can apply it during risk analysis to recognise plausible-but-uncertain causal relationships. While the HOT-PIE taxonomy can help to review known uncertainties (e.g. finding other causal mechanisms related to the known uncertainties), its strength lies in prompting the discovery of unknown uncertainties during safety assessment, which the stakeholders may not have discovered by using the existing safety assessment techniques. The intent of HOT-PIE taxonomy is to prompt stakeholders to shift unknown uncertainties into known uncertainties so that they can start making assessment if the uncertainties are safety critical.

Example. For example, in section 4.2.6, the HOT-PIE taxonomy may be able to prompt the NIMROD project team to consider the Operating Condition causal mechanism under the Environment primary causal condition. This might have prompted the safety analysts to consider the impact, trade-off and uncertainty of having air-to-air refuelling capability on the aircraft for extended operation. A follow-up on this uncertainty could have exposed the design flaw in the system modification of having the cross-feed duct too close to the fuel pipe.

To help stakeholders recognise uncertainty, we consider three questions here that can better prepare stakeholders to decide when to use the taxonomy and where to look out for epistemic uncertainties.

- When does epistemic uncertainty occur in a safety assessment?

- Where does epistemic uncertainty occur in a safety assessment?
- Where does epistemic uncertainty occur in a causal relationship?

4.2.7.1. When does Epistemic Uncertainty Occur in a Safety Assessment?

Epistemic uncertainty occurs when there is a lack of knowledge during a safety assessment. This happens when the relevant information is either not available or not sufficient. Using the RSAF safety assessment from section 1.1.3 as an example below:

Examples where information is not available include:

- contractor does not have equipment specification that is provided by sub-contractors,
- expertise not present during the safety assessment, and
- unable to predict all the operation profiles that the system will operate in.

Examples where information is not sufficient include:

- using preliminary system design instead of the final one,
- using initial concept of operation before the final concept is available further down the lifecycle, and
- using flight trial report that only consists of 3 flight profiles out of the desired 4, as the last flight profile was not conducted due to bad weather.

Epistemic uncertainty can also be present if there is suspicious that reliable information now may not be accurate or sufficient in the future. Usually, stakeholders are alert to epistemic uncertainty when there is insufficient information to support a safety assessment. However, even when an information provided for a safety assessment is sufficient during an analysis, the accuracy of the information may change with time and affect the validity of the safety assessment.

4.2.7.2. Where does Epistemic Uncertainty Occur in a Safety Assessment?

This has been discussed extensively in section 3.3.1 under the CMSS model. To recall, epistemic uncertainty can occur in information that may or may not be generated solely for the safety assessment.

Firstly, epistemic uncertainty could be from information that is generated solely for the safety assessment, i.e. in the safety causal models. This includes experimental results, flight trial observations and contractors’ answers to safety queries.

Secondly, epistemic uncertainty could be from information that is generated from other analyses beyond the safety assessment, i.e. in other system models. For example, in the CHALLENGER space shuttle accident, the launching temperature was part of the data retrieved from the space shuttle system (i.e. a form of system model), rather than part of the safety causal model. Other examples of information not directly related to the safety assessment include personal experiences, mental models, system design documents and equipment specifications.

4.2.7.3. Where does Epistemic Uncertainty Occur in a Causal Relationship?

In section 3.3.2, we have established that safety assessment is carried out with the aim of identifying causal relationships that could be hazardous. Uncertainties in such causal relationships can be in two areas: nodes and linkages. To help in explaining the differences, we have reproduced the CRMSA model described in chapter 3 as a reference (see Figure 44):

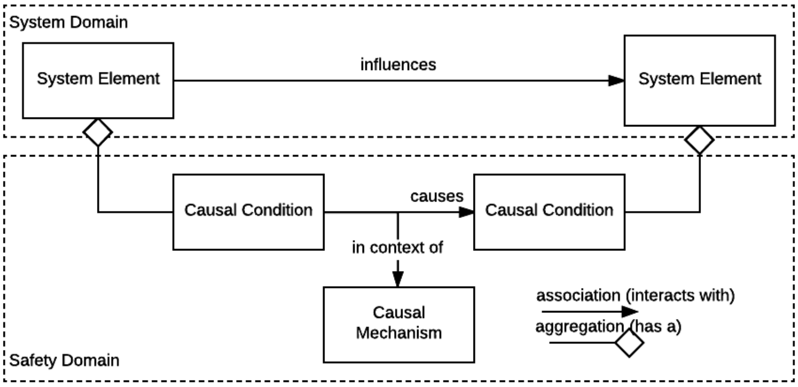


Figure 44. Causal Relationship Model of Safety Assessment

Uncertainty about Nodes. Uncertainty in safety assessment can be located on the nodes in a causal relationship, when they do not have prior relationship with one another. From Figure 44, this can refer to the system elements in the system domain or the causal conditions in the safety domain. For example, there may be a new platform (which is a new system element node) joining a network of systems which can create uncertainty about the safety of the whole network. In the safety domain, stakeholders may recognise that maintenance engineers can operate the system wrongly and dangerously using the maintenance mode, which can be a system hazard. This may not have been considered before and hence, an unsafe maintenance by the engineers can constitute a new causal condition node.

Uncertainty about Linkages. Uncertainty in safety assessment can also be located on the linkage in a causal relationship. From Figure 38, this can refer to the influence between system elements in the system domain or the causal mechanism in the safety domain. For example, for a technical safety assessment, there may be uncertainty about the network messages between a newly installed system element and a legacy system element. This can happen when information about the network interfaces between the new and legacy systems are not sufficiently provided by the system manufacturers. As a result, this creates uncertainty during the hazard analysis on the network messages communicated between the two systems.

4.3. Initiative 2: A Systematic Prioritisation Process to Document Uncertainty Analyses

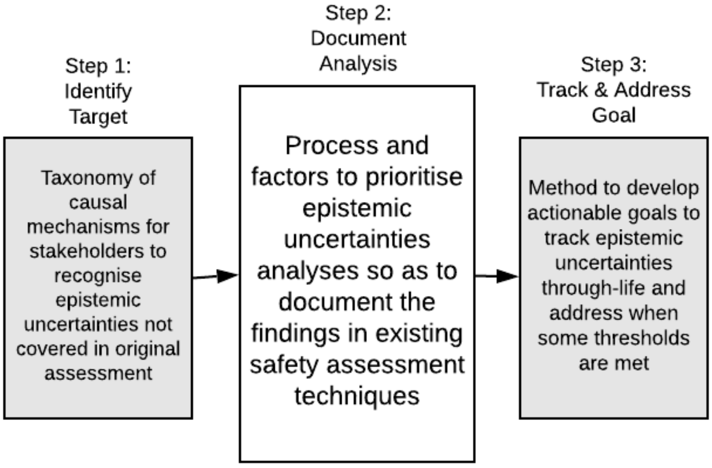

Step 1: Identify Target
Taxonomy of causal mechanisms for stakeholders to recognise epistemic uncertainties not covered in original assessment
Step 2: Document Analysis
Process and factors to prioritise epistemic uncertainties analyses so as to document the findings in existing safety assessment techniques
Step 3: Track & Address Goal
Method to develop actionable goals to track epistemic uncertainties through-life and address when some thresholds are met

Figure 45. T.A.G. Approach Second Step: Document Analysis

To support step 2 of the T.A.G. approach, we introduce a process and factors to prioritise uncertainty analyses that stakeholders would be conducting after identifying these uncertainties (see Figure 45). Such a prioritisation process is needed to help stakeholders to systematically document the epistemic uncertainty discovered from the previous step.

The second principle of managing epistemic uncertainties states that epistemic uncertainties in the underlying models of safety assessment should be documented (see section 3.2.2). Such documentation should trace the reasons why uncertainties are chosen to be investigated further and be integrated with existing process as much as possible.

We have mentioned in section 3.2.2.1. that there is no established approach to documenting uncertainties in current safety assessment techniques. In the beginning of this chapter, we have also emphasised the importance of integrating any new approach with existing safety assessment techniques. For Step 2 of the T.A.G. approach, we aim to integrate our process with the risk analysis process in existing safety assessment techniques. This will be elaborated In section 4.3.1.

With limited resources during system development, stakeholders may end up discarding less certain causal relationships during safety assessment and run the danger of ignoring safety critical uncertainties (see section 3.2.2.2). We have also raised the concern that stakeholders' do not usually document the reasons they choose or discard causal relationships for analysis during safety assessment (see section 3.2.2.3). Since prioritisation is unavoidable in complicated system lifecycle where resources are limited to accomplish the many tasks, it is important to document such considerations for traceability reason. To be explicit about such selections, we have developed factors for stakeholders to consider when prioritising which uncertainty analyses they would manage first. For example, there may not be enough manpower or expertise to manage a certain uncertainty; or there may not be enough time to investigate an uncertainty due to an imminent deadline. These factors shall be elaborated in section 4.3.2.

4.3.1. Process to Integrate Uncertainty Management in Existing Techniques

To understand the process to integrate the T.A.G. approach in existing safety assessment techniques, we will first describe a generic process of how through-life safety assessment is currently conducted (see section 4.3.1.1). This is followed by a description of the generic process of managing epistemic uncertainties in through-life safety assessment (see section 4.3.1.2).

4.3.1.1. Generic Process of Through-life Safety Assessment

In section 2.2.1, we have described the typical safety assessment process in system development lifecycle. Taking the example of the MoD CADMID acquisition lifecycle mentioned in chapter 3, there are multiple milestones across the lifecycle where safety assessments are conducted. Such assessments would result in the generation of safety artefacts such as safety case evidence and reports. Safety artefacts are something that are observed or intentionally produced during safety assessment. As illustrated in Figure 46, multiple safety cases are generated at every milestone, which are refined with the latest assessment about the safety of the system.

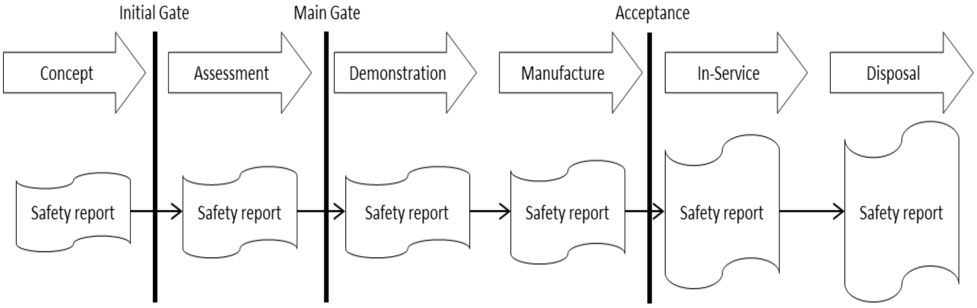


Figure 46. MoD CADMID Cycle

Using the CADMID lifecycle as an example, we have developed a process flow to show a generic process of safety assessment throughout the system lifecycle (see Figure 47).

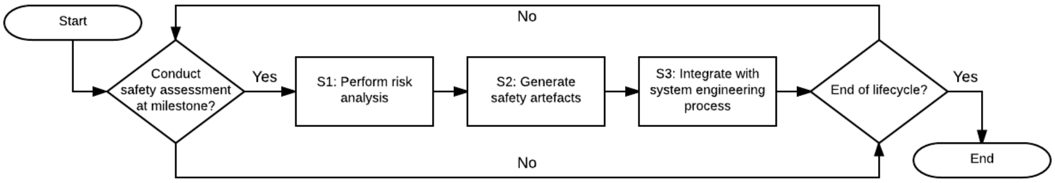

Start
No
Conduct safety assessment at milestone?
Yes
S1: Perform risk analysis
S2: Generate safety artefacts
S3: Integrate with system engineering process
End of lifecycle?
Yes
End
No

Figure 47. Generic Process Flow for Through-Life Safety Assessment

From the figure, we can see that safety assessment generally consists of three steps (represented by the rectangular boxes). The requirement to conduct safety assessments at different milestones will repeat until the system reaches the end of its lifecycle (e.g. decommissioned, disposal). The three generic steps within a safety assessment are as follows:

Step 1 (S1): Perform Risk Analysis. A safety assessment consists of one or more risk analyses. The nature of risk analyses depends on the types of safety assessment techniques. An overview of risk analyses is captured in section 2.2.1.

Step 2 (S2): Generate Safety Artefacts. During a risk analysis, objects are generated from investigations and observations. These objects are collectively known as safety artefacts. Example of safety artefacts include list of hazards, risk mitigation plans, outstanding action items and residual risk report.

Step 3 (S3): Integrate with System Engineering Process. The generated safety artefacts will be feedback to the system engineering process for further cause of actions if they affect the system lifecycle. For example, additional flight trial may have to be conducted after discovering more fight-critical hazards in the system of interest. More people and resources may need to be activated to conduct the flight trial as part of ongoing system engineering process that is beyond the scope and control of the stakeholders conducting the safety assessment.

Example. For example, in the RSAF, a safety assessment technique that is conducted as part of developing the safety assessment report is the zonal safety analysis. A process flow showing the zonal safety analysis is shown in Figure 48. According to the standard SAE ARP 4761 [7], the risk analysis in the zonal safety analysis broadly includes the following three steps: prepare design guideline, inspect installation and inspect for system/item interference. Through these steps, safety artefacts such as measurement results and propose corrective actions are generated. These artefacts would then be feedback to the system engineering process where necessary. For example, modification may need to be carried out due to certain interference between components. This would have to be considered in conjunction with other system engineering activities. The result from the zonal safety analysis will be attached as part of the safety case for the aircraft.

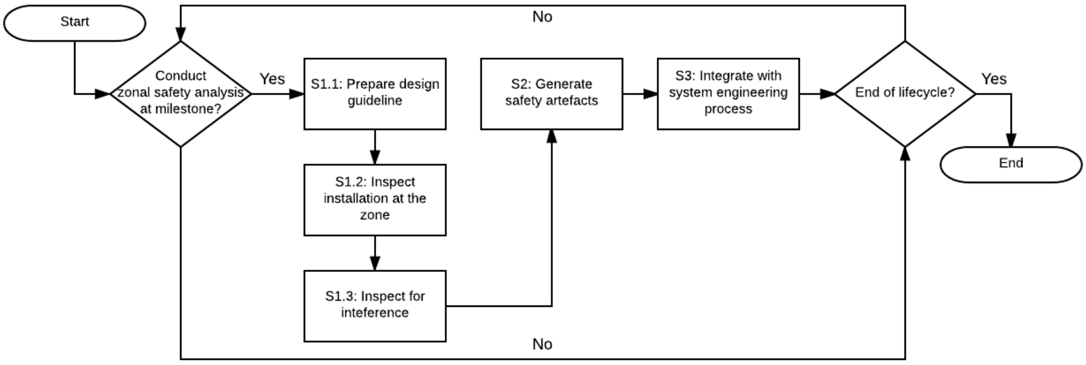

Start
Conduct zonal safety analysis at milestone?
Yes
No
S1.1: Prepare design guideline
S1.2: Inspect installation at the zone
S1.3: Inspect for inteference
S2: Generate safety artefacts
S3: Integrate with system engineering process
End of lifecycle?
Yes
No
End

Figure 48. Process Flow for Zonal Safety Analysis

4.3.1.2. Generic Process of Integrating T.A.G. Approach in Safety Assessment

We have described the process of integrating the T.A.G. approach in risk analysis at the introduction of this chapter. In this section, we would provide a more detailed process flow to show this integration (see Figure 49). The highlighted portion represents the additional elements that have been added as part of the T.A.G. approach, as compared to the generic process of through-life safety assessment presented in Figure 47.

Step 1 (S1): Perform Risk Analysis. The usual risk analysis per existing safety assessment technique. In addition, T.A.G. approach is integrated here by inserting the following sub steps:

> **Step 1a (S1a): Identify Epistemic Uncertainty**. This is Step 1 of the T.A.G. approach. Stakeholders can use the HOT-PIE taxonomy as a reference (see section 4.2) to identify both known and unknown epistemic uncertainties. Stakeholders can also make use of the CMSS and CRMSA (see chapter 3) to help in finding uncertainties in the safety causal models, system models and the model types in the safety assessment.
>
> For example, there may be individuals during the safety assessment who are familiar with the system human resource model regarding the workforce. During safety assessment, such HR experts may be prompted by the '*Human'* primary causal condition under the HOT-PIE taxonomy to bring up concerns about the expertise and experience of the workforce (i.e. uncertainty in the system model). This can help to identify plausible-but-uncertain causal relationships related to the workforce. Using the CRMSA, the safety analysts can describe this plausible-but-uncertain causal relationship explicitly and manage it as part of the T.A.G. approach.
>
> **Step 1b (S1b): Document Epistemic Uncertainty**. This is part of Step 2 of the T.A.G. approach. Any epistemic uncertainty that has been recognised, regardless if it

is assessed to be safety critical, should be tagged and stored in a T.A.G. database. This database is needed in the next step for prioritisation.

Step 1c (S1c): Prioritise Uncertainty Analyses. The T.A.G. database will document causal relationships with epistemic uncertainties that may potentially be safety critical. We refer to them as plausible-but-uncertain causal relationships. Stakeholders are not able to commit if such plausible-but-uncertain causal relationships are safety critical due to a lack of information at the point of conducting the risk analysis. They would also have to assess if the risk is at least tolerable for these uncertainties to be tracked further as explained in section 4.1.2. Due to potential lack of resources, stakeholders may need to prioritise these uncertainties for further tracking. The factors to consider when prioritising the tracking of these uncertainties will be elaborated in section 4.3.2.

Step 2 (S2): Generate Safety Artefacts. Besides generating safety artefacts from safety analyses (e.g. list of hazards, risk mitigation plans, outstanding action items and residual risk report), a database of prioritised uncertainties from step 1c will also be generated. All this information should be integrated into the same existing reporting channel and transferred to the system engineering process in the next step.

Step 3 (S3): Integrate with System Engineering Process. Besides the usual course of actions due to the risk analysis, the T.A.G. approach would also generate action plans to track and address epistemic uncertainties. The requirements from such action plans would be feedback to the system engineering process for further actions.

Example. For example, back to the running example of the RSAF safety assessment, a process flow that shows how T.A.G. approach could be integrated with the zonal safety analysis is shown in Figure 50. As shown in the figure, uncertainty can be recognised in each of the three steps of the zonal safety analysis. Hence, uncertainty tagging can occur in any of the three steps when uncertain causal relationships are discovered. After completing the third step (i.e. inspect for interference), the T.A.G. database shall be consolidated with other safety artefacts from the zonal analysis.

We will revisit the process flows in later chapters where the T.A.G. approach is integrated with safety assessment techniques under the system (see chapter 5) and component (see chapter 6) viewpoints.

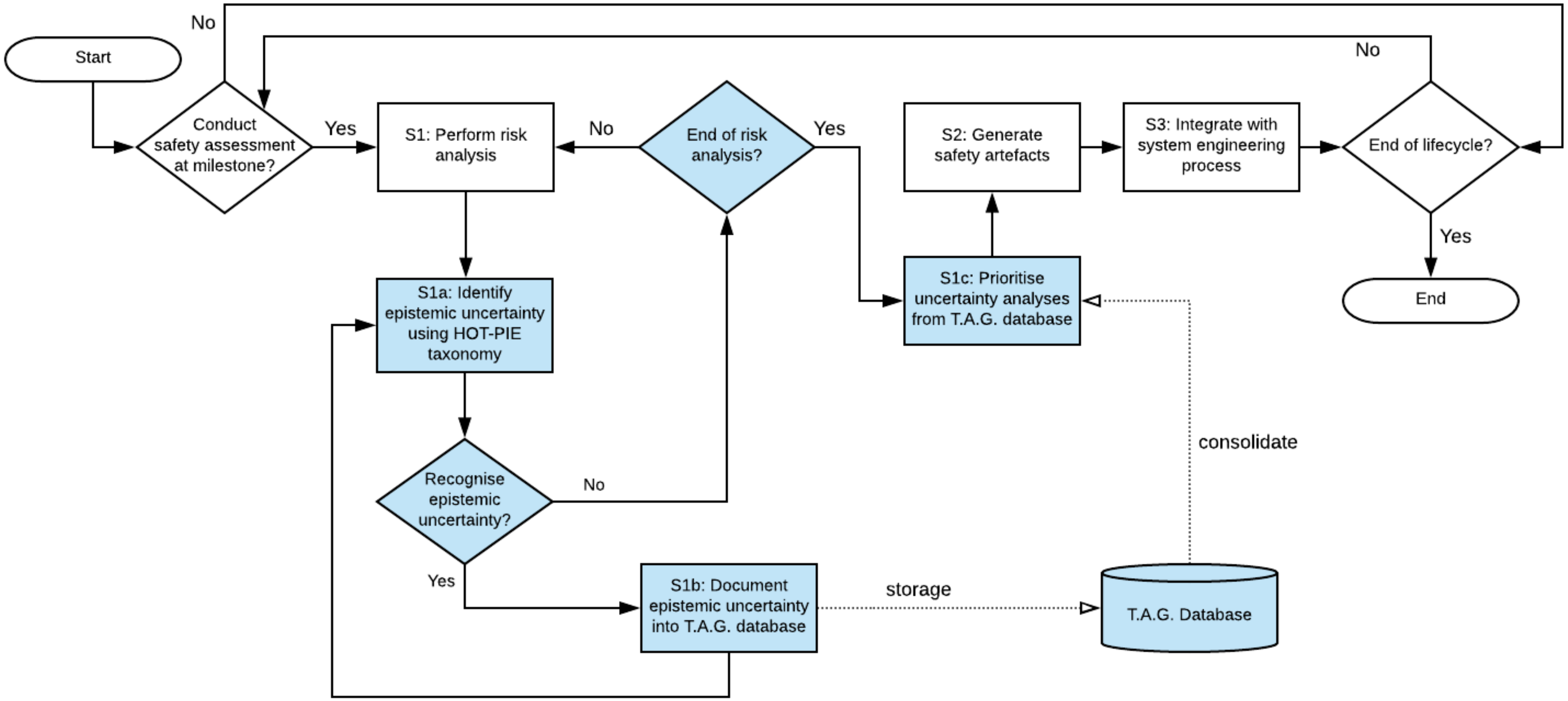

Start
No
Conduct safety assessment at milestone?
Yes
S1: Perform risk analysis
No
End of risk analysis?
Yes
S2: Generate safety artefacts
S3: Integrate with system engineering process
No
End of lifecycle?
Yes
End
S1a: Identify epistemic uncertainty using HOT-PIE taxonomy
S1c: Prioritise uncertainty analyses from T.A.G. database
consolidate
Recognise epistemic uncertainty?
No
Yes
S1b: Document epistemic uncertainty into T.A.G. database
storage
T.A.G. Database

Figure 49. Process Flow to Integrate T.A.G. Approach in Safety Assessment

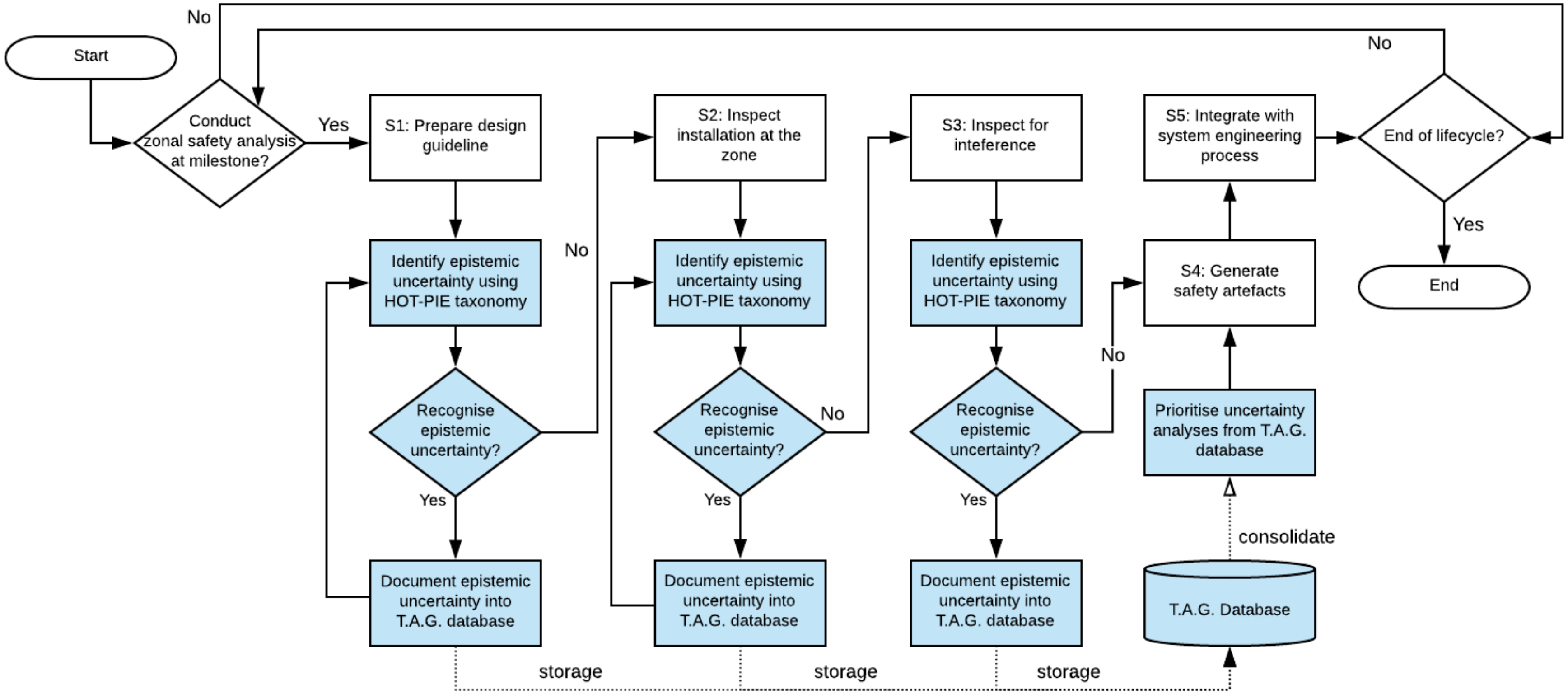

Start
No
Conduct zonal safety analysis at milestone?
Yes
S1: Prepare design guideline
S2: Inspect installation at the zone
S3: Inspect for inteference
S5: Integrate with system engineering process
No
End of lifecycle?
Yes
End
Identify epistemic uncertainty using HOT-PIE taxonomy
No
Identify epistemic uncertainty using HOT-PIE taxonomy
Identify epistemic uncertainty using HOT-PIE taxonomy
S4: Generate safety artefacts
No
Recognise epistemic uncertainty?
Recognise epistemic uncertainty?
No
Recognise epistemic uncertainty?
Prioritise uncertainty analyses from T.A.G. database
Yes
Yes
Yes
consolidate
Document epistemic uncertainty into T.A.G. database
Document epistemic uncertainty into T.A.G. database
Document epistemic uncertainty into T.A.G. database
T.A.G. Database
storage
storage
storage

Figure 50. Process Flow to Integrate T.A.G. Approach in Zonal Safety Analysis

4.3.2. Factors to Prioritise Uncertainty Analyses

Ideally, stakeholders would want to analyse all identified plausible-but-uncertain causal relationships from the safety assessments. As mentioned, this is most likely not possible in practice due to limitation of resources such as time and manpower. Hence, stakeholders need a way to prioritise the list of causal relationships identified from the first step of the T.A.G. approach.

To help stakeholders prioritise the analyses to be conducted on the plausible-but-uncertain causal relationships without overly complicating the existing process, we aim to develop a set of questions for stakeholders to consider during prioritisation based on some factors. For that, we have adopted the Goal Question Metric (GQM) approach [23] to create a set of guided questions and matrix for stakeholders to consider when prioritising the uncertainty analyses. The approach provides a structured way to develop factors (i.e. the metric) to help in our goal of prioritising the epistemic uncertainty identified from initiative one. While the approach is initially designed for software measurement, it can be easily adopted for our research. To summarise, the GQM approach comprises the following three steps:

1. **Goal**. Define the goal that specifies the purpose of the measurement
2. **Question**. Refine the goal into questions to determine factors that break down the issues into components
3. **Metric**. Develop a measurement method as a metric to answer the questions.

It is important to realise that such prioritisation inadvertently involves bias opinions among the stakeholders that are conducting the ranking. Hence, it is important to document as much as possible such opinions that influence the stakeholders' decisions.

Plausible-but-uncertain causal relationships that are of lower priority will be given less attention as compared to causal relationships that are of higher priority. A more proactive action plans may be expected for those relationships that are judged to be of higher priority. Such proactive action would involve coming out with "monitoring technique" and "monitoring activity" to track the uncertainty and its potential effect on the causal relationship. This is akin to developing a hazard tracking system for the high-priority uncertainty.

For example, there could be uncertainty about operational profile that a system can function safely in. A predetermined "monitoring technique" to monitor this uncertainty can be the 'flight trial' and the "monitoring activity" to monitor this uncertainty is the 'opportunity to conduct the trial beyond a certain flight boundary when operating overseas'. Both "monitoring technique" and "monitoring activity" will be tracked as a pair in the proactive action plan.

We will proceed to describe the three steps in the GQM approach to develop the guided questions for prioritising uncertainty analyses.

4.3.2.1. Goal of Prioritisation

Our goal is to develop factors to help users compare and select the analyses to find out more information regarding the identified epistemic uncertainties. These analyses are based on the uncertainties that are identified in the first step of the T.A.G. approach.

Example. In the CHALLENGER example, after identifying the uncertainty in section 4.2.3, one analysis could be to investigate the uncertainty about the plausible design issue with the O-ring at different temperature.

4.3.2.2. Factors for Prioritisation

In our approach, we have recommended priority ranking of uncertainties as the preferred approach. However, we acknowledge that there could be other ways to rank uncertainties. One other approach of ranking uncertainties is to be explicit about how much an uncertainty may affect more than one causal relationships in the safety assessment. An uncertainty that affects multiple causal relationships can be exhibiting more dependencies and as such could be ranked as a higher priority to be managed.

We base our prioritisation on the literature survey of the characteristics of epistemic uncertainty (see section 2.1.2.1). These characteristics are summarised in Table 23.

Table 23. Literature Survey on the Characteristics of Uncertainties

Author	**Observations (details in section 2.1.2)**
D. H. McQuiston, Johnston & Bonoma, Kirsch & Kutschker	Novelty, complexity and importance affect participation and influence behaviour
Peng et al, Novak & Eppinger	3 project complexity dimensions based on OIPT: project size, project novelty, task interdependence.
Tatikonda & Rosenthal	Project varies along two dimensions: technology novelty and project complexity
Shenhar & Dvir	Consider four project characteristics to identify and manage risk: "novelty, technology, complexity and pace,"
Svejvig & Anderson	Rethinking project management based on the following features: Learnability, multiplicity, temporality, complexity, uncertainty and sociability

Author	Observations (details in section 2.1.2)
Saunders et al	Determinants of uncertainty based on: Environmental, individual, complexity, information, temporal and capability
Galbraith, Teuteberg	Organisation information processing model – the greater the task uncertainty, the greater the amount of information that must be processed to achieve a given level of performance

From Table 23, we derive two key factors to consider when prioritising the plausible-but-uncertain causal relationships to follow up: criticality and the expected effort (see Figure 51). By considering these factors, stakeholders would be able to make a better assessment on which uncertainties to analyse first. Generally, the more critical an uncertainty is to cause harm, the higher the priority to manage it first. As for the expected effort, it may be tempted to focus on uncertainty that requires the least effort. However, it can be dangerous if stakeholders decide to avoid uncertainties that require a lot of expected efforts.

For example, an uncertainty that is critical but requires a lot of expected effort may imply that stakeholders need to cater for more resources to analyse the uncertainty, and not the discard it since it can potentially cause a lot of harm. Hence, how the expected effort can affect the prioritisation depends on the context surrounding the uncertainty.

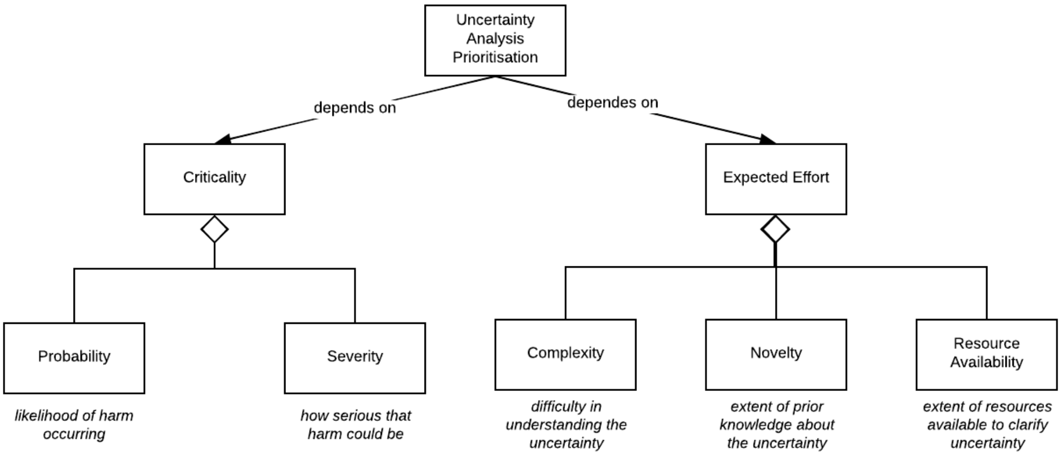

Uncertainty Analysis Prioritisation
depends on
dependes on
Criticality
Expected Effort
Probability
Severity
Complexity
Novelty
Resource Availability
likelihood of harm occurring
how serious that harm could be
difficulty in understanding the uncertainty
extent of prior knowledge about the uncertainty
extent of resources available to clarify uncertainty

Figure 51. Factors to Prioritise Uncertainty Analyses

Next, we will describe criticality and expected effort in detail.

Factors to Measure Criticality. In safety, the criticality of an uncertainty analysis is measured by its level of risk. In section 2.2.1, safety risk is defined as a product of the likelihood of harm occurring and how serious that harm could be. Similarly, the criticality of an uncertainty analysis can be determined by the probability of occurrence and severity of any hazard due to the plausible-but-uncertain causal relationship. Logically, an uncertainty analysis that is of higher criticality should be given higher priority to clarify its epistemic uncertainty.

Example. For example, during system development, higher priority may be given to clarifying an uncertainty concerning the technical functions of a weapon system compared to the uncertainty in the number of engineers to conduct a flight trial. The technical functions have direct consequences on the hazard of the weapon to cause harm, whereas the manpower to conduct the flight trial would most likely affect the efficiency of the trial and not so much of a safety risk.

Stakeholders need to be aware that uncertainty can also affect their assessment of criticality. In practice, the actual criticality of the uncertainty analysis depends on the real risk, which may differ from the estimated risk. This is because the real risk depends on known uncertainties, as well as unknown uncertainties (which we may not even be aware) that affect the probability and/or severity of a causal relationship. It is possible that the real risk could be higher than the estimated risk. Hence, the T.A.G. approach is important to find out more about known and unknown uncertainties such that the stakeholders can make a more informed assessment about the real risk.

Factors to Measure Expected Effort. The expected effort depends on the amount of additional knowledge needed to close the gap between the current state of uncertainty and the threshold of acceptable uncertainty. This is determined by the complexity, novelty and resource availability of the context associated with the uncertainty.

- **Complexity**. Complexity refers to the difficulty in understanding the uncertainties in a condition. The harder to appreciate the uncertainties in a condition (e.g. developing a system in a complicated safety-critical system), the more complex is the condition. Hence, if all else remains equal, the condition that is more complex tends to need more effort to clarify the uncertainties.

  From the literature survey, there seems to be multiple definitions of complexity in project management. In his assessment of uncertainty, McQuiston [37] refers to complexity as '*how much information the organisation must gather to make an accurate evaluation of the system*'. Shenhar [44] states that complexity is about finding out how complicated are the product, the process and the project involved. Tatikonda [43] defines project complexity as '*the nature, quantity and magnitude of organisational subtask and subtask interactions posed by the project*'.

  Peng [39] uses the Organisational Information Processing Theory (OIPT) to conclude that information processing needs are affected by complexity issues such as product size and task interdependence. Product size refers to the number of parts in the product design and task interdependence looks at the influence of any given task on other tasks. Tatikonda, on the other hand, defines three attributes under project complexity, namely project difficulty (level of task performed), objective novelty (novelty of task objectives) and technology interdependence (interdependency of task units). While Tatikonda classifies novelty as a component under project complexity, there are other literature that elevate novelty as one of the main causal determinant besides complexity.

We have constructed our guided questions (see Table 25) about complexity per the three uncertainty elements defined in the CMSS (see section 3.3.1): safety causal model, system model and model types. Under each uncertainty element, the extent of complexity is assessed by factors such as size, difficulties and interdependency.

- **Novelty**. Novelty refers to the extent of prior knowledge about a condition. The lesser the prior knowledge or experience (e.g. developing a new system to operate in a new environment), the more novel is the condition. Since stakeholders cannot depend too much on prior knowledge, condition that is more novel tend to need more efforts to clarify the uncertainties.

  Novelty is defined by McQuiston as '*the lack of experience of individuals in the organisation with similar situations.*'. Shenhar considers novelty as '*how intensely new are crucial aspects of the project*'. Tatikonda refers to technology novelty as the '*newness to the development organisation of the technologies employed in the product and process development effort'*. Peng considers project novelty to involve novelty of product or process, lack of information about markets and customers and the ambiguity of project goals. Based on these definitions, we have selected the 'newness' of the target and the extent of prior experience to be the key factors to assess novelty.

  Like complexity, we have also constructed the guided questions (see Table 25) about 'newness' based on the three uncertainty elements defined in the CMSS (see section 3.3.1): safety causal model, system model and model types. Under each uncertainty element, the extent of novelty is assessed by factors such as system novelty, process novelty and objective novelty.

  We have also included a fourth category, experience, to emphasise the influence of prior experience. A group of stakeholders with prior experience of managing similar systems may be able to provide more information as compared to a group of less experienced stakeholders. However, having experience in other similar systems may not always equate to being more knowledgeable about the current system of interest. Further analysis would still have to be conducted in the context of the current system.

- **Resource Availability**. Resource availability refers to the extent of resources available to understand a condition. It is more difficult to solicit information if there are no dedicated resources to help stakeholders to learn more about the uncertainties. We consider resource availability as having the capability (e.g. skills, infrastructure, manpower) and time to solicit information about the uncertainties.

  Having the capability is not enough if there is no time to do the learning. This has been stated by many researchers, such as Shenhar, Svejvig [45] and Saunders [46]. For example, in system safety lifecycle, stakeholders must meet numerous deadline. They may not have enough time to focus on all the uncertainties that have been discovered due to a lack of time. Hence, urgency is an important factor under resource availability.

Example. For example, in the RSAF safety assessment example from section 4.2, we have discovered a series of plausible uncertainties. There may be a need to prioritise which uncertainties to analyse first due to resource constraint in developing the safety assessment report. Besides considering the criticality, the safety analysts should also consider the expected efforts to analyse the uncertainty. The analysis for this example will continue in the next section of this chapter.

4.3.2.3. Matrix to Calculate the Priority

Based on the factors of criticality and expected efforts, we have created a list of guided questions for stakeholders to prioritise uncertainties for tracking (see Table 25). Stakeholders can use these questions to prioritise and choose which of the plausible-but-uncertain causal relationships to focus on in the third step of the T.A.G. approach.

To make a more objective decision, we introduce an overall priority score based on the two factors of criticality and expected effort. The following instructions would help stakeholders to calculate the priority score:

1. Identify the uncertainty in the targeted causal relationship.
2. Analyse the criticality of the target based on the guided questions. Score the criticality from 0 (low criticality) to 1 (high criticality).
3. Analyse the complexity of the target based on the guided questions. Score the complexity from 0 (low complexity) to 1 (high complexity).

4. Analyse the novelty of the target based on the guided questions. Score the novelty from 0 (low novelty) to 1 (high novelty).

5. Analyse the resource availability of the target based on the guided questions. Score the resource availability from 0 (high resource availability) to 1 (low resource availability).

6. Calculate the overall expected effort by average out the sum of the three scores under complexity, novelty and resource availability (assuming the factors have equal weightage).

7. Locate the quadrant on the prioritisation matrix (see Figure 52) that corresponds to the score for criticality and expected effort.

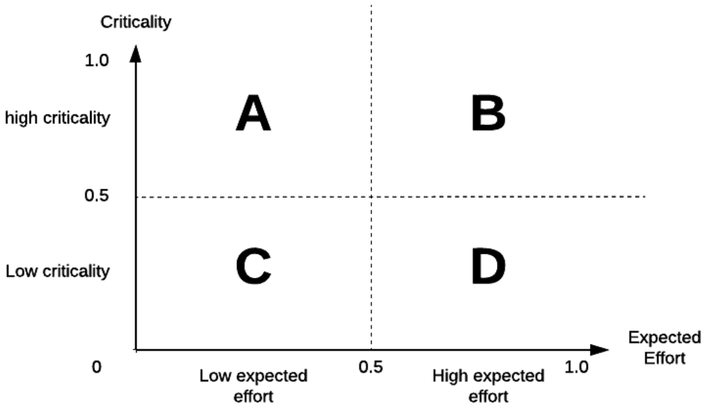


Figure 52. Prioritisation Matrix for Uncertainty Analysis

Example. Using the examples in section 4.2, the expected efforts to analyse the uncertainty can be measured quantitatively by average out the scores (between 0 and 1) for each of the factors: complexity, novelty and resource availability for each uncertainty. Analysis of an uncertainty that is more complex, more novel and less learnable would be given a higher score. For instance, we have compared between three of the uncertainties that consider Human, Technology and Information causal condition in the HOT-PIE taxonomy (see Table 24). Note that, while these three examples belong to different case studies, we have artificially analysed them together as a form of comparison on the expected efforts.

Table 24. Example of Expected Efforts Calculation to Analyse Uncertainties

Example	A	B	C
HOT-PIE Primary Condition	Human	Technology	Information
Uncertainty	uncertainty about the expertise of the SMEs in carrying out the safety assessment (section 4.2.1)	uncertainty about the design concern of the O-ring performance in low temperature (section 4.2.3)	uncertainty about compliance with standards by subcontractors during modification of the aircraft (section 4.2.5)

Example	A	B	C
HOT-PIE Primary Condition	Human	Technology	Information
Complexity	0.3	0.6	0.8
Novelty	0.3	0.7	0.4
Resource Availability	0.4	0.6	0.8
Expected Effort	0.3	0.6	0.7

Complexity. Comparing the three examples, the most complex analysis would be to find out the level of compliance among the subcontractors in example C. The efforts to find out the manufacturers of components and subcomponents, and subsequently deciding on which standards they need to comply with, can be complicated. Hence it is given a score of 0.8. Relatively, analysing the O-ring for example B would be less complex than example C. Hence it is given a complexity value of 0.6. Since it is quite straight forward to find out the experience level of the SMEs in example A, it is considered a relatively simpler task than the other two with a complexity score of 0.3.

Novelty. In terms of novelty, the task of analysing previous data for design concern in the O-ring might be something that the engineers have little experience in. Hence, example B is given the highest novelty score of 0.7. Relatively, communicating with contractors and subcontracts in example C, as well as SMEs in example A, are activities that stakeholders would most likely have done before. Hence, both example A and C have lower scores for novelty compared to example B.

Resource availability. The hardest uncertainty to find out more information would be example C as it involves multiple third parties (contractors and subcontractors) outside the team that is conducting the safety analysis. Hence, it is given a resource availability score of 0.8. Example B requires the team to retrieve past design and operational data that should be readily available. Finding if the SMEs have the necessary expertise in example A is also relatively easier compare to the other two examples.

Expected effort. Averaging out the scores for complexity, novelty and resource availability, it is observed that analysing the expertise of the SMEs in example A requires the least effort at a score of 0.3. This is followed by analysing the O-ring design in example B which has a score of 0.6. The task of checking the compliance of contractors and subcontractors in example C is expected to require the most effort out of the three examples.

The prioritisation matrix for the three examples is plotted in Figure 53 using the calculated scores of criticality and expected efforts. Stakeholders can use it to prioritise the order in which the three uncertainties will be managed.

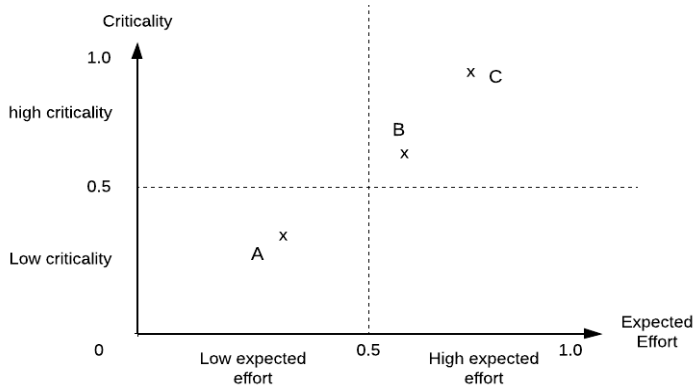


Figure 53. Prioritisation Matrix Example

Since prioritisation is subjective, stakeholders can be making decisions that are bias and based on certain mindsets. Hence, it is important for stakeholders to be explicit about the strategies that have been considered during prioritisation. Stakeholders should also be mindful that it can be a challenging task to prioritise the causal relationships as it is highly dependent on the information available.

While we cannot guarantee that two safety analysts could produce consistent results, the guided questions in the T.A.G. approach are introduced to provide some level of consistency by having a structured way for the safety analysts to consider uncertainty. While it is feasible, applying an average of the overall priority score (based on the factors of criticality and expected effort) from each safety analyst may produce skewed and misleading results when it comes to prioritisation. An alternate approach could be to conduct qualitative discussion and sharing among the safety analysts that are using the T.A.G. approach to build consensus in selecting the more important uncertainties to focus on first. This will be an area we will focus on to improve the T.A.G. approach in the future.

Example. As an example, stakeholders may decide to prioritise the uncertainty analyses based first on the criticality of the uncertainties and secondly by the amount of efforts needed. This is summarised in the prioritisation matrix in Figure 54.

With this strategy, stakeholders would decide to put high priorities to analyses in quadrants A and B, rather than C and D (see Figure 52). Take note that this is but one of the many ways that stakeholders can prioritise the uncertainty analyses. The final choice of where to focus on depends on the context facing the stakeholders during the analysis.

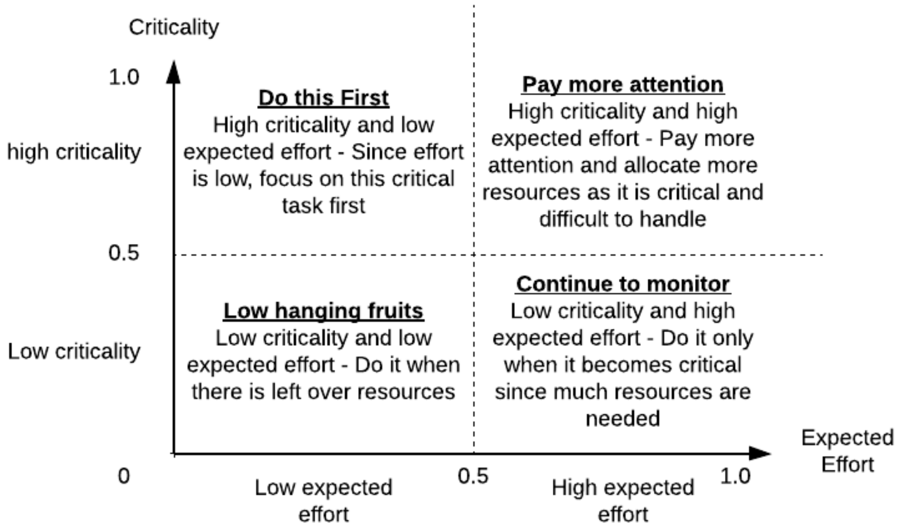

Criticality
1.0
high criticality
0.5
Low criticality
0
Do this First
High criticality and low expected effort - Since effort is low, focus on this critical task first
Pay more attention
High criticality and high expected effort - Pay more attention and allocate more resources as it is critical and difficult to handle
Low hanging fruits
Low criticality and low expected effort - Do it when there is left over resources
Continue to monitor
Low criticality and high expected effort - Do it only when it becomes critical since much resources are needed
Expected Effort
Low expected effort
0.5
High expected effort
1.0

Figure 54. An Example of Prioritisation for Uncertainty Analysis

To summarise, we have developed an approach to prioritise the plausible-but-uncertain causal relationships based on a set of guided questions (see Table 25) and a prioritisation matrix chart (see Figure 52). Stakeholders can make use of this approach to prioritise the uncertainty analyses according to the quadrants in the prioritisation matrix.

Table 25. Guided Questions to Prioritise Uncertainty Analyses

Target Uncertainty to be analysed: < Describe the targeted uncertainty here >	
Factors	**Questions**
Criticality: Assess the criticality of the target to affect system safety from the stakeholders' perspectives.	
Probability	Q1. How likely will this target cause harm to the system?
Severity	Q2. How serious are the consequences if this target occurs?
Complexity: Assess the complexity of the system related to the target from the stakeholders' perspectives.	
Safety Causal Model	Q3. How complex are the safety causal models from where the target is identified? (e.g. no. of related causal conditions, hazard interdependency, residual risk interdependency)
System Model	Q4: How complex are the system models from where target is identified? (e.g. system structure / product size / product design, process / task interdependency, process design, technology interdependency, project management difficulties)
Model Type	Q5: How complex are the model types from where target is identified? (e.g. patterns, templates and conventions used by the safety and system models)
Novelty: Assess the novelty of knowledge related to the target from the stakeholders' perspectives.	
Safety Novelty	Q6: How much of past information was used to define the safety causal model from where the target is identified? (e.g. causal factors, hazard list)
System Novelty	Q7: How much of past information was used to define the system model from where the target is identified? (e.g. system legacy, process, technology, objective)
Model Type Novelty	Q8: How much of past information was used to define the model type from where the target is identified? (e.g. patterns, templates, conventions)
Experience	Q9: Are there other prior experiences related to the target?
Resource Availability: Assess the resource availability related to the target from the stakeholders' perspectives.	
Capability Urgency	Q10: Do we have the capability to learn the knowledge needed? (e.g. manpower, skills, infrastructure and support, environment, other resources) Q11: Do we have the time required to learn the knowledge needed? (e.g. project deadline, lifecycle milestones)

4.4. Initiative 3: A Structured Method to Track Action Plan and Address Goal

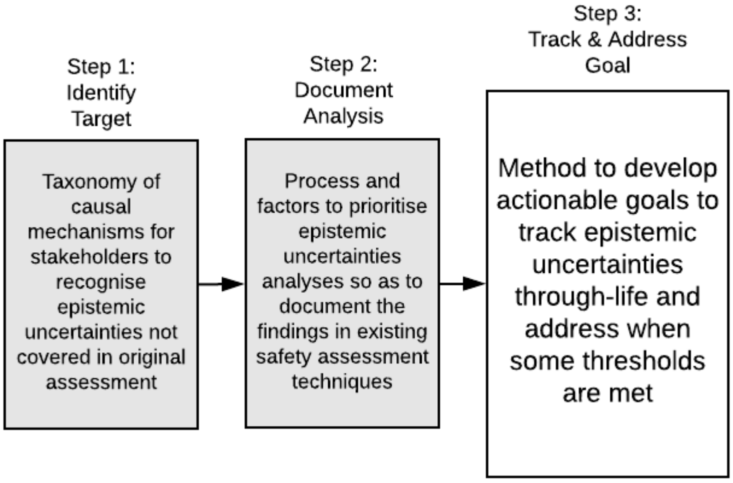

Step 1:
Identify
Target
Step 2:
Document
Analysis
Step 3:
Track & Address
Goal
Taxonomy of causal mechanisms for stakeholders to recognise epistemic uncertainties not covered in original assessment
Process and factors to prioritise epistemic uncertainties analyses so as to document the findings in existing safety assessment techniques
Method to develop actionable goals to track epistemic uncertainties through-life and address when some thresholds are met

Figure 55. T.A.G. Approach Third Step: Track and Address Goal

To support step 3 of the T.A.G. approach, we implement a method to develop actionable goals that stakeholders can implement as part of the system engineering process (see Figure 55). The aim is to track plausible-but-uncertain causal relationship through-life till certain thresholds are met to prompt the stakeholders to address the uncertainty. The portion of the process where T.A.G. approach step 3 takes place is shown shaded in Figure 56.

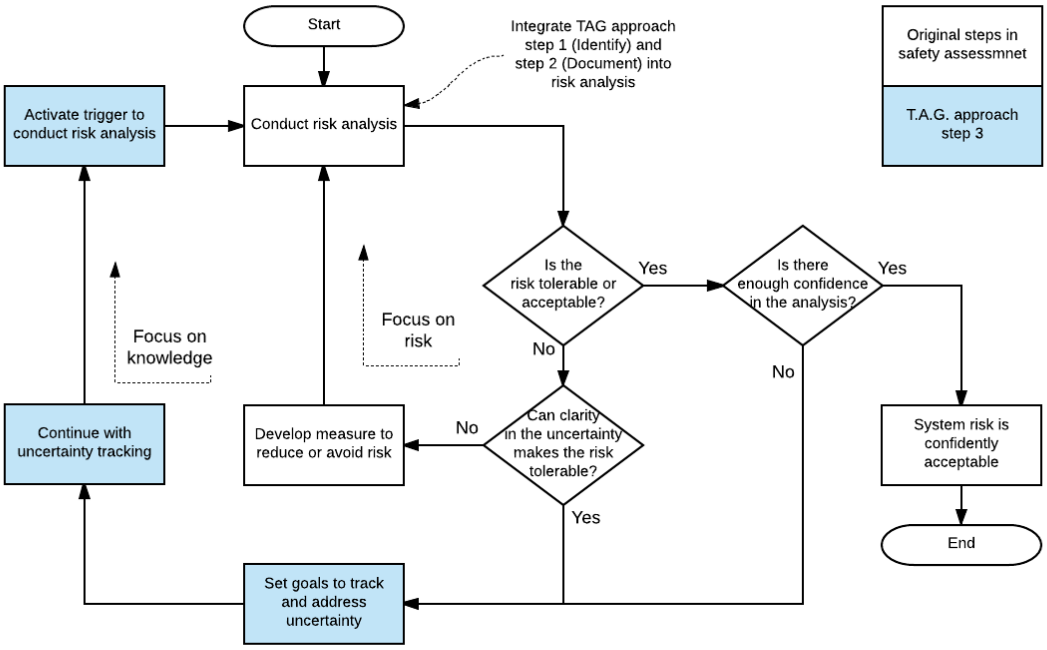

Start
Integrate TAG approach step 1 (Identify) and step 2 (Document) into risk analysis
Original steps in safety assessmnet
T.A.G. approach step 3
Activate trigger to conduct risk analysis
Conduct risk analysis
Is the risk tolerable or acceptable?
Yes
No
Is there enough confidence in the analysis?
Yes
No
Focus on knowledge
Focus on risk
Can clarity in the uncertainty makes the risk tolerable?
No
Yes
Continue with uncertainty tracking
Develop measure to reduce or avoid risk
System risk is confidently acceptable
End
Set goals to track and address uncertainty

Figure 56. T.A.G. Approach Step 3: Track and Address Goal

The third principle of managing epistemic uncertainties states that epistemic uncertainties in the underlying models of safety assessment should be tracked and addressed in a systematic manner through-life (see section 3.2.3). The demand to track and address the uncertainty systematically is further elaborated here:

- **Demand of a Systematic Plan to Track the Uncertainty**. We have established in section 3.2.3.1 that system safety engineering demands a systematic and through-life management of knowledge. In order not to leave it to chance, tracking of epistemic uncertainties should be planned and this plan should follow a structured approach that spells out the activities to be conducted as part of the tracking. To do that, we have developed a guide for stakeholders to develop a systematic action plan (see section 4.4.1). The factors to be considered under the action plan is covered in section 4.4.2.
- **Demand of a Systematic Plan to Address the Uncertainty**. As mentioned in section 3.2.3.2, epistemic uncertainties can change with time. Stakeholders need to be prepared when such changes occur. To help in being responsive to such changes, the plan to manage the epistemic uncertainties should include the actions needed when certain thresholds with regards to the uncertainties are met. For example, the plan could specify the right people and resources that should be available to address an uncertainty. The factors in the action plan to address the uncertainty is elaborated in section 4.4.2. Some of the possible outcomes when addressing epistemic uncertainty are described in section 4.4.3.

4.4.1. The GQTA Approach

Like the prioritisation effort earlier, we also develop a set of questions for stakeholders to consider when constructing the action plan to manage epistemic uncertainties. However, instead of directly using the GQM approach, we have modified the approach to emphasis the two areas of tracking and addressing of uncertainties. To do that, we have dropped the ‘M’ (i.e. matric) in the GQM approach and added the ‘T’ and ‘A’ to represent ‘Track’ and ‘Address’ respectively. This is known as the GQTA approach, which stands for Goal Question Track and Address approach. The following summarises the four steps under the GQTA approach:

1. **Goal** – From the prioritisation based on criticality and expected effort, stakeholders would have selected a list of plausible-but-uncertain causal relationships. The goal is to manage such uncertainty conditions throughout the system safety lifecycle proactively with an action plan.

Example. After deciding to analyse an uncertainty in section 4.3 (assuming it has been given high priority), the following goal is to develop the plan that comprises two parts: 1) to decide on the monitoring technique that tracks the plausible design flaws in the system after multiple modifications, and 2) to decide on the monitoring activity that addresses the design flaws when certain thresholds are met.

2. **Question** – Refine the goal into questions that break down the issues to prepare for tracking and addressing. Since system lifecycle and the epistemic uncertainties can evolve quickly with time, stakeholders may not have a lot of time to construct a comprehensive action plan to track and address the targeted epistemic uncertainties. We have proposed important areas to focus on when deriving the action plan and these are described in section 4.4.2.

3. **Track** – Develop the tracking guide based on the questions from the previous step. We have suggested four focus areas to help the stakeholders develop the action plan. This is described in section 4.4.2.

4. **Address** – Like tracking, the focus areas to address the uncertainty is also described in section 4.4.2. Besides having an action plan, it is important to consider the possible outcomes when called upon to address the uncertainties. This is discussed in section 4.4.3.

4.4.2. Factors to Consider in Action Plan

We have surveyed literature in section 2.1.2.2 on ways to track and address uncertainties and this is summarised in Table 26.

Table 26. Literature Survey of Ways to Track and Address Uncertainties

Author	Approach (details in section 2.1.2)
Eisenhardt & martin, Helfat, Zollo & Winter	Dynamic capabilities theory – resources and capabilities must be constantly reallocated and reoptimized to adapt to changing environment
Weick	Decomposition of organizing – enactment, selection and retention.
Teece	Conceive the concept of dynamic capabilities as the ability to sense, seize and adapt to exploit competences and address changing environment
Petit & Hobbs	Conceptual framework to study management of uncertainty in project portfolios
Saunders et al	Present three approaches (structural, behavioural and relational) to contend with uncertainties in project life and describe ways that practitioners can identify, analyse and act on it
Diana	Apply sense and respond in risk and security domain – risk signal that can be captured (sense) provide an opportunity to mitigate risk (respond). Sense and respond methods involve monitoring and observing, incessant rehearsing, understanding and interpreting data, deciding how to respond, and producing a response
Gothelf and Seiden	Identify five most important principles underpinning the sense and respond approach

From the literature surveyed, we observe that most of the models can be broadly grouped under two distinct phases: track and address. This supports our third principle of managing epistemic uncertainties – '*epistemic uncertainties should be tracked and addressed in a systematic manner through-life*'. Correspondingly, our action plan will be constructed based on these two phases. In addition, we have also based our action plan on the conceptual framework to study project uncertainty by Petit [50] as it recommends a range of core activities for sensemaking. As a result, we have combined our principles to track and address with the conceptual framework by Petit to derive four focus areas. These focus areas are highlighted in Figure 57.

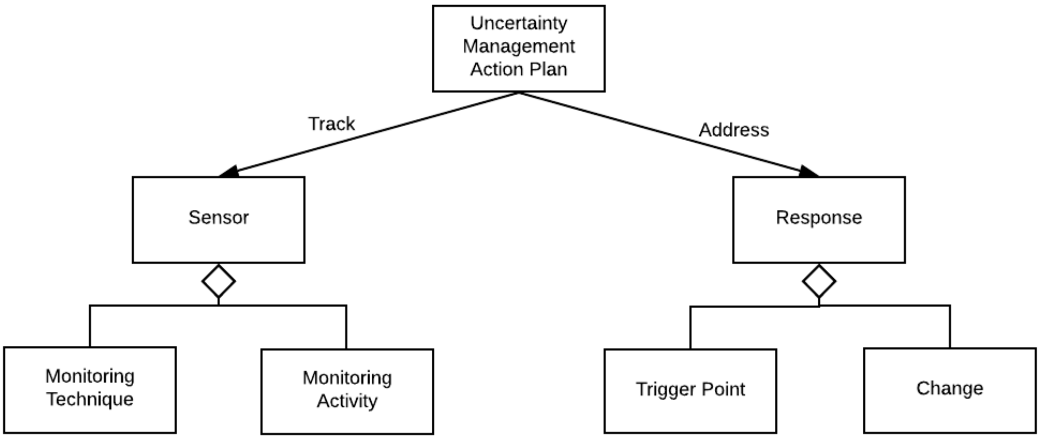


Figure 57. Focus Areas in Uncertainty Management Action Plan

An action plan should document the approach to 1) track sensors and 2) address responses that come from the sensing. For tracking, the stakeholders must decide on the type of monitoring technique to be constructed and the monitoring activity needed for the sensing task. To address the responses well, the stakeholders must set trigger points to make decision and readily adapt when there are any changes needed to keep the system safe. These four focus areas are elaborated in this section.

4.4.2.1. Decide Monitoring Technique

The first area focuses on deciding on the monitoring technique needed to track the epistemic uncertainties. In this case, we use the term ‘monitoring technique’ to describe any means to detect changes related to the epistemic uncertainty. Examples of monitoring technique include machines (e.g. optical sensors) to detect certain physical property and people (e.g. operators or maintenance engineers) to observe certain phenomena. The key questions for considerations are:

- What are the monitoring techniques needed to track the epistemic uncertainties?
- How are the monitoring techniques able to track the epistemic uncertainties?

Example. In the CHALLENGER example, to track the plausible design concern of the O-ring in low temperature, the monitoring technique could be the people and their tasks are to review the design and operation data. They would be given the responsibility of sensing if sufficient information is available to review the uncertainty over the design concern.

4.4.2.2. Develop Monitoring Activity

The second area focuses on developing the monitoring activity to track the uncertainties using the monitoring technique. There can be many factors to consider here. The key questions to consider include:

- What types of monitoring technique are to be taken?
- When should these monitoring activities be collected from the techniques?
- Who is responsible to do the tracking?
- What are the structures and supporting resources to put in place for the tracking?
- What are the skills, experiences and attitudes needed to do the tracking?

Example. In the CHALLENGER example, to track the plausible design concern in the O-ring, the people tasked to consolidate the design and operation data would need to decide when sufficient material has been consolidated. They would also need to be experienced in soliciting the right documents that are needed for the analysis. They should report regularly to the safety committee that oversees the safety of the space shuttle.

4.4.2.3. Set Trigger Points

The third area focuses on the level of preparedness to make decision due to some trigger points. Such trigger points can be periodic (e.g. presenting result during monthly or quarterly safety review board), or due to any change of state or event that goes beyond a certain predetermined threshold (e.g. in the CHALLENGER example, a reason to address the design concern of the O-ring would be the discovery of plausible hazard when carrying out the monitoring activity). The key questions for considerations are:

- What are the trigger points that require proactive response?
- Who are responsible to decide on the respond actions?
- What are the governance and supporting structure to put in place to make the decision?
- What are the skills, experiences and attitudes needed to make the decision?

4.4.2.4. Adapt to Change

The fourth area focuses on the readiness to address the uncertainty and make changes to keep the system safe. The key questions for considerations are:

- Who is responsible to review the changes that will be put in place?

- How prepared and responsive should the system be in addressing the uncertainty?
- What are the structure and support resources to put in place to address the uncertainty?
- What are the skills, experiences and attitudes needed to address the uncertainty?

Example. In the CHALLENGER example, the safety committee overseeing the safety of the space shuttle should be responsible for making any recommendation or changes to address the uncertainty. In addition, the committee needs to weight the balance between the level of risk and the need to consolidate more information about the uncertainty. If the system needs to be modified upon the uncertainty analysis, the entire system engineer process would have to be reviewed to assess the impact on the system lifecycle.

4.4.2.5. Planning Guide

The planning guide to develop action plan is consolidated in Table 27. This guide can be used by stakeholders to develop the tracking plan to monitor the uncertainty conditions, as well as the follow up actions needed to address the uncertainties. We have also developed an expanded version of the guide in Annex B to provide more examples to consider for tracking and addressing epistemic uncertainties. This expanded version can be used for training purposes or as reference material during safety assessment where it is practical to do so.

The proposal by Saunders [51] to consider structural, behavioural and relational approaches can be considered in the expended version. Structural issues refer to the structure and processes to manage uncertainty. Behavioural issues refer to the attitude and mental state when managing uncertainty. Relational issues focus on the interpersonal communication to manage uncertainty.

Table 27. Guide to Set Goals for Action Plan

Target Uncertainty: < Describe the targeted uncertainty here >		**Guided Questions to set Goals for Action Plan**
1: **Track Sensors**	1a: **Decide Monitoring Techniques** – What shall we track?	Q1. What are the monitoring techniques needed to track the epistemic uncertainties? Q2. Why are the monitoring technique able to meet the goal of managing the uncertainty condition?
	1b: **Form Monitoring Activity** – How do we track?	Q3. What types of monitoring activity to be taken? Q4. When should these monitoring activity be collected from the techniques? Q5. Who are responsible to do the tracking?

Target Uncertainty: < Describe the targeted uncertainty here >		**Guided Questions to set Goals for Action Plan**
		Q6. What are the structures and supporting resources to put in place for the tracking?
		Q7. What are the skills, experiences and attitudes needed to do the tracking?
2: **Address Responses**	2a: **Set Trigger Points** – When and how shall we decide to respond?	Q8: What are the trigger points from the monitoring activity that require proactive response?
		Q9. Who are responsible to decide on the respond actions?
		Q10. What are the governance and supporting structure to put in place to make the decision?
		Q11. What are the skills, experiences and attitudes needed to make the decision?
	2b: **Adapt to Change** – What are the possible responses?	Q12. Who is responsible to review the changes that will be put in place?
		Q13. How prepared and responsive should the system be in addressing the uncertainty?
		Q14. What are the structure and support resources to put in place to address the uncertainty?
		Q15. What are the skills, experiences and attitudes needed to address the uncertainty?

4.4.3. Possible Outcomes When Addressing Uncertainty

Besides having an action plan to explicitly document how to address the uncertainty, it would help stakeholders to better manage the uncertainties if they can appreciate and anticipate the possible outcomes when they address the uncertainty. Since these outcomes depend largely on how much the residual uncertainty has evolved, we would first present the common ways that an uncertainty may change with time in section 4.4.3.1. This would follow by a description of the possible outcomes when addressing epistemic uncertainty in section 4.4.3.2.

4.4.3.1. How does Residual Uncertainty change with Time?

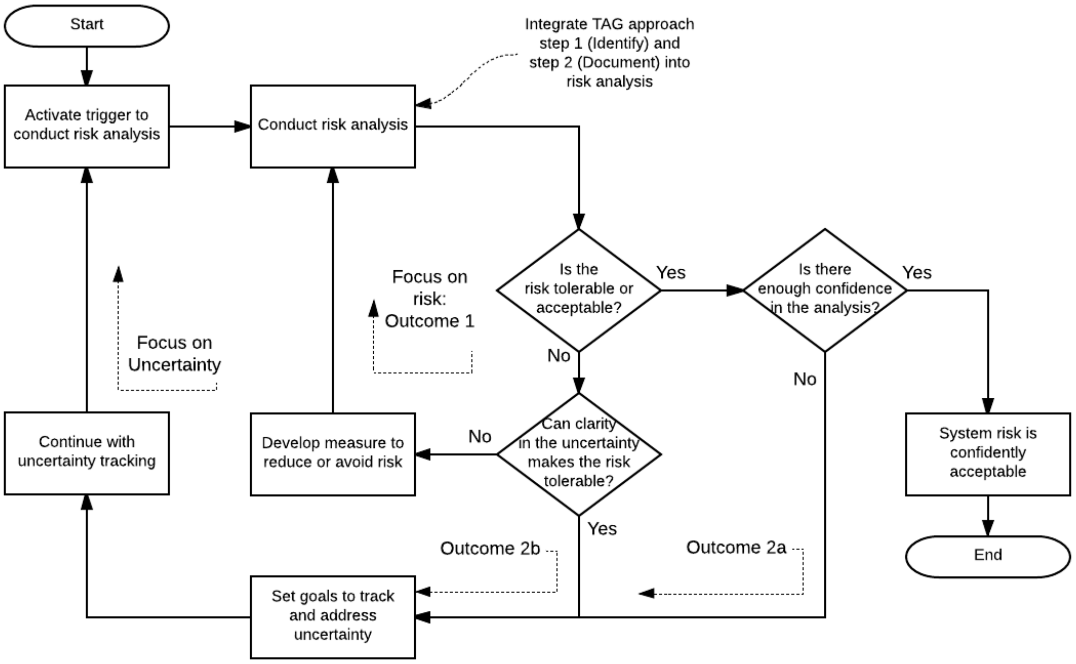
The uncertainty that remains after a safety assessment has been undertaken can be considered as residual uncertainty. Here are three cases in which such residual uncertainty may change with time:

- **Residual Uncertainty Reduces with Greater Clarity**. Some residual uncertainties may become lesser with time, e.g. as experiments produce test results, stakeholders clarify safety concerns with specialists, and new component developers lock down their interface specifications. Such relevant information can allow stakeholders to make better judgement about safety risk, albeit at a later phase of the system lifecycle. Having more information increases confidence but may not directly make a system safer. Whether the additional information makes the system safe or unsafe depends on the follow-on hazard analysis.
- **Residual Uncertainty Creates Other Related Uncertainty**. Some residual uncertainties may evolve when more information is available. For example, when mitigating argument or evidence is presented with regards to an uncertainty, a new residual uncertainty may be created. For example, there could be uncertainty about using a weapon system in a specific operating environment A. Operators may report their experiences of using the same weapon system in another environment B for the same purpose. This may reduce the uncertainty around the usage of the weapon system, but it may result in a new residual uncertainty about the similarity between environment A and B. This can trigger a follow-up effort to derive monitoring technique or activity to compare the differences between both operating environments.
- **Discovery of Previously Unknown Uncertainty**. As residual uncertainties are being tracked under the T.A.G. approach, stakeholders may unexpectedly discover other previously unknown uncertainties. This may trigger a response to address the newly discovered epistemic uncertainty. Since this uncertainty is new, the stakeholders would have to trigger the process to address the risk associated with the uncertainty.

4.4.3.2. What are the Possible Outcomes when Addressing Uncertainty?

When one of the above scenarios happens to the residual uncertainty, it may reach a certain predefined threshold tracked by the T.A.G. approach such that it triggers the stakeholders to address the uncertainty. The generic steps to address uncertainty when triggered is shown in Figure 58. It is similar to Figure 56 at the start of section 4.4, except that the starting point is at the trigger that activates the risk analysis to address the uncertainty.

When being triggered to address the uncertainty, stakeholders would conduct a risk analysis to assess the risk that is still facing the system, as well as the assessing the confidence in the analysis. At this moment, stakeholders would have to decide if the risk is confidently within the tolerable or acceptable region. If it is not, a decision would have to be made to either focus on managing the risk or clarifying the uncertainty. These two outcomes are explained further here.



Start
Integrate TAG approach step 1 (Identify) and step 2 (Document) into risk analysis
Activate trigger to conduct risk analysis
Conduct risk analysis
Focus on risk: Outcome 1
Is the risk tolerable or acceptable?
Yes
Is there enough confidence in the analysis?
Yes
Focus on Uncertainty
No
No
Continue with uncertainty tracking
Develop measure to reduce or avoid risk
No
Can clarity in the uncertainty makes the risk tolerable?
System risk is confidently acceptable
Yes
Outcome 2b
Outcome 2a
End
Set goals to track and address uncertainty

Figure 58. Possible Outcomes when Addressing the Uncertainty

Outcome 1 – Focus on Risk by Implementing System Change. If the level of system risk is assessed to be unacceptable, stakeholders would need to derive strategies to reduce the risk. Possible approaches include modifying the system design and changing the ways that the system is being operated. An example is to derive safety measure to reduce the risk (e.g. putting up speed bumps at an accident-prone junction to force motorist to reduce speed). Another example could be to use risk mitigation action to avoid the risk (e.g. putting up signs to warn motorists of an accident-prone junction). Stakeholders would have to conduct the safety analysis again to conclude if the new measure or mitigation does lower the risk to a tolerable or acceptable level.

Outcome 2 – Focus on Uncertainty by Continuing with the T.A.G Approach. The T.A.G. tracking on the epistemic uncertainty may need to continue either to increase confidence or reduce risk.

- **Outcome 2a – Increasing Confidence.** Regardless of the level of risk, there may still not be enough confidence in the assessment due to the presence of epistemic uncertainties. For example, during hazard analysis, risk assessment was conducted based on the latest system design. The risk may be acceptable but there are outstanding epistemic uncertainties in the system design (e.g. lack of subsystem specifications from subcontractors) that need to be followed up to provide the necessary confident with the risk assessment.

- **Outcome 2b – Reducing Risk.** In this second case, clarifying certain epistemic uncertainties may be needed to make an unacceptable risk tolerable or acceptable.

To end the uncertainty tracking, stakeholders must be satisfied that the epistemic uncertainties being tracked by the T.A.G. approach have either been eliminated or that the risk is assessed to be confidently acceptable despite the decision to stop tracking these known uncertainties. The T.A.G. approach would be applied in multiple cases studies across the next two chapters to demonstrate how it can complement established safety assessment techniques to manage epistemic uncertainties.

4.5. **Initial Evaluation and Summary**

The overall evaluation of this research is discussed in detail under Chapter 7. In this section, we highlight and evaluate specifically the T.A.G. approach discussed in this chapter.

In section 4.2, we have established the HOT-PIE taxonomy of causal mechanism that comprises six primary causal conditions of over three hundred abstract causal mechanisms. These causal mechanisms can be used to prompt stakeholders to either appreciate known uncertainties better or recognise unknown uncertainties that may not be obvious prior embarking on the T.A.G approach. Through the many examples in the section, we have shown that the taxonomy can help stakeholders to identify uncertainties that can lead to hazards not discovered in the original safety assessment.

In section 4.3, we have established the prioritisation process to integrate the T.A.G. approach with existing safety assessment process. A generic process flow to tag uncertainty is developed, which can be customised with existing safety assessment technique. Two major factors to prioritise uncertainty analyses (i.e. criticality and expected efforts) are also established to help stakeholders to decide on which uncertainties to manage. We have shown, with the running examples from the RSAF safety assessment, CHALLENGER and NIMROD case studies, that the prioritisation process can be integrated with existing safety assessment technique such as the zonal analysis. We have also demonstrated how three epistemic uncertainties from the examples can be prioritised by considering the criticality and expected efforts to manage the uncertainties.

In section 4.4, we have established a structured method to construct an action plan to track and address epistemic uncertainties that are derived from the first two steps of the T.A.G. approach. The method advocates four important activities in the action plan: deciding on monitoring technique, forming the monitoring activity, setting trigger points and adapting to change. Again, using the running examples, we have demonstrated the formulation of the action plan to track and address the uncertainty about the plausible design flaw in previous aircraft modifications. This action plan shall provide the stakeholders the means to track and address the uncertainty even when it changes with time.

In order to make the T.A.G. approach more feasible for industrial usage, we have consolidated the three steps, together with the key concepts, into a concise user guide. This is attached in Annex K. This guide is also used in the next two chapters when applying the T.A.G. approach on case studies that involves established safety assessment techniques.

The T.A.G. approach requires the stakeholders to consider epistemic uncertainties in such depth that they would probably have not considered during normal risk analysis, as well as to track and address these uncertainties through life which they may not have done before. This inevitably would 'cost' the stakeholders to commit more resources such as time and manpower into the safety assessment. Like evaluating other safety assessment techniques, it is not easy to conduct a quantitative cost-benefit analysis on implementing the T.A.G. approach. Instead, a qualitative assessment using questionnaire is applied in this research to solicit feedback on the cost-benefit of implementing the T.A.G. approach. This will be further elaborated under the main evaluation in Chapter 7.

Chapter 5 – Application of T.A.G. Approach in Safety Assessments from System Viewpoints

5.1. Introduction

With the understanding of the T.A.G. approach from chapter 4 and the relationships between system and system elements under CMSS (see Figure 28), we will next evaluate the feasibility of integrating the approach with existing safety assessment techniques. To cover a broad range of safety assessment techniques, our evaluation will focus on safety assessments conducted at both system and component viewpoints. In this chapter, we apply the T.A.G. approach on the Yongwen railway accident from a top-down perspective by assuming that we do not have hindsight about the facts in the accident. Hence, the uncertainties that will be identified using the T.A.G. approach may or may not be directly related to the actual cause of the railway accident. Facts about the actual accident are used for comparison after we have completed the exercise to identify the uncertainties using the HOT-PIE taxonomy; and are not considered as known information during the T.A.G. approach.

System vs Component Viewpoints. From IEEE 42010, recall that viewpoint, or model type, "*establishes the conventions for constructing, interpreting and analysing the view* (or model) *to address concerns framed by that viewpoint*". Examples of viewpoint conventions include languages, notations, model kind, modelling methods and analysis techniques. Such model type determines the expressive power of the system model. The applications of T.A.G. approach at both viewpoints shall be evaluated in two separate chapters (see Table 28):

Table 28. Summary of Chapters that Apply T.A.G. Approach

Chapter	Viewpoint	Safety Assessment Technique	Referenced Scenario
5	System	Systems-Theoretic Process Analysis	Yongwen railway accident analysis by Song et al [24]
6	Component	Fault Tree Analysis and Failure Modes and Effects Analysis	ARP 4761 aircraft design example [7]

- **Safety Assessments from System Viewpoints**. In this chapter, chapter 5, we focus on safety assessments from system viewpoints. We integrate the T.A.G. approach with the Systems-Theoretic Process Analysis (STPA) hazard analysis technique. This is based on the STPA analysis on the Yongwen railway system by Song et al. The HOT-PIE taxonomy is applied on the existing STPA risk analysis to assess the application of the T.A.G. approach to discover plausible-but-uncertain causal relationships.

- **Safety Assessments under Component Viewpoints**. In the next chapter, chapter 6, we focus on safety assessments from component viewpoints. We integrate the T.A.G. approach with the Fault Tree Analysis (FTA) and Failure Modes and Effects Analysis

(FMEA) techniques. This is based on the aircraft design example in ARP 4761. Unlike system viewpoint that consider broader system issues such as technology and processes, component viewpoints create models that are more targeted such as electrical circuit diagram or physical design specifications. The challenge is to be aware of the scope and limitations of each model, as well as appreciate the relationships between models that potentially can influence each other and become safety critical (e.g. when reviewing an electrical circuit diagram, the safety analyst may assume that the air-condition is well regulated and overheating of the electrical components is not of safety concern). We will elaborate more about safety assessments from component viewpoints in the next chapter.

Accident Analysis vs Hazard Analysis. In this chapter, we shall integrate the T.A.G. approach with the STPA safety assessment technique from the system viewpoint. As shown in Table 28, we have chosen to base our study on the Yongwen railway accident analysis by Song et al as they are able to demonstrate the important steps in STPA using a case study. In his research, Song conducts an inductive hazard analysis to develop safety requirements to manage hazards that are discovered using the STPA technique. As analysing the Yongwen railway accident can be retrospective based on hindsight and facts of the accident, Song's hazard analysis has injected a sense of artificiality in order to demonstrate that the STPA technique can indeed discover hazards. These hazards could either be directly related to the actual accident or additional safety risk beyond the actual causes of the accident.

Similarly, when we apply the T.A.G. approach on the Yongwen railway accident, we approach the analysis by assuming that we do not have hindsight about the facts of the accident. The facts from the accident can be used for comparison but should not be considered as known information during the safety assessment.

Why do we choose STPA? To evaluate safety assessment from the system perspective, we need a safety assessment technique that focuses on using models that are based mostly on system viewpoints. The STPA is one of such techniques. The STPA uses abstract safety causal models that highlight interactions between different systems, as well as unsafe system causal relationships in the form of control loops and constraints. This allows the elements in the STPA to be integrated with the CRMSA model (introduced in section 3.3.2). We will show in this chapter that the risk analyses conducted in the STPA can be easily represented using the CRMSA model and this helps to integrate the T.A.G. approach into the analysis.

While the STPA examines hazards from the system viewpoints, the system issues that are being analysed are mostly bounded once the system control structure is finalised in the early stages of the risk analysis. In addition, there isn’t systematic process in STPA to highlight structural linkages or causal relationships that are plausible but uncertain due to lack of information. This provides the opportunity to use the T.A.G. approach to complement STPA to identify more safety-critical system issues with the use of the HOT-PIE taxonomy, as well as track the epistemic uncertainties beyond the STPA process, throughout the system lifecycle till they are being addressed when information is available.

Summary of the Yongwen Railway Accident. On July 23, 2011, A high-speed train from Beijing to Fuzhou (train no: D301) collided into the back of another train from Hangzhou to Fuzhou (train no: D3115) along the Yongwen railway line at 20:30 China Standard Time (CST) near the suburbs of Wenzhou. The two trains collided at the track circuit 5829AG, causing 40 fatalities and 172 injuries. The sequence of events leading to the accident is described in Annex C.

While the direct cause of the collision was due to the error in the track signalling system, it was accepted that there are other contributing factors to the accident that would have existed in the complicated railway system. The accident represents an appropriate example of a complicated social-technical system with numerous subsystems that are constantly interacting with each other, which can potentially be safety-critical.

Moving forward, we will first describe the STPA on the railway system conducted by Song et al in section 5.2. With this understanding, we will next evaluate how epistemic uncertainties can be managed using the T.A.G. approach together with STPA in section 5.3. We will conclude the chapter by evaluating the contributions of the T.A.G. approach in managing uncertainties together with STPA in section 5.4.

5.2. **Safety Assessment using STPA**

According to chapter 2.2.1, Systems-Theoretic Process Analysis (STPA) is a hazard analysis technique based on Systems-Theoretic Accident Model and Processes (STAMP). To refresh, STAMP is an accident causality model that is based on the study of systems that are interdependent on each other. It is applied in safety assessment by considering factors such as software, human, technology and organisation. These are factors that directly or indirectly affect the safety of social-technical system. According to Leveson [66], the basic activity in STAMP is to find constraints in the system. In STAMP model, the cause of accident is considered as “*the result of a lack of constraints imposed on the system design and on operations*”.

STPA uses control loop to identify constraints that, when violated, can result in unsafe situations. Using the generic process of safety assessment from section 4.3, we construct a similar process flow for STPA (see Figure 59). The figure highlights the two unique steps in the STPA technique that focus on risk analysis (step 1a and step 1b). These steps are described in detail under the STPA primer [182].

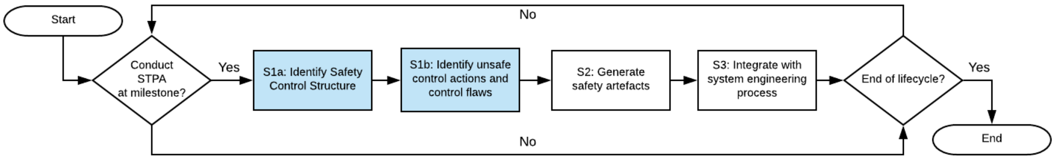

Start
Conduct STPA at milestone?
Yes
S1a: Identify Safety Control Structure
S1b: Identify unsafe control actions and control flaws
S2: Generate safety artefacts
S3: Integrate with system engineering process
End of lifecycle?
Yes
End
No
No

Figure 59. Process Flow for STPA Safety Assessment

Step 1a (S1a): Identify Safety Control Structure. This is the first risk analysis to be conducted in STPA. It focuses on constructing the hierarchical safety control structure (SCS) related to the hazard and the constraints necessary to control the hazard.

Step 1b (S1b): Identify Unsafe Control Actions and Control Flaws. This is the second risk analysis to be conducted in STPA. It focuses on identifying unsafe control actions (UCAs) and the root causes, known as control flaws (CFs), for the UCAs by looking for any failures and dysfunctional interactions.

Step 2 (S2): Generate Safety Artefacts and Requirements. Generate and document the safety artefacts and requirements that will be integrated with the larger system engineering process.

Step 3 (S3): Integrate with System Engineering Process. The generated safety artefacts will be feedback to the system engineering process for further cause of actions if they affect the system lifecycle.

For the rest of this section, we will describe the SCS, UCAs and CFs that have been highlighted by Song et al during their STPA analysis on the Yongwen railway accident.

5.2.1. Yongwen Railway Safety Control Structure

The SCS depends on how the system of interest is being modelled during the safety assessment. It is largely derived from the system models that are available during the STPA assessment. There are 3 steps in constructing the SCS:

1. Identify main components
2. Determine who controls who and what (control and feedback)
3. Add details of constraints such as component responsibilities and process models

The railway SCS related to the accident according to Song et al is reproduced here in Figure 60. Based on the SCS, safety constraints are identified during the safety assessment (see Table 29**Error! Reference source not found.**). The subsequent analysis focuses on determining the UCAs and CFs that can lead to hazards that are known as '*inadequate control processes to maintain the safety constraints*'.

Table 29. Safety Constraints for Railways System Components

Railway system components	Safety constraints
Train control centre	• Acquire correct and fresh data about train position, speed and occupation • Control passing signals in block sections and send moving authorities to ATP correctly based on acquired data
ATP on board	• Control operation of train according to signals provided by ground system • Prevent train from entering the block section occupied by another train
Train dispatcher	• Sending control commands to railway stations and drivers • Monitoring the operation of trains and the occupation of tracks
Drivers	• Operate trains safely according to operation procedures and control commands issued by train dispatcher • Report operation information and problems to dispatcher and railway stations
Ministry of railway	• Make regulations and standards on safe operation of trains • Provide oversights on the execution of regulations and standards • Implement technical review and certification of equipment
Shanghai railway bureau	• Enforce safety regulations and working standards in its railways • Provide adequate education and training to its staff • Oversee railway operation and report incidents and accidents to ministry

As it is not possible to analysis all the railway system components for the rest of our evaluation, we will narrow the focus on the interactions between train control centre (TCC) equipment and the ATP on the train (i.e. the first two components in **Error! Reference source not found.**). We will use these two components to evaluate the process of integrating the T.A.G. approach into STPA.

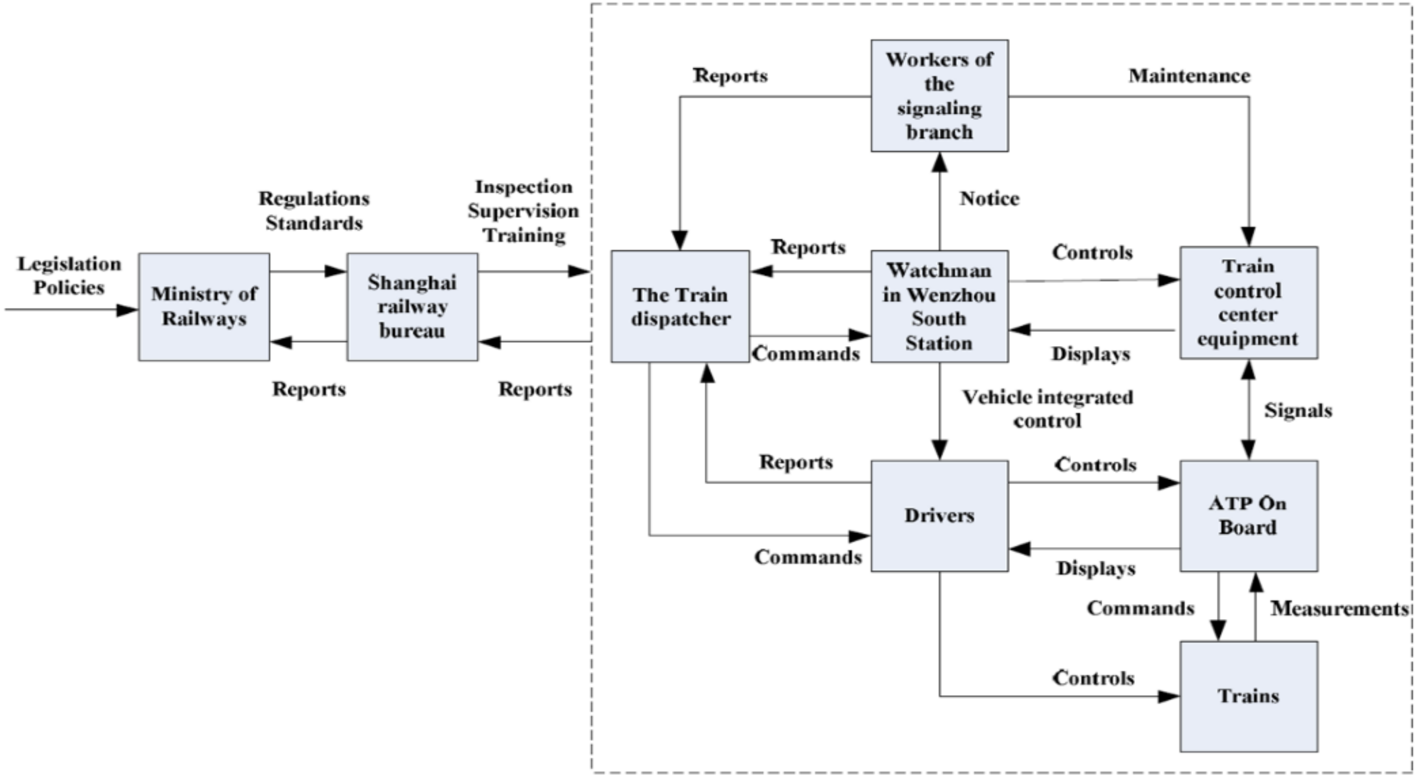


Figure 60. Basic Railway Safety Control Structure in the Yongwen Accident

5.2.2. Unsafe Control Actions between TCC and ATP

In the STPA analysis, the UCAs are identified by looking for causal conditions that can violate the safety constraints. A summary of the UCAs identified by Song et al for both the train control centre and on-board ATP are shown in Table 30.

Table 30. Unsafe Control Actions for the TCC and ATP

Train control centre	On-board ATP
1) Train CC didn't get information that 5829AG section was occupied by D3115 2) Passing signal in the 5829AG section was wrong 3) Send moving authorities to the ATP on D301 while the section was occupied by D3115	1) ATP on D3115 stopped train near faulted 5829AG track section 2) ATP on D301 didn't take any action to prevent train from entering block section occupied by D3115

Example. For example, the UCA "*passing signal in the 5829AG section was wrong*" can cause the TCC safety constraint "*acquire correct and fresh data about train position, speed and occupation*" to be violated, while the UCA "*ATP on D301 didn't take any action to prevent train from entering block section occupied by D3115*" can cause the ATP safety constraint "*prevent train from entering the block section occupied by another train*" to be violated.

5.2.3. Control Flaws leading to the Unsafe Control Actions

After identifying the UCAs, the CFs are identified by looking for causal conditions that can cause the UCAs. A summary of the CFs identified by Song et al for both the train control centre and on-board ATP are shown in Table 31.

Table 31. Control Flaws for the TCC and ATP

Train control centre	On-board ATP
False track occupation information was provided to the computer due to component failure	Code sent to ATP on D3115 by 5829 track circuit was abnormal
Design of equipment could not ensure correctness and freshness of data in face of a thunder strike and component failure	Process model of the ATP on D301 was inconsistent with actual process

Example. For example, the UF "*false track occupation information was provided to the computer due to component failure*" can cause the TCC UCA "*passing signal in the 5829AG section was wrong*", while the UF "*code sent to ATP on D3115 by 5829 track circuit was abnormal*" can cause the ATP UCA "*ATP on D3115 stopped train near faulted 5829AG track section*".

5.3. Integrate T.A.G. approach into STPA

To evaluate the application of the T.A.G. approach, we will continue to focus on the causal relationships between train control centre and ATP on the train from the previous section. We will first describe the process flow of tagging uncertainties as part of the STPA (section 5.3.1). Next, we follow the process flow to complement STPA with the three steps in T.A.G. approach described in chapter 4:

- Step 1. Identify uncertainties using HOT-PIE taxonomy. This is conducted for both the SCS (section 5.3.2) and the UCAs and CFs (section 5.3.3).
- Step 2. Prioritise the uncertainty analyses based on the factors introduced in the T.A.G approach (section 5.3.4), and
- Step 3. Generate track and address action plans using the guided questions introduced in the T.A.G. approach (section 5.3.5).

5.3.1. Process of Tagging Uncertainties in STPA

Using the generic process of tagging uncertainties in safety assessment from section 4.3, we construct a similar process flow to tag epistemic uncertainties in STPA (see Figure 61). To emphasise what has been mentioned earlier in section 3.3, the T.A.G. approach is not a separate standalone method to conduct safety assessment but rather it introduces additional steps to complement the existing safety assessment techniques. This is again evidenced in Figure 61, which shows how the T.A.G. approach is integrated with the existing process flow for STPA. For the rest of this section, we will describe in detail how epistemic uncertainties during STPA analysis for the Yongwen railway system can be identified, documented, tracked and addressed by using the T.A.G. approach.

5.3.2. Step 1a: Identify and Tag Uncertainties from SCS

According to Figure 61, the first step of the T.A.G. approach is to tag uncertainties that are discovered while identifying the SCS. This is conducted with the help of the HOT-PIE taxonomy (a copy of the taxonomy can be found in the T.A.G. approach User Guide in Annex K). In this section, we evaluate the use of the taxonomy to identify epistemic uncertainties related to the SCS. Uncertainty Observation 1 is about new linkage among nodes with existing causal relationship (see section 5.3.2.1) and Uncertainty Observation 2 is about linkage between nodes that do not have prior relationships with each other (see section 5.3.2.2). The differences between the two have been discussed in section 4.2.7.

5.3.2.1. Uncertainty Observation 1: Fatigue among Workers

From the SCS in Figure 60, it is observed that there is a causal relationship whereby "workers maintain the TCC equipment". This is an existing causal relationship between the worker and the TCC equipment. We use the CRMSA model to represent this causal relationship at the system domain in Figure 62. Here, we can search for uncertainty in linkages between nodes.

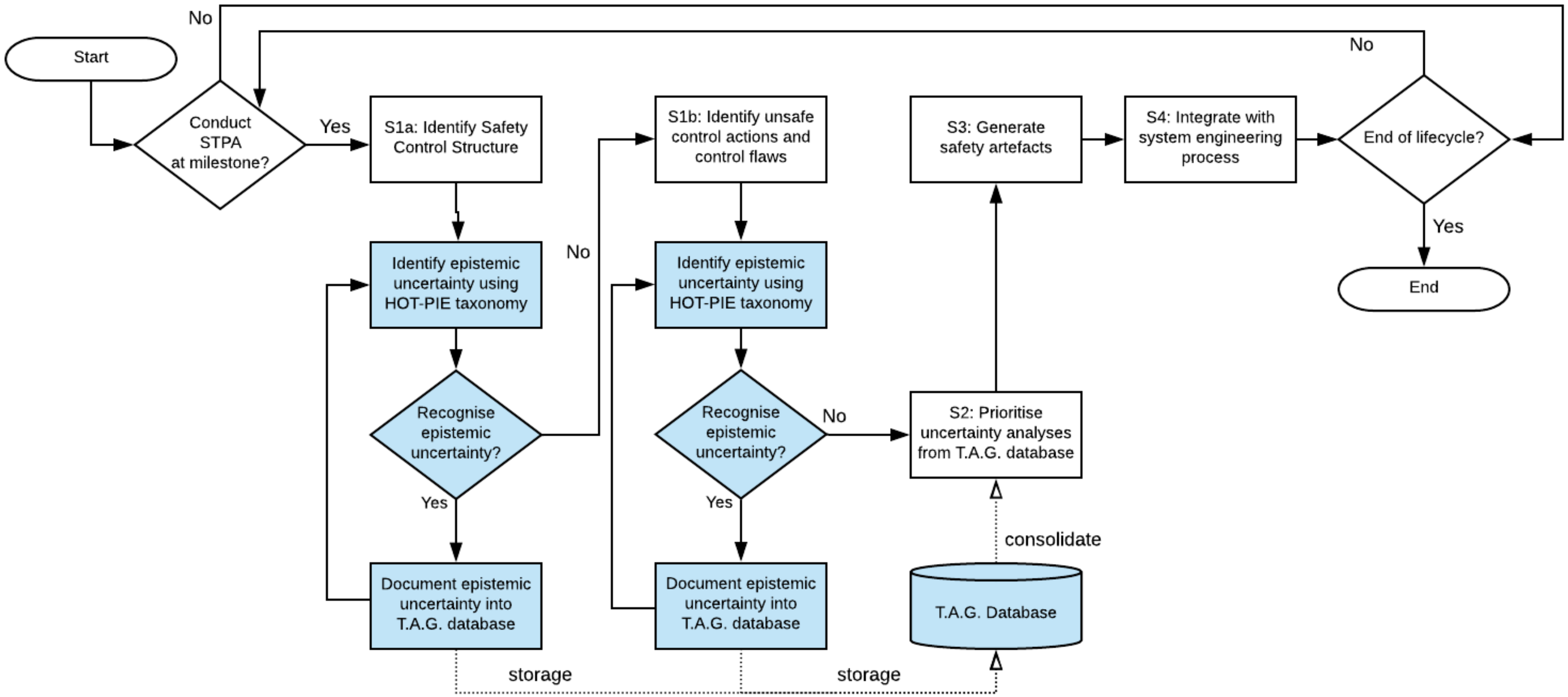

Start
Conduct STPA at milestone?
No
Yes
S1a: Identify Safety Control Structure
Identify epistemic uncertainty using HOT-PIE taxonomy
Recognise epistemic uncertainty?
Yes
No
Document epistemic uncertainty into T.A.G. database
S1b: Identify unsafe control actions and control flaws
Identify epistemic uncertainty using HOT-PIE taxonomy
Recognise epistemic uncertainty?
Yes
No
Document epistemic uncertainty into T.A.G. database
S2: Prioritise uncertainty analyses from T.A.G. database
consolidate
T.A.G. Database
storage
storage
S3: Generate safety artefacts
S4: Integrate with system engineering process
End of lifecycle?
No
Yes
End

Figure 61. Process Flow to Tag Epistemic Uncertainties in STPA Safety Assessment

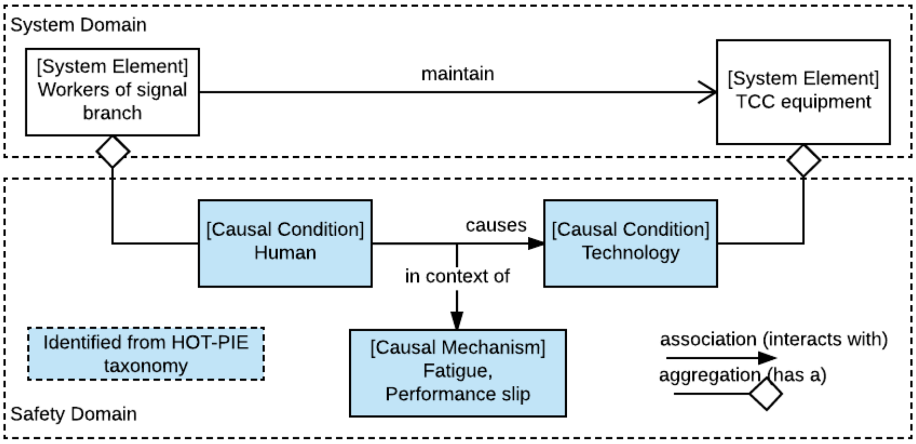

System Domain
[System Element] Workers of signal branch
maintain
[System Element] TCC equipment
[Causal Condition] Human
causes
[Causal Condition] Technology
in context of
Identified from HOT-PIE taxonomy
[Causal Mechanism] Fatigue, Performance slip
association (interacts with)
aggregation (has a)
Safety Domain

Figure 62. Uncertainty about Fatigue among Workers

We refer to the HOT-PIE taxonomy in the user guide to recognise plausible causal relationships. Searching through the *Human* primary condition for the workers and the *Technology* primary condition for the TCC equipment, we found two relevant plausible-but-uncertain causal mechanisms:

- Fatigue (under secondary casual condition H2: Mental State), and
- Performance slip (under secondary casual condition H3: Action).

Example. We choose *Fatigue* as a plausible-but-uncertain causal mechanism. Hence, a plausible-but-uncertain causal relationship could be related to the fatigue level of the workers when maintaining the TCC equipment. Pertinent questions to consider include:

1. Are the workers getting sufficient rest prior to any maintenance or repair job on the railway system?

2. Is there an established system to monitor the fatigue level of the workers during maintenance?

3. Are there criteria in place to stop work when working condition is undesirable such as due to bad weather or prolong working hours?

5.3.2.2. Uncertainty Observation 2: Infrastructure Support by Shanghai Railway Bureau

From the SCS in Figure 60, we have observed that there is no prior system linkage between the Shanghai railway bureau and the TCC equipment. However, with the HOT-PIE taxonomy, we discover new causal relationship between both system components (see Figure 63). Here, we search for uncertainty in the linkage between nodes with no prior identified causal relationships.

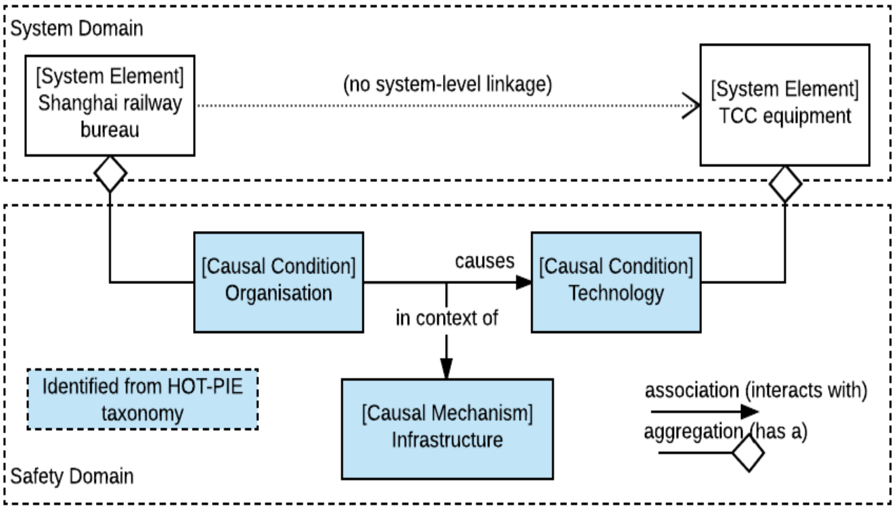

System Domain
[System Element] Shanghai railway bureau
(no system-level linkage)
[System Element] TCC equipment
[Causal Condition] Organisation
causes
[Causal Condition] Technology
in context of
Identified from HOT-PIE taxonomy
[Causal Mechanism] Infrastructure
association (interacts with)
aggregation (has a)
Safety Domain

Figure 63. Uncertainty about Infrastructure Support by Shanghai Railway Bureau

We refer to the HOT-PIE taxonomy in the user guide to recognise plausible causal relationships. Searching through the *Organisation* primary condition for the Shanghai railway bureau and the *Technology* primary condition for the TCC equipment, we found a relevant plausible-but-uncertain causal mechanism:

- Interdependent infrastructure (under secondary casual condition O3: Resource).

Example. We choose *Interdependent Infrastructure* as a plausible-but-uncertain causal mechanism between Shanghai railway Bureau and the TCC equipment. Pertinent questions to consider include:

1. What are the roles and responsibilities of the Shanghai railway bureau in ensuring reliable communication from the TCC equipment?

2. Is the technical infrastructure support provided by the Shanghai railway bureau to the Wenzhou south station sufficient and timely?

3. How often is the TCC equipment maintained and serviced according to the policies and standards laid down by the Shanghai railway bureau?

5.3.3. Step 1b: Identify and Tag Uncertainties from UCAs and CFs

According to Figure 61, the second step is to tag uncertainties that are discovered while identifying the UCAs and CFs. While the process for tagging uncertainties may look like Step 1, the focus has shifted from system models (i.e. looking at SCSs) to safety causal models (i.e.

looking at UCAs and CFs). This shows that the T.A.G. approach is flexible enough to discover plausible-but-uncertain causal relationships in both system and safety domains.

This is conducted with the help of the HOT-PIE taxonomy in the user guide attached in Annex K. In this section, we evaluate the use of the taxonomy to identify epistemic uncertainties related to the UCAs and CFs between the TCC and ATP on the train. Uncertainty Observation 3 is about new linkage among nodes with existing causal relationship (see section 5.3.3.1) and Uncertainty Observation 4 is about linkage between nodes that do not have prior relationships with each other (see section 5.3.3.2).

5.3.3.1. Uncertainty Observation 3: Software Algorithm during Signal Abnormality

In section 5.2.3, we have highlighted a causal relationship whereby the CF "*false track occupation information was provided to the computer due to component failure*" can cause the TCC UCA "*passing signal in the 5829 section was wrong*". We use the HOT-PIE taxonomy to determine that the causal mechanism for this causal relationship can be classified as a form of electrical property (under secondary casual condition T2: Property). We represent this relationship using the CRMSA in Figure 64. Here, we can search for uncertainty in linkage between existing nodes.

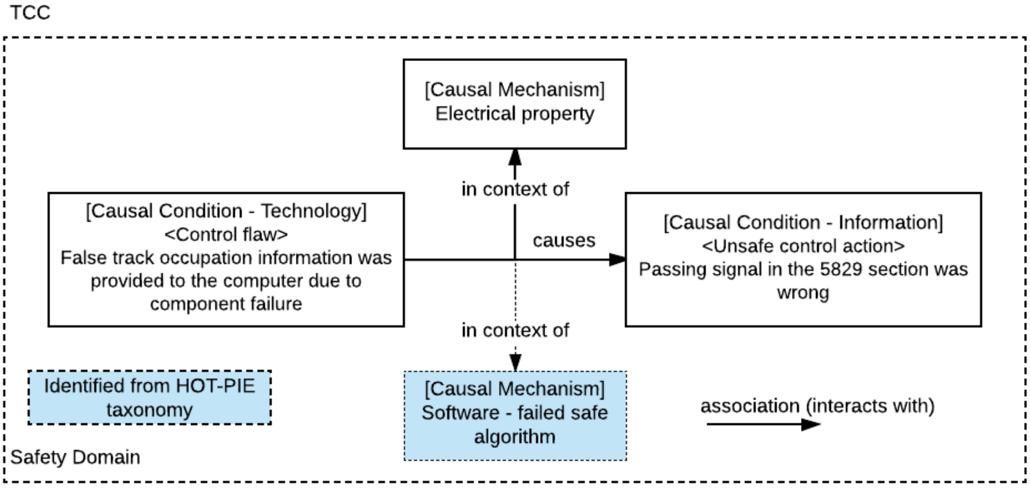

TCC
[Causal Mechanism]
Electrical property
in context of
[Causal Condition - Technology]
<Control flaw>
False track occupation information was provided to the computer due to component failure
causes
[Causal Condition - Information]
<Unsafe control action>
Passing signal in the 5829 section was wrong
in context of
Identified from HOT-PIE taxonomy
[Causal Mechanism]
Software - failed safe algorithm
association (interacts with)
Safety Domain

Figure 64. Uncertainty about Software Algorithm during Signal Abnormality

Next, we refer to the HOT-PIE taxonomy in the user guide to recognise plausible causal relationships. Considering the CF is a *Technology* primary condition and the UCA is an *Information* primary condition, we identify a relevant plausible-but-uncertain causal mechanism using the taxonomy:

- Software issue (under secondary casual condition T1: Machine).

Example. We choose *Software issue* as a plausible-but-uncertain causal mechanism. Pertinent questions to consider include:

1. How much knowledge do the operators and engineers have about the software functions in the train operation?

2. Can the software in the train operation provide a fail-safe algorithm to mitigate any risk due to hardware failure?

3. Can the software in the train operation handle abnormality such as wrong passing signal? If so, how?

5.3.3.2. Uncertainty Observation 4: Human Intervention during Signal Abnormality

In section 5.2.2, we have highlighted the UCAs between the TCC and the ATP on train. However, there is no direct causal relationship established between any UCAs. That does not mean that there are no uncertainties between the UCAs that can be safety-critical. Using the HOT-PIE taxonomy and represent it using the CRMSA model, we aim to discover new causal relationships between the UCAs (see Figure 65). Here, we can search for uncertainty in linkage between new nodes.

In this illustration, we consider the TCC UCA "*Send moving authorities to the ATP on D301 while the section was occupied by D3115*" can potentially cause the ATP UCA "*ATP on D301 didn't take any action to prevent train from entering block section occupied by D3115*".

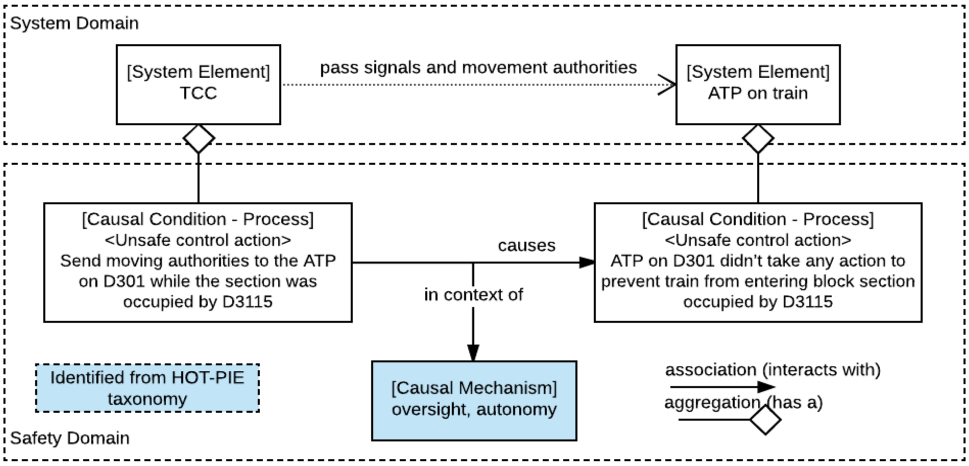

System Domain
[System Element] TCC
pass signals and movement authorities
[System Element] ATP on train
[Causal Condition - Process] <Unsafe control action> Send moving authorities to the ATP on D301 while the section was occupied by D3115
causes
in context of
[Causal Condition - Process] <Unsafe control action> ATP on D301 didn't take any action to prevent train from entering block section occupied by D3115
Identified from HOT-PIE taxonomy
[Causal Mechanism] oversight, autonomy
association (interacts with)
aggregation (has a)
Safety Domain

Figure 65. Uncertainty about Human Intervention during Signal Abnormality

Next, we refer to the HOT-PIE taxonomy in the user guide to recognise plausible causal relationships. Consider both UCAs as *Process* primary causal conditions, we found two relevant plausible-but-uncertain causal mechanism:

- Oversight (under secondary casual condition P1: Nature), and
- Extent of autonomy (under secondary casual condition P1: Nature).

Example. We choose *Oversight* as a plausible-but-uncertain causal mechanism. Pertinent questions to consider include:

1. How much human oversight is there during the train operation, especially to handle abnormality?

2. To what extent can humans intervene in the autonomous train operation?

3. How established is the command and control structure among the staff running the train operation?

5.3.3.3. Summary

To summarise, we have identified four epistemic uncertainties with the help of the HOT-PIE taxonomy (see Table 32). While observations 1, 2 and 4 may not be directly related to the actual cause of the railway accident, observation 3 (uncertainty about the software algorithm during signal abnormality) may plausibly lead to the discovery of the error in the track signalling system that causes the accident. Hence, the HOT-PIE taxonomy can potentially help safety analysts to recognise safety hazards to a system.

Table 32. Summary of Epistemic Uncertainties Identified using HOT-PIE Taxonomy

Observation	1	2	3	4
Uncertainty	Fatigue among workers	Infrastructure support by Shanghai Railway Bureau	Software algorithm during signal abnormality	Human intervention during signal abnormality
HOT-PIE primary causal condition (cause)	Human	Organisation	Technology	Process
HOT-PIE primary causal condition (effect)	Technology	Technology	Information	Process
HOT-PIE causal mechanism	H2: Mental State (Fatigue)	O3: Resource (Infrastructure)	T1: Machine (Software)	P1: Nature (Oversight)

To move on, we assume that the estimated risk due to these four observations are at least tolerable such that the stakeholders can embark on managing the epistemic uncertainties. The next step is to prioritise the uncertainty analyses needed to manage these uncertainties.

5.3.4. Step 2: Document and Prioritise Uncertainty Analyses

Following the process flow from Figure 61, the next step is to prioritise the uncertainty analyses for the four epistemic uncertainties found in the previous sections. This is conducted with the help of the guided questions under the T.A.G. approach User Guide (see Annex K). In this section, we will evaluate the criticality and expected effort to manage Uncertainty Observation 1, while the assessment for the remaining 3 Uncertainty Observations are attached in Annex D.

5.3.4.1. *Uncertainty Observation 1: Fatigue among Workers*

The responses to the guided questions in the User Guide are shown in Table 33. It is to note that the responses to the guided questions are fictional.

Table 33. Responses to Guided Questions for Uncertainty Observation 1

Factors	Questions
Criticality: Assess the criticality of the target to affect system safety from the stakeholders' perspectives.	
Probability	Q1. How likely will this target cause harm to the system? May not be often as it happens only when maintenance need to be conducted. Need information on maintenance frequency to make better judgement
Severity	Q2. How serious are the consequences if this target occurs? Fatigue can lead to performance slip and directly cause harm to the engineers and the rail system.
Complexity: Assess the complexity of the system related to the target from the stakeholders' perspectives.	
Safety Causal Model	Q3. How complex are the safety causal models from where the target is identified? (e.g. no. of related causal conditions, hazard interdependency, residual risk interdependency) No safety causal model involved.
System Model	Q4: How complex are the system models from where target is identified? (e.g. system structure / product size / product design, process / task interdependency, process design, technology interdependency, project management difficulties) Simple system model of workers maintaining TCC equipment
Model Type	Q5: How complex are the model types from where target is identified? (e.g. reach and depth of patterns, templates and conventions used by the safety and system models) It follows conventional model of workers maintaining equipment when there is equipment failure. Possible complexity lies in the type of maintenance work being conducted.

Factors	Questions
Novelty: Assess the novelty of knowledge related to the target from the stakeholders' perspectives.	
Safety Novelty	Q6: How much of past information was used to define the safety causal model from where the target is identified? (e.g. causal factors, hazard list) No safety causal model involved
System Novelty	Q7: How much of past information was used to define the system model from where the target is identified? (e.g. system legacy, process, technology, objective) Maintenance history and worker conditions during maintenance should be available from experience.
Model Type Novelty	Q8: How much of past information was used to define the model type from where the target is identified? (e.g. patterns, templates, conventions) Maintenance and repair tasks should be following standard orders and manuals.
Experience	Q9: Are there other prior experiences related to the target? Not sure if workers are familiar with all the repairs that are expected from them.
Resource Availability: Assess the resource availability related to the target from the stakeholders' perspectives.	
Capability	Q10: Do we have the capability to learn the knowledge needed? (e.g. manpower, skills, infrastructure and support, environment, other resources) Yes, easy task to collect information. No specialised skill set required.
Urgency	Q11: Do we have the time required to learn the knowledge needed? (e.g. project deadline, lifecycle milestones) Yes, there should be sufficient time to find out the information.

From the responses, we assess the criticality, complexity, novelty, and resource availability of managing the epistemic uncertainty regarding the level of fatigue among the workers. We make use of the instructions in section 4.3.2.3 to calculate the priority scores (reproduced below for reference).

1. Identify the uncertainty in the targeted causal relationship.
2. Analyse the criticality of the target based on the guided questions. Score the criticality from 0 (low criticality) to 1 (high criticality).
3. Analyse the complexity of the target based on the guided questions. Score the complexity from 0 (low complexity) to 1 (high complexity).

4. Analyse the novelty of the target based on the guided questions. Score the novelty from 0 (low novelty) to 1 (high novelty).
5. Analyse the resource availability of the target based on the guided questions. Score the resource availability from 0 (high resource availability) to 1 (low resource availability).
6. Calculate the overall expected effort by average out the sum of the three scores under complexity, novelty and resource availability (assuming the factors have equal weightage).
7. Locate the quadrant on the prioritisation matrix that corresponds to the score for criticality and expected effort.

The criticality is given a score of 0.7 which implies that the uncertainty is of high risk (see Table 34). The expected effort is calculated from the average score of complexity (0.2), novelty (0.3) and resource availability (0.2). The average expected effort is calculated as 0.2, which implies that the task to analyse the uncertainty may not take a lot of effort. Based on the scores for criticality and expected effort, we plotted the Uncertainty Observation on the prioritisation matrix chart (see Figure 66).

Table 34. Scores based on Prioritisation Factors

Factors	Score
Criticality	0.7
Complexity	0.2
Novelty	0.3
Resource Availability	0.2
Expected effort	0.2

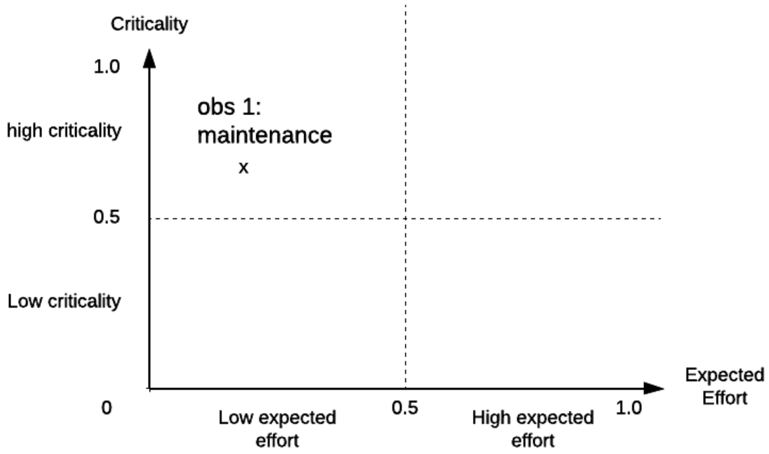


Figure 66. Uncertainty Observation 1 Plotted on Prioritisation Matrix Chart

We indicate on the top left quadrant the point that corresponds to the score of 0.7 criticality and 0.2 expected effort.

5.3.4.2. Summary

The same prioritisation process was conducted on Uncertainty Observations 2 to 4 (see Annex D). The following Table 35 summarises the calculation of priority for all the observations.

Table 35. Summary of Prioritisation for the Uncertainty Observations

Observation	1	2	3	4
Uncertainty	Fatigue among workers	Infrastructure support by Shanghai Railway Bureau	Software algorithm during signal abnormality	Human intervention during signal abnormality
Criticality	0.7	0.4	0.9	0.8
Complexity	0.2	0.2	0.7	0.6
Novelty	0.3	0.2	0.6	0.5
Resource Availability	0.2	0.7	0.7	0.6
Expected effort	0.2	0.4	0.7	0.6

Using the scores for criticality and expected efforts, the 4 uncertainty observations are plotted on the prioritisation matrix chart as shown in Figure 67.

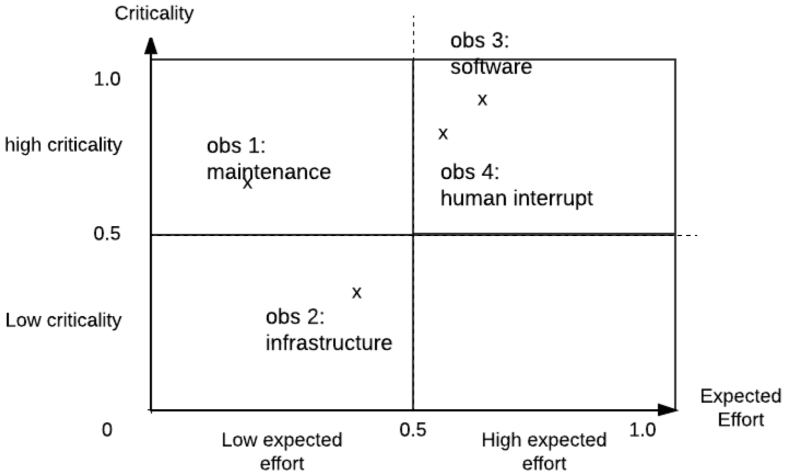


Figure 67. Uncertainty Observation 1 to 4 Plotted on Prioritisation Matrix Chart

Example. For example, while this may not be the only strategy, stakeholders may choose to focus first on analyses that are of high criticality and follow by those that required lower efforts. If this is the strategy adopt, the priority of action would be as follows:

Focus on Top Left Quadrant First. Since the expected effort is low and the uncertainty is critical, stakeholders should consider managing the Uncertainty Observation 1 about the working condition of the signal engineers first.

Pay More Attention to Top Right Quadrant: While the uncertainties are critical, the expected efforts are considerable. Hence, the stakeholders are advised to allocate more resources to manage the uncertainties from Uncertainty Observation 3 (software issue) and 4 (human intervention issue). By having an overview of the resources needed to manage these critical uncertainties, stakeholders can make a more informed and confident decision about the efforts needed to track and address these uncertainties concurrently.

Low Hanging Fruits in the Bottom Quadrants: While the expected effort is low, the uncertainty is assessed to not be as critical as the others. Hence, the stakeholders are advised to focus on Uncertainty Observation 2 (infrastructure support) only if there are resources readily available to manage the uncertainty. If there are any shared resources with other uncertainties, priority should go to uncertainties that are in quadrants with higher criticality.

5.3.5. Step 3: Generate Track and Address Action Plan

After prioritising the plausible-but-uncertain causal relationship, the next step is to generate the action plan to track and address the uncertainties throughout the system lifecycle. This is usually conducted towards the end of each safety assessment, where uncertainties causal relationships and other safety artefacts such as residual risk and safety requirements are consolidated. These uncertainties will be tracked throughout the system engineering lifecycle till either the next safety assessment or some pre-defined points that are triggered for the stakeholders to address them.

In this section, we will generate the action plan to track and address Uncertainty Observation 3 regarding software algorithm during signal abnormality. This observation is chosen as it is relatively critical compared to the other observations and the high expected effort would most likely require a structured plan to track and address the uncertainty.

5.3.5.1. Uncertainty Observation 3: Software Algorithm during Signal Abnormality

The action plan is formulated by responding to the series of questions in the T.A.G. approach user guide (see Annex K). The details of the responses are attached in Table 36.

Table 36. Guide to Develop Action Plan for Uncertainty 3

Target Uncertainty: Uncertainty about software algorithm during signal abnormality
1: **Track Sensors**

Target Uncertainty: Uncertainty about software algorithm during signal abnormality	
1a: **Decide Monitoring Technique** – What shall we track?	Q1. What are the monitoring techniques needed to track the epistemic uncertainties? • Software algorithm, software operation onboard the train and train station • Involves proper allocation and management of resources to understand the role of software in controlling signal abnormality during train operation. • Separate manpower (software engineers) are needed to conduct the investigation. Q2. Why are the monitoring techniques able to meet the goal of managing the uncertainty condition? • Knowing the software algorithm will provide clarity on the software functions for train operation. This will help in the investigation into the possibilities of having any failed-safe software modules to mitigate the risk due to signal abnormalities.
1b: **Form Monitoring Activity** – How do we track?	Q3. What types of monitoring activity are to be taken? • Investigate the safety impact of software modules on physical train operations, especially look for any software that mitigate the risk of false signals Q4. When should these monitoring activity be collected from the techniques? • Can be a paper study by reviewing the software manuals and documentation. • Unsure if observation of software performance is needed during operation. This will require more resources and expertise. Q5. Who are responsible to do the tracking? • Software engineers need to be assigned to capture and analyse the software information Q6. What are the structures and supporting resources to put in place for the tracking? • A structured management oversight appointed by the Shanghai railway bureau needs to be in place to ensure data is feedback to the safety team for analysis. • Regular update to be presented at the monthly safety forum. • Regular data collection is needed to capture the investigation on software modules. Q7. What are the skills, experiences and attitudes needed to do the tracking? • Require software engineers that are competent in analysing software modules in train operations.
2: **Address Responses**	
2a: **Set Trigger Points** – When and how shall we decide to respond?	Q8: What are the trigger points from the monitoring activity that require proactive response? • Discovery of software modules that affect the safety of railway operation. • Discovery of software modules that potentially mitigate risk by providing failed-safe algorithm or highlight signal abnormality. Q9. Who are responsible to decide on the respond actions?

Target Uncertainty: Uncertainty about software algorithm during signal abnormality	
	- Shanghai railway bureau safety management committee. Q10. What are the governance and supporting structure to put in place to make the decision? - Results and observations to be presented in the monthly safety forum. - If trigger points are met, software engineers should escalate result directly to the safety committee through a safety incident report. Q11. What are the skills, experiences and attitudes needed to make the decision? - Knowledge about software operations and software functions. - Knowledge about latest development in the railway system and operation. - Having the authority or access to higher management that has the authority to review train operations.
2b: **Adapt to Change** – What are the possible responses?	Q12. Who is responsible to review the changes that will be put in place? - Project safety engineers, together with software engineers and operators. Q13. How prepared and responsive should the system be in addressing the uncertainty? - If no data is observed in the next 2 months, safety committed needs to decide on a different action plan if the uncertainty is still considered to plausibly be safety-critical. - Depends on the findings regarding the software modules, project safety committee needs to decide on the urgency to escalate the issue to higher management. Q14. What are the structure and support resources to put in place to address the uncertainty? - Process to escalate safety concerns will follow the existing safety report channels. - Incident report to be submitted when trigger points are met. - Potentially a review of the software-hardware interface components may be needed depending on the investigation. Q15. What are the skills, experiences and attitudes needed to address the uncertainty? - Knowledge about software operations and functions.

Using the guide, we have developed the following track and address action plan.

Track Portion of the Action Plan. Tracking the uncertainty involves investigating software modules and algorithm onboard the train and train station. This is because knowing the software algorithm will provide clarity on the software functions for train operation. This will help in the investigation of having any fail-safe software modules to mitigate the risk due to signal abnormalities.

A detailed resource allocation and management plan needs to be developed after a better understanding of the software in controlling signal abnormality during train operation. To do that, dedicated software engineers are needed to conduct the investigation. The software engineers should be competent in analysing software modules in train operations.

The software engineers are responsible to investigate the safety impact of software modules on physical train operations, especially focus on any inherence software that can mitigate the risk of false signals. This can be a paper study by reviewing the software manuals and documentation or if needed, a more in-depth observation of software performance is needed during operation. Note that the second option will incur more resources and expertise.

A structured management oversight appointed by the Shanghai railway bureau needs to be in place to ensure data is feedback to the safety team for analysis. Regular data collection is expected to capture the investigation on software modules and regular updates are to be presented at the monthly safety forum.

Address Portion of the Action Plan. The uncertainty will trigger a response when there is a discovery of software modules that either affect the safety of railway operation or potentially mitigate risk by providing failed-safe algorithm or highlight signal abnormality.

The Shanghai railway bureau safety management committee shall be responsible to decide on the respond actions. Results and observations shall be presented in the monthly safety forum. The process to escalate safety concerns will follow the existing safety report channels. When trigger points are met, dedicated software engineers shall escalate result directly to the safety committee through a safety incident report.

A review committee that includes project safety engineers, together with software engineers and operators, shall be responsible to review any changes that will be put in place to address the uncertainty. The review committee should have knowledge about software operations and software functions, as well as the latest development in the railway system and operation. It should also have the authority or access to higher management that has the authority to review train operations.

If no data is observed in the next 2 months, the review committee needs to decide on a different action plan if the uncertainty is still considered to plausibly be safety-critical. Depends on the findings regarding the software modules, the committee needs to decide on the urgency to escalate the issue to higher management. There may potentially be a review of the software-hardware interface components depending on the investigation.

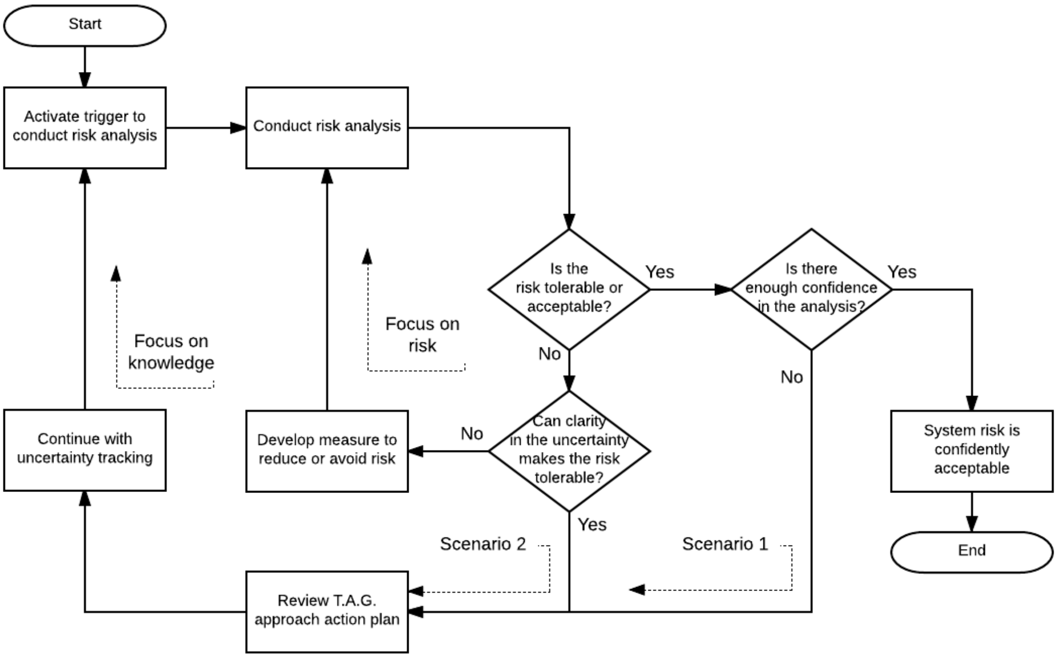

Start
Activate trigger to conduct risk analysis
Conduct risk analysis
Is the risk tolerable or acceptable?
Yes
No
Is there enough confidence in the analysis?
Yes
No
Focus on knowledge
Focus on risk
Continue with uncertainty tracking
Develop measure to reduce or avoid risk
No
Can clarity in the uncertainty makes the risk tolerable?
Yes
System risk is confidently acceptable
End
Scenario 2
Scenario 1
Review T.A.G. approach action plan

Figure 68. Possible Scenarios when Addressing Uncertainty in Software Algorithm

Possible Actions when Trigger to Address Uncertainty. When being triggered to address the uncertainty (see Figure 68), stakeholders would conduct a risk analysis to assess the risk that is still facing the system, as well as the assessing the confidence in the analysis. At this moment, stakeholders would have to decide if the assessed risk due to the software algorithm uncertainty is confidently within the tolerable or acceptable region. If it is not, a decision would have to be made if one or both of the following needs to be conducted:

- **Focus on Risk through System Change**. If the assessed risk has become unacceptable, stakeholders would need to derive strategies to reduce the risk. Possible approaches include modifying the software or changing the ways that the system is being operated to mitigate the software limitation. An example could be to implement software that prompt and alert the railway watchers when there is signal abnormality. This is assuming that the stakeholders have accessed to the software code. Stakeholders would have to conduct the safety analysis again to conclude if the new measure or mitigation does lower the risk to a tolerable or acceptable level.

- **Focus on Knowledge through Uncertainty Clarification.** The T.A.G. tracking on the epistemic uncertainty may need to continue under certain scenarios. The following are two possible scenarios to solicit more information to clarify the epistemic uncertainties.

    - **Scenario 1 – Increasing Confidence.** Even when the assessed risk is tolerable, there may still not be enough confidence in the assessment due to the residual uncertainties. For example, the investigation team may have obtained the preliminary software document that indicate the presence of the software modules to manage signal abnormality. However, since it is preliminary, the team decides to continue and track the uncertainty till the final software document is available to validate that the safety functions have indeed been implemented.

- **Scenario 2 – Reducing Risk.** In the second scenario, clarifying certain epistemic uncertainties may directly help to reduce assessed risk. For example, after investigating the software modules, it has been discovered that software written separately by two subcontractors may have contradicted each other. As a result, more investigation needs to be conducted by soliciting information from these two subcontractors. Hence, there is a need to continue and track the uncertainty due to software algorithm as new information is needed to lower the system risk into the tolerable region.

To end the uncertainty tracking, stakeholders must be satisfied that the uncertainty due to software algorithm being tracked by the T.A.G. approach has either been eliminated or that the risk is assessed to be confidently acceptable despite the decision to stop tracking the uncertainty in the software algorithm.

Uncertainty tracking will stop when there is sufficient clarity related to the uncertainty such that the safety analysts can use the knowledge to influence safety assessment. For example, the uncertainty may no longer pose a safety concern and hence drop from the tracking. In another example, the uncertainty, after clarification, may have invalidated an assumption made during a previous safety assessment. In such case, the safety analysts would need to review the impact of the invalidated assumption on the safety assessment, while continuing to track the uncertainty.

5.4. Initial Evaluation and Summary

In this chapter, we have highlighted the contribution of the research specifically from applying the T.A.G. approach with STPA on the Yongwen railway accident case study. Using the T.A.G. approach, we could discover four plausible-but-uncertain causal relationships that can potentially be safety-critical and one of them can plausibly lead to the discovery of the error in the track signalling system that causes the accident. Two of the causal relationships are linkage among nodes with existing causal relationship while the other two are linkages between nodes that do not have prior relationships with each other. Using the prioritisation factors of criticality and expected effort, the uncertainties were prioritised according to the analyses that are needed to clarify the unknowns. Action plans were also developed based on the guided questions in the T.A.G. approach user guide, which allow users to track and address the uncertainties through-life.

While the four plausible-but-uncertain causal relationships may be identifiable using the existing STPA technique, this would not be conducted in a systematic way like the T.A.G. approach. The ways that the SCS, UCAs and CFs are currently presented may not allow easy discovery of plausible causal relationships. Furthermore, as STPA requires users to form close control loops when identifying hazards, it may not be easy to identify and document epistemic uncertainties that may not contribute to close control loops. In comparison, the T.A.G. approach provides an easier approach to capture epistemic uncertainties, as long as one can define the causal relationship associated with the uncertainty, based on certain causal mechanism.

Safety analysts may be concerned that extensive resources are needed to apply the T.A.G. approach on existing safety assessment techniques. While the concern is valid, such uncertainties, if not managed well, can potentially lead to safety-critical consequences. Furthermore, we have not discovered any feasible and systematic methods to manage such uncertainties in safety assessment during our literature survey. It is not clear if there is any alternate approach that uses less resources and efforts than the T.A.G. approach and provide an acceptable level of confidence in managing the uncertainties. As part of the overall evaluation, we would revisit the resource concern in chapter 7.

Chapter 6 – Application of T.A.G. Approach in Safety Assessments from Component Viewpoints

6.1. Introduction

While the previous chapter focuses on conducting safety assessments from a system viewpoint, we focus on conducting safety assessments from the component viewpoint in this chapter (see Table 37).

Table 37. Summary of Chapters that Apply T.A.G. Approach

Chapter	Viewpoint	Safety Assessment Technique	Referenced Scenario
5	System	Systems-Theoretic Process Analysis	Yongwen railway accident analysis by Song et al [24]
6	Component	Fault Tree Analysis and Failure Modes and Effects Analysis	ARP 4761 aircraft design example [7]

For safety assessment with models from the component viewpoints, besides discovering epistemic uncertainty, the T.A.G. approach can help stakeholder to be more aware of the following characteristics of component models:

- **Limitation of Component Model**. When analysing components, we can expect the safety assessment techniques to be more focus and narrow in their expressive power. For example, a fault tree could be constructed for a component by focusing solely on the electrical properties of the component. It does not normally consider other factors such as human error or environmental conditions. By considering the different causal conditions in the HOT-PIE taxonomy, it helps the stakeholders to appreciate the focus and limitation of the models (including the safety assessment techniques) that are being used when analysing a component.

- **Association Relationship between Component Models.** Since a component model may be limited by its expressive power, it is often dependent or influenced by other component models (refer to the discussion of association relation between models in section 3.3.1.2). For example, a fault-tree that considers electrical circuit diagram for components may depend on other models that assume certain way of operating and maintaining the components. Separately, there could be a human factor analysis that has considered the situations in which the operators or engineers may not follow the specific procedure due to issues like distraction or fatigue.

  While both the fault tree and the human factor analysis are independent analyses, the findings from the human factor analysis can have safety implications on the assumptions and assessment using the fault tree. This association between the fault tree and the human factor analysis may not commonly surfaced and tracked in safety assessment, especially if the safety analysts conducting the assessment are electrical engineers focusing on analysing the electrical circuit diagram. The HOT-PIE taxonomy can be a tool to discover such associations between different component models, such that the stakeholders can derive action plan to track and response accordingly when an observation from one model affects the safety assessment in another model.

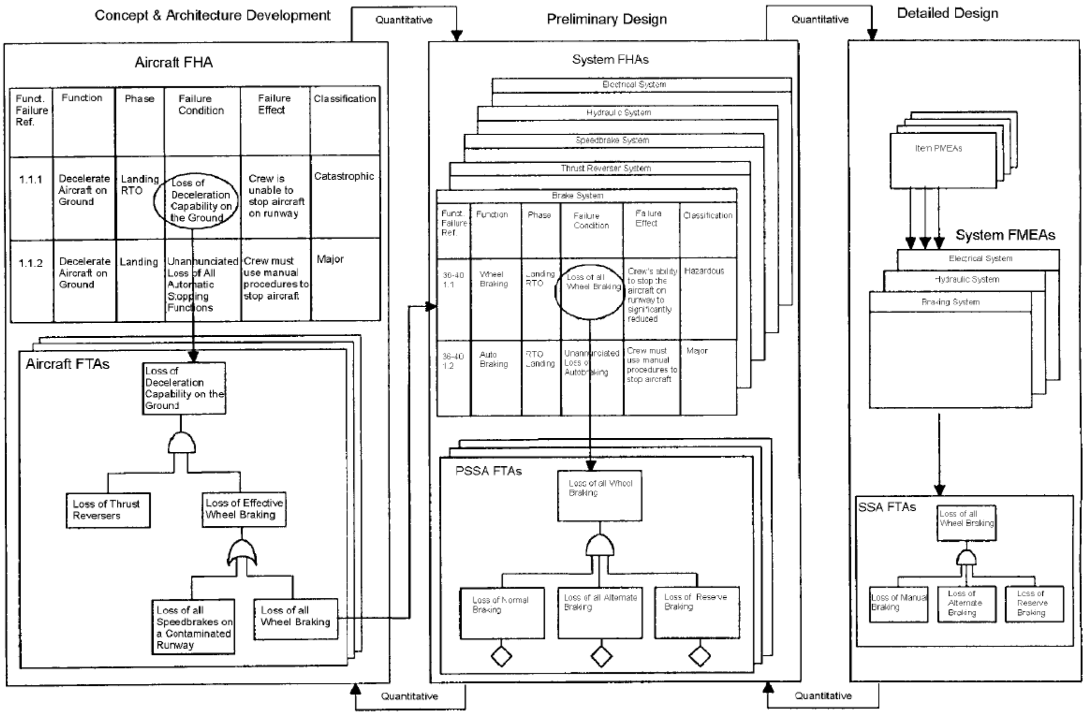
ARP 4761: Aircraft Design Example. We have chosen to base our study on the ARP 4761 standard, which is an industrial standard for conducting safety assessment to certify civil aircraft. The examples in the standard provide suitable references at the component viewpoint level. The standard describes the three common phases of safety assessment across the system lifecycle: Function Hazard Assessment (FHA), Preliminary System Safety Assessment (PSSA) and System Safety Assessment (SSA). Details of these three phases are described in section 2.2.1.2.

The standard also includes a worked example of a typical safety assessment for a fictitious aircraft design. As described in Appendix L of ARP 4761 standard, the fictitious aircraft design example focuses on analysing functions that can potentially leads to the loss of the aircraft. The functions were complicated enough to apply the different safety assessment techniques and concise enough to show the process flow when using the techniques. The analysis focuses upon a specific system and eventually to specific components within the system.

The safety assessment was applied throughout the three phases of FHA, PSSA and SSA with different safety assessment techniques (see Figure 69).

- **Function Hazard Assessment (FHA)**. First, the FHA was applied on models looking from system viewpoints by considering system functions related to the loss of aircraft. Since we have discussed system viewpoints in chapter 5, we will not present the FHA in this chapter.
- **Preliminary System Safety Assessment (PSSA)**. Next, the PSSA was applied on the Wheel Brake System (WBS) component using the FTA technique. We will describe the FTA during PSSA in section 6.2. After which, we will evaluate how the T.A.G. approach can be integrated with this component viewpoint FTA in section 6.3.
- **System Safety Assessment (SSA)**. Finally, the SSA was applied on the same WBS component using the Failure Modes and Effects Analysis (FMEA) technique. We will describe the FMEA during SSA in section 6.4. After which, we will evaluate how the

T.A.G. approach can be integrated with this component viewpoint FMEA in section 6.5.



Funct. Failure Ref.	Function	Phase	Failure Condition	Failure Effect	Classification
1.1.1	Decelerate Aircraft on Ground	Landing RTO	Loss of Deceleration Capability on the Ground	Crew is unable to stop aircraft on runway	Catastrophic
1.1.2	Decelerate Aircraft on Ground	Landing	Unannunciated Loss of All Automatic Stopping Functions	Crew must use manual procedures to stop aircraft	Major

Funct. Failure Ref.	Function	Phase	Failure Condition	Failure Effect	Classification
36-40 1.1	Wheel Braking	Landing RTO	Loss of all Wheel Braking	Crew's ability to stop the aircraft on runway to significantly reduced	Hazardous
36-40 1.2	Auto Braking	RTO Landing	Unannunciated Loss of Autobraking	Crew must use manual procedures to stop aircraft	Major

Figure 69. Safety Assessment Techniques Applied on FHA, PSSA and SSA (extracted from ARP 4761)

6.2. Safety Assessment using FTA

In this section, we focus on the FTA technique applied on the WBS component during preliminary system safety assessment (PSSA). According to ARP 4761, the PSSA is a "*top down approach to determine how failures can lead to the functional hazards identified by the FHA*". The PSSA is normally conducted throughout the preliminary design phase of the system lifecycle. The outputs from PSSA are feedforward as inputs to the SSA during system development.

In this example, the FTA technique is applied during PSSA to identify failures at the WBS that might cause the aircraft failure conditions identified during FHA. Using the generic process of safety assessment from section 4.3, we construct a similar process flow for FTA (see Figure 70). The figure highlights three unique steps in the FTA technique that focus on risk analysis (step 1a to 1c). In the later part of the chapter, we will show how the T.A.G. approach can be integrated with the FTA process.

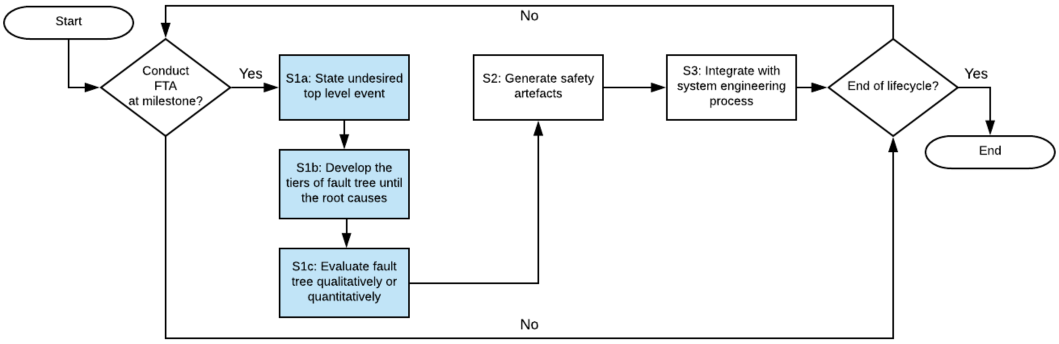

Start
Conduct FTA at milestone?
Yes
No
S1a: State undesired top level event
S1b: Develop the tiers of fault tree until the root causes
S1c: Evaluate fault tree qualitatively or quantitatively
S2: Generate safety artefacts
S3: Integrate with system engineering process
End of lifecycle?
Yes
No
End

Figure 70. Process Flow for FTA Safety Assessment

Step 1a (S1a): State Undesired Top-Level Event. In this example, the undesired top-level event identified in the WBS FTA is the "*unannunciated loss of all wheel braking*".

Step 1b (S1b): Develop Tiers of Fault Tree until Root Causes. The focus is to extend the fault event to the next lower level which are "*minimum, immediate, necessary, and sufficient*" to cause the top-level event. The fault tree shall be developed downward until the root causes are established or until it is considered unnecessary to develop further.

For the example, the WBS FTA is shown in Figure 71 (a detailed description of the WBS is attached in Annex F). The description of the fault tree is extracted from ARP 4761 and reproduced here for reference:

> *"This tree includes recognition of the fact that loss of all braking is considered hazardous and a design decision was made to cover all possible loss regardless of annunciation. This is reflected by the addition of the "Loss of Annunciation Capability" event which is ANDed with the "Loss of All Wheel Braking" event ... The probability of "Loss of Annunciation Capability" is set to 1.0. ... (The figure) also shows the downward development of the "Normal Brake" system design by budgeting failure rates to the identified contributors to "Normal Brake System Does Not Operate" (Loss of Hydraulic Supply, Loss of Hydraulic Components, and Loss of BSCU Ability to Command Braking). The Loss of BSCU Ability to Control Braking event is further broken down into the two major elements of BSCU Failure and Loss of Power to BSCU ..."*

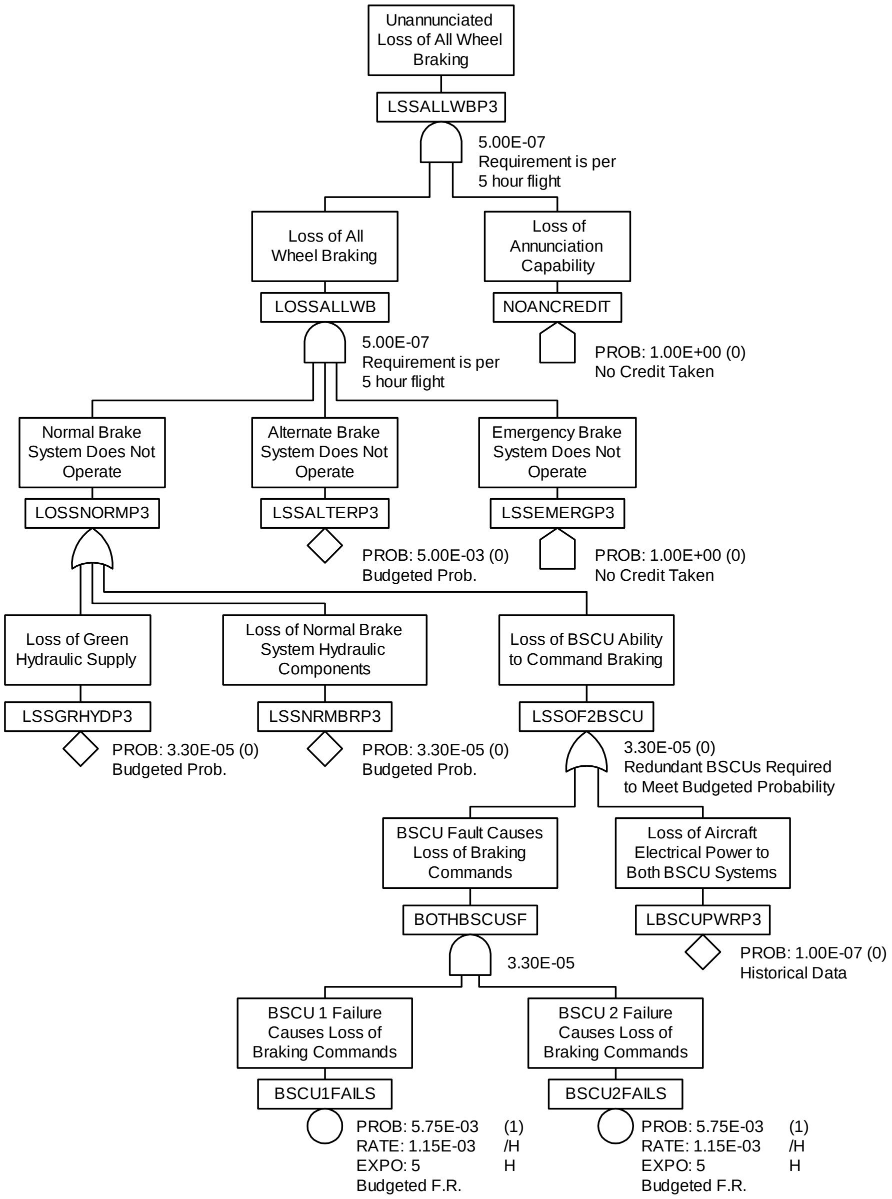

Unannunciated Loss of All Wheel Braking
LSSALLWBP3
5.00E-07
Requirement is per 5 hour flight
Loss of All Wheel Braking
LOSSALLWB
Loss of Annunciation Capability
NOANCREDIT
PROB: 1.00E+00 (0)
No Credit Taken
5.00E-07
Requirement is per 5 hour flight
Normal Brake System Does Not Operate
LOSSNORMP3
Alternate Brake System Does Not Operate
LSSALTERP3
PROB: 5.00E-03 (0)
Budgeted Prob.
Emergency Brake System Does Not Operate
LSSEMERGP3
PROB: 1.00E+00 (0)
No Credit Taken
Loss of Green Hydraulic Supply
LSSGRHYDP3
PROB: 3.30E-05 (0)
Budgeted Prob.
Loss of Normal Brake System Hydraulic Components
LSSNRMBRP3
PROB: 3.30E-05 (0)
Budgeted Prob.
Loss of BSCU Ability to Command Braking
LSSOF2BSCU
3.30E-05 (0)
Redundant BSCUs Required to Meet Budgeted Probability
BSCU Fault Causes Loss of Braking Commands
BOTHBSCUSF
3.30E-05
Loss of Aircraft Electrical Power to Both BSCU Systems
LBSCUPWRP3
PROB: 1.00E-07 (0)
Historical Data
BSCU 1 Failure Causes Loss of Braking Commands
BSCU1FAILS
PROB: 5.75E-03 (1)
RATE: 1.15E-03 /H
EXPO: 5 H
Budgeted F.R.
BSCU 2 Failure Causes Loss of Braking Commands
BSCU2FAILS
PROB: 5.75E-03 (1)
RATE: 1.15E-03 /H
EXPO: 5 H
Budgeted F.R.

Figure 71. WBS FTA: Unannunciated Loss of All Wheel Braking (extracted from ARP 4761)

Step 1c (S1c): Evaluate Fault Tree Quantitatively or Qualitatively. In the ARP 4761 example, the failure budget (measured by probability of failure) for the WBS components are calculated using the fault tree and feedback to the developer for design considerations. Calculation of the failure budget will not be discussed further as it is beyond the scope of our research on managing epistemic uncertainty.

Step 2 (S2): Generate Safety Artefacts. Generate and document the safety artefacts and requirements that will be integrated with the larger system engineering process.

Step 3 (S3): Integrate with System Engineering Process. The generated safety artefacts will be feedback to the system engineering process for further cause of actions if they affect the system lifecycle.

6.3. Integrate T.A.G. approach into FTA

PSSA is usually carried out during preliminary design phase of the system lifecycle. At this phase, there could still be many unknowns in the preliminary design where we can expect epistemic uncertainties. The T.A.G. approach can help stakeholders identify and document such uncertainties early and continue to track and address them later in the system lifecycle.

To evaluate the application of the T.A.G. approach, we use the WBS FTA constructed during the PSSA (see section 6.2). We will first describe the process flow of tagging uncertainties as part of the FTA (section 6.3.1). Next, we follow the process flow to complement FTA with the three steps in T.A.G. approach described in chapter 4:

- Step 1. Identify uncertainties using HOT-PIE taxonomy. This risk analysis is conducted while the WBS FTA is being generated (section 6.3.2).
- Step 2. Prioritise the uncertainty analyses based on the factors introduced in the T.A.G approach (section 6.3.3),
- Step 3. Generate track and address action plans using the guided questions introduced in the T.A.G. approach (section 6.3.4).

6.3.1. Process of Tagging Uncertainties in FTA

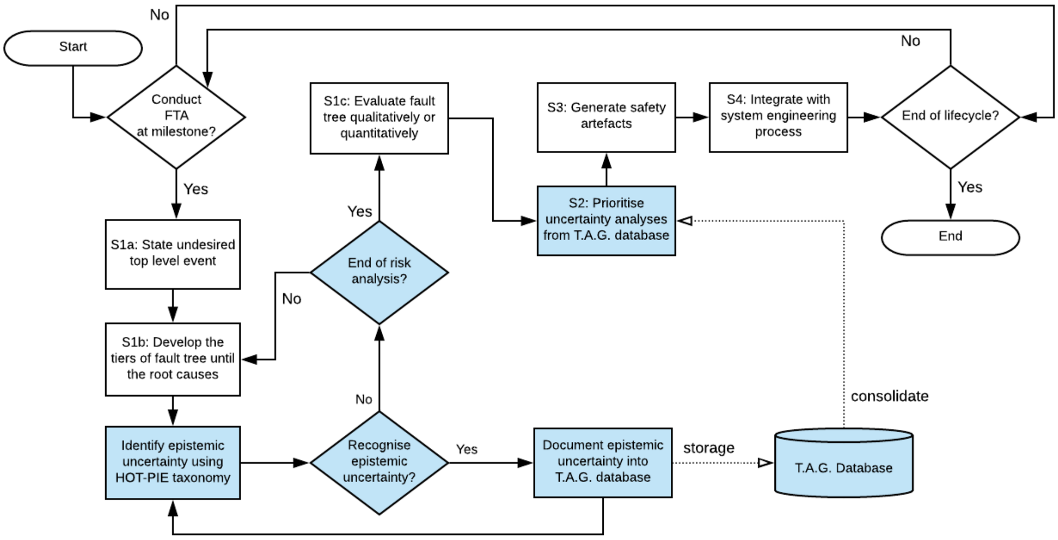

Start
No
Conduct FTA at milestone?
Yes
S1a: State undesired top level event
S1b: Develop the tiers of fault tree until the root causes
Identify epistemic uncertainty using HOT-PIE taxonomy
Recognise epistemic uncertainty?
No
Yes
End of risk analysis?
No
Yes
S1c: Evaluate fault tree qualitatively or quantitatively
Document epistemic uncertainty into T.A.G. database
storage
T.A.G. Database
consolidate
S2: Prioritise uncertainty analyses from T.A.G. database
S3: Generate safety artefacts
S4: Integrate with system engineering process
End of lifecycle?
No
Yes
End

Figure 72. Process Flow to Tag Epistemic Uncertainties in FTA Safety Assessment

Using the generic process of tagging uncertainties in safety assessment from section 4.3, we construct a similar process flow to tag epistemic uncertainties in FTA (see Figure 72). Once again, the process flow shows that the T.A.G. approach is not a separate standalone method to conduct safety assessment but rather it introduces additional steps to complement the existing FTA.

For the rest of this section, we will describe in details how epistemic uncertainties in the WBS FTA can be identified, documented, tracked and addressed by using the T.A.G. approach.

6.3.2. Step 1: Identify and Tag Uncertainties from Fault Tree

According to Figure 72, the first step of the T.A.G. approach is to tag uncertainties that are discovered while constructing the fault tree. Before uncertainties can be tagged, stakeholders need to identify the undesired top-level event (step 1a), as well as develop the fault tree (step 1b). The tagging is conducted with the help of the HOT-PIE taxonomy (a copy of the taxonomy can be found in the T.A.G. approach User Guide in Annex K). In this section, we evaluate the use of the taxonomy to identify epistemic uncertainties associated with the fault tree. Uncertainty Observation 1 is observed from the first layer of the fault tree (see section 6.3.2.1) and Uncertainty Observation 2 is observed from the second layer of the fault tree (see section 6.3.2.2).

6.3.2.1. Uncertainty Observation 1: Slow Reaction Due Limited Vision

From the fault tree in Figure 71, we focus on the first layer of the fault tree with the failure event– "*normal brake system does not operate*". We use the CRMSA model to describe the causal relationships related to the sub-branches of this failure event in Figure 73.

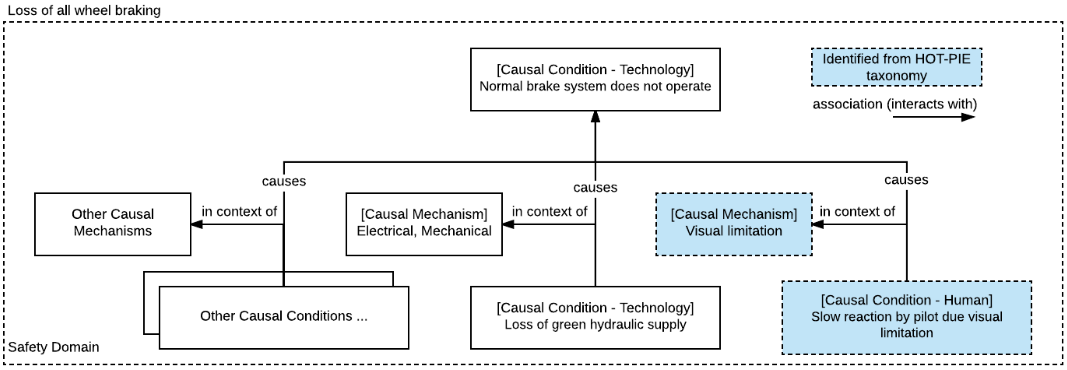

Loss of all wheel braking
[Causal Condition - Technology]
Normal brake system does not operate
Identified from HOT-PIE taxonomy
association (interacts with)
causes
causes
causes
Other Causal Mechanisms
in context of
[Causal Mechanism]
Electrical, Mechanical
in context of
[Causal Mechanism]
Visual limitation
in context of
Other Causal Conditions ...
[Causal Condition - Technology]
Loss of green hydraulic supply
[Causal Condition - Human]
Slow reaction by pilot due visual limitation
Safety Domain

Figure 73. Uncertainty about Reaction Time due Limited Vision

Next, we refer to the HOT-PIE taxonomy in the user guide to recognise plausible causal relationships. Searching through the *Human* primary condition, we found a relevant plausible-but-uncertain causal mechanism:

- Visual limitation (under secondary casual condition H2: Mental state)

Example. A plausible-but-uncertain causal relationship could be related to the limited vision on the pilot that may result in slower reaction when operating the normal brake system (see the highlighted causal relationship in Figure 73). Pertinent questions to consider include:

1. Can the vision of the pilot from the cockpit be obstructed such that they fail to notice dangers outside the cockpit and affect the operation of the normal brake system?

2. If so, what kind of conditions can lead to such bad vision from the cockpit?

3. How much reaction time does the pilot have, to operate the normal brake system?

6.3.2.2. Uncertainty Observation 2: Logic Error in Command Messages

From the fault tree in Figure 71, we move to the second layer of the fault tree with the failure event – “*loss of BSCU ability to command braking*”. The Brake System Control Unit (BSCU) is the computer that monitors signals from various WBS components to provide correct brake functions for normal operation. We use the CRMSA model to describe the causal relationships related to the sub-branches of this failure event in Figure 74.

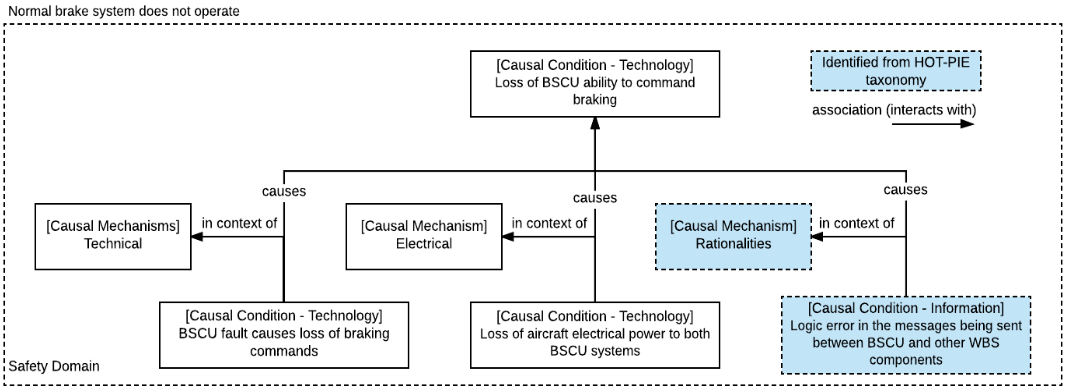

Normal brake system does not operate
[Causal Condition - Technology] Loss of BSCU ability to command braking
Identified from HOT-PIE taxonomy
association (interacts with)
causes
causes
causes
[Causal Mechanisms] Technical
in context of
[Causal Mechanism] Electrical
in context of
[Causal Mechanism] Rationalities
in context of
[Causal Condition - Technology] BSCU fault causes loss of braking commands
[Causal Condition - Technology] Loss of aircraft electrical power to both BSCU systems
[Causal Condition - Information] Logic error in the messages being sent between BSCU and other WBS components
Safety Domain

Figure 74. Uncertainty about the Logic in the BSCU Command Messages

Next, we refer to the HOT-PIE taxonomy in the user guide to recognise plausible causal relationships. Searching through the *Information* primary condition, we found a relevant plausible-but-uncertain causal mechanism:

- Rationalities regarding logic error (under secondary casual condition I2: Knowledge).

Example. A plausible-but-uncertain causal relationship could be related to some logic error in the message that is communicated between the BSCU and other WBS components (see the highlighted causal relationship in Figure 74). Pertinent questions to consider include:

1. Besides the power supply, are there other inputs or unintended messages that can be received by the BSCU from other components and result in loss of brake functions?

2. Has the design team assessed all of the interconnect specification related to messages received and transmitted by the BSCU?

6.3.2.3. Summary

To summarise, we have identified two epistemic uncertainties with the help of the HOT-PIE taxonomy (see Table 38). We have assumed that the estimated risk due to these two observations are at least tolerable such that the stakeholders can embark on managing the epistemic uncertainties. The next step is to prioritise the uncertainty analyses needed to manage these uncertainties.

Table 38. Summary of Epistemic Uncertainties Identified in WBS FTA

Observation	**1**	**2**
Uncertainty	Slow reaction due limited vision	Logic error in command messages
HOT-PIE primary causal condition (cause)	Human	Information
HOT-PIE primary causal condition (effect)	Technology	Technology
HOT-PIE causal mechanism	H2: Mental state (Visual limitation)	I2: Knowledge (Rationalities)

6.3.3. Step 2: Document and Prioritise Uncertainty Analyses

Following the process flow from Figure 72, the next step is to prioritise the uncertainty analyses for the two epistemic uncertainties found in the previous section. This is conducted with the help of the guided questions from the T.A.G. approach user guide (see Annex K). The assessment for the Uncertainty Observation is attached in Annex E.

The following Table 39 summarises the calculation of priority for the 2 observations.

Table 39. Summary of Prioritisation for the Uncertainty Observations from WBS FTA

Observation	Uncertainty	Criticality	Complexity	Novelty	Resource	Expected effort
1	Slow reaction due limited vision	0.8	0.6	0.3	0.4	0.4
2	Logic error in command messages	0.7	0.8	0.7	0.7	0.7

Using the scores for criticality and expected efforts, the 2 uncertainty observations are plotted on the prioritisation matrix chart as shown in Figure 75. Stakeholders can make use of the chart to decide which uncertainties to focus on and allocate the proportionate resources to manage them.

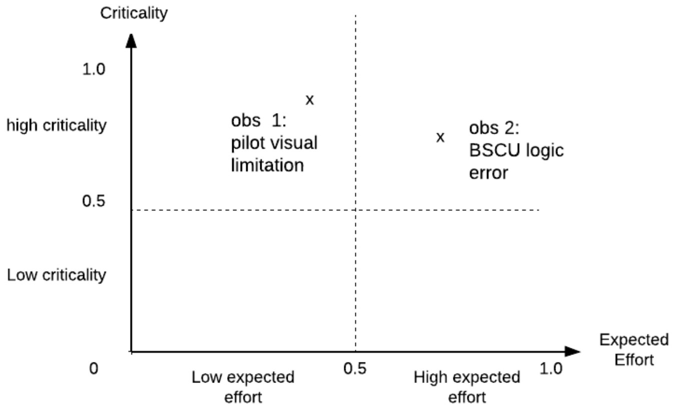


Figure 75. Uncertainty Observation 1 and 2 Plotted on Prioritisation Matrix Chart

6.3.4. Step 3: Generate Track and Address Action Plan

After prioritising the plausible-but-uncertain causal relationships, the next step is to generate the action plan to track and address the uncertainties after completion of the PSSA. These uncertainties will be tracked throughout the system engineering lifecycle till some pre-defined points are triggered for the stakeholders to address them.

The action plan is formulated by answering a series of questions from the T.A.G. approach user guide from Annex K. In this section, we will generate the action plan to track and address Uncertainty Observation 2 regarding logic error in command messages. This observation is chosen as it requires more expected effort as compared to Uncertainty Observation 1. Hence, a structured plan to track and address the uncertainty would be important.

6.3.4.1. Uncertainty Observation 2: Logic Error in Command Messages

The details of developing the action plan for Uncertainty Observation 2 are described in Annex G. Using the guide, we have developed the following track and address action plan:

Track Portion of the Action Plan. Tracking the uncertainty involves investigating the interfaces between the BSCU and other WBS components, such as the Interface Control Document (ICD). This is because having the ICD will allow the engineers to investigate the command messages that are passed between the BSCU and the other WBS components. This is needed to sieve out any potential unintended logic error in the command messages that are communicated between the components during operations.

To do that, dedicated communication engineers are needed to conduct the investigation. The communication engineers should be competent in analysing electronic signals in WBS operation.

The communication engineers are responsible for investigating the communication messages that are transmitted between the BSCU and the other WBS components and checking if there are any scenarios where correct messages are transmitted based on the technical specifications but the messages lead to unintended WBS operation. This can be a paper study by reviewing the ICD and interface documentation. It is unsure if actual test on the physical system is necessary to validate the design performance. This will incur more resources and expertise.

In this example, a structured management oversight appointed by the management needs to be in place to ensure data is feedback to the safety team for analysis. Regular update is expected at the weekly safety meeting.

Address Portion of the Action Plan. The uncertainty will trigger a response when there is a discovery of command message that either has logic error or affects the safety of the WBS operation.

The management safety committee shall be responsible for deciding upon the respond actions. When trigger points are met, the communication engineers shall escalate result directly to the management via incident report.

A review committee that includes project manager and communication engineers shall be responsible to review any changes that will be put in place to address the uncertainty. The review committee should have knowledge about communication engineering and WBS operation. It should also have accessed to higher management that has the authority to review WBS operation.

If no data is observed within a month, the project manager needs to decide on a different action plan if the uncertainty is potentially still safety-critical. Depends on the findings regarding the software modules, the project manager needs to decide on the urgency to escalate the issue to higher management. There may potentially be a review of the software-hardware interface components depending on the investigation.

Possible Actions when Trigger to Address Uncertainty. Taking reference from section 4.4.3, if the assessed risk due to uncertainty about the logic error in command message is still not confidently within the tolerable or acceptable region, a decision would have to be made if one or both of the following needs to be conducted:

- **Focus on Risk through System Change**. If the level of system risk has become unacceptable, stakeholders would need to derive strategies to reduce the risk. One approach could be to modify the software in either the BSCU or the WBS component to eliminate the logic error in the message. This is assuming that the stakeholders have accessed to the software code. Stakeholders would have to conduct the safety analysis again to conclude if the new measure or mitigation does lower the risk to a tolerable or acceptable level.

- **Focus on Knowledge through Uncertainty Clarification.** The T.A.G. tracking on the epistemic uncertainty may need to continue under certain scenarios. The following are two possible scenarios to solicit more information to clarify the epistemic uncertainties.
    - **Scenario 1 – Increasing Confidence.** Even when the assessed risk is tolerable, there may still not be enough confidence in the assessment due to the residual uncertainties. For example, the investigation team may have obtained the preliminary software document that indicate the logic behind the messages sent between the BSCU and the control valves. However, since it is preliminary, the team decides to continue and track the uncertainty till the final software document is available to validate that the implemented software code has no logic error.
    - **Scenario 2 – Reducing Risk.** In the second scenario, clarifying certain epistemic uncertainties may directly help to reduce assessed risk. For example, it is discovered that the software documentation for the green pump is incomplete. As a result, more investigation needs to be conducted to solicit the information from the relevant contractor. Hence, there is a need to continue and track the uncertainty as more information is needed to lower the system risk into the tolerable region.

To end the uncertainty tracking, stakeholders must be satisfied that the uncertainty due to logic error in the command message being tracked by the T.A.G. approach has either been eliminated or that the risk is assessed to be confidently acceptable despite the decision to stop tracking the uncertainty.

6.4. Safety Assessment using FMEA

In this section, we focus on the FMEA conducted on the WBS component during system safety assessment (SSA). According to ARP 4761, the SSA is a "*a systematic, comprehensive evaluation of the implemented system to show that relevant safety requirements are met*". The SSA is conducted to verify that the system that has been developed conforms to the safety requirements generated from the FHA and PSSA. It is usually conducted in the development phase of the system lifecycle, after the conceptual and design phases.

In this example, a piece-part WBS FMES is constructed during SSA to identify failure modes at the WBS and the effects they have on the aircraft braking function. The WBS FMES represents a higher level FMEA that is derived from the component FMEAs within the WBS. Using the generic process of safety assessment from section 4.3, we construct a similar process flow for FMEA (see Figure 76). The figure highlights three unique steps in the FMEA technique that focus on risk analysis (step 1a to 1c). In the later part of the chapter, we will show how the T.A.G. approach can be integrated with the FMEA process.

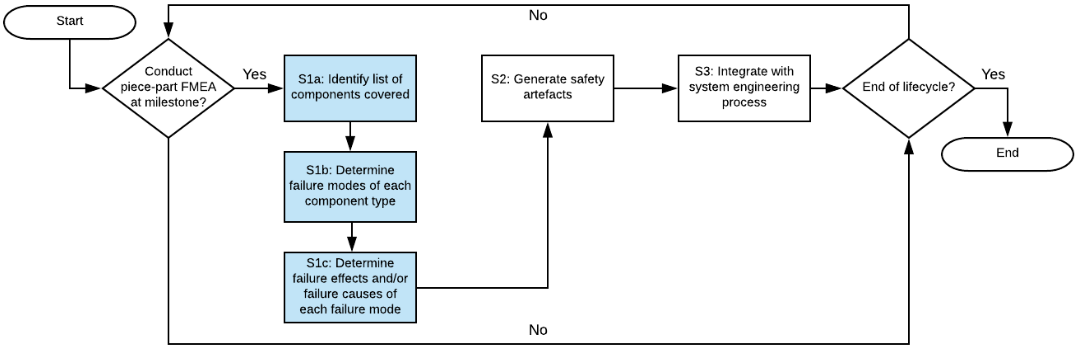

Start
No
Conduct piece-part FMEA at milestone?
Yes
S1a: Identify list of components covered
S2: Generate safety artefacts
S3: Integrate with system engineering process
End of lifecycle?
Yes
End
S1b: Determine failure modes of each component type
S1c: Determine failure effects and/or failure causes of each failure mode
No

Figure 76. Process Flow for FMEA Safety Assessment

Step 1a (S1a): Identify List of Components Covered. The WBS system description is attached in Annex F. We have reproduced a diagram of the WBS components here for reference (see Figure 77). The WBS comprises the BSCU computer that is connected to two independent hydraulic pumps via control valves. These pumps are in turn connected to the pedals and the aircraft wheels.

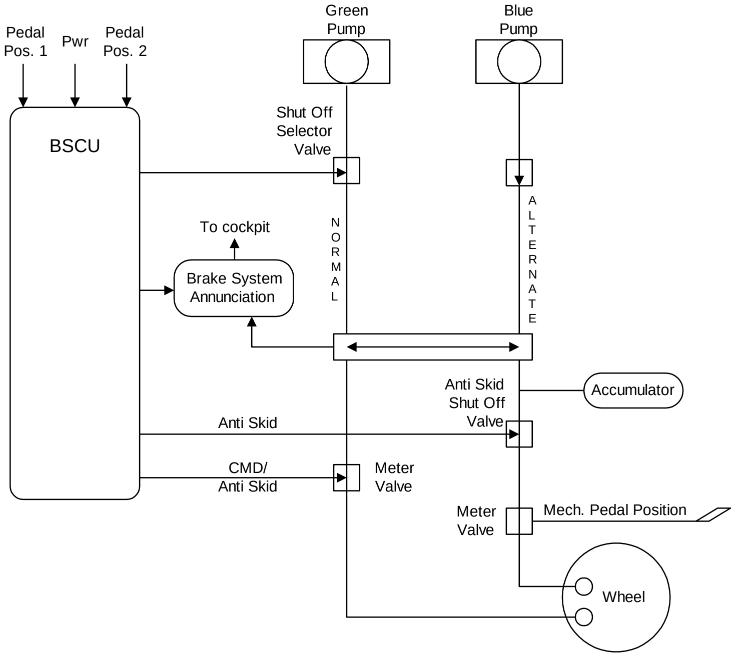
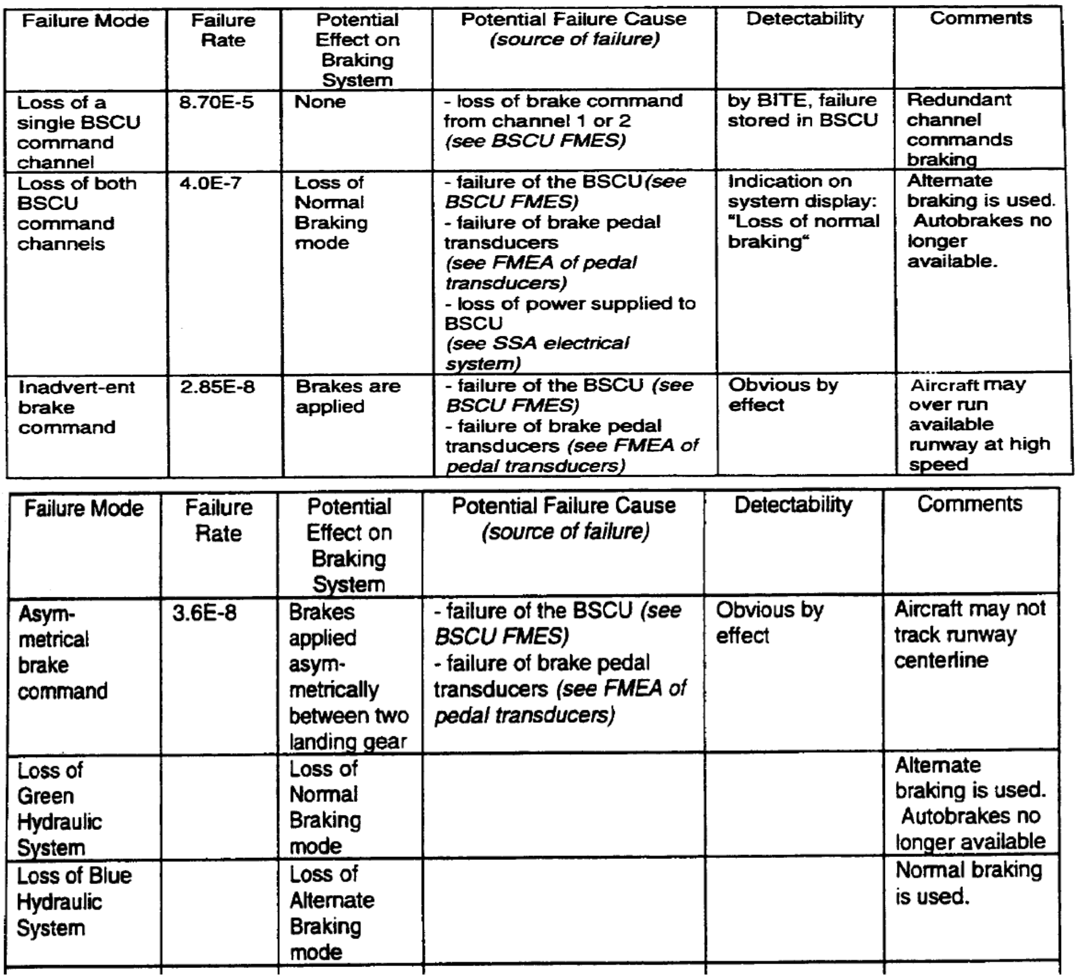


Figure 77. Wheel Break System (WBS) Components Breakdown (extracted from ARP 4761)

Step 1b (S1b): Determine Failure Modes of Each Component Type.

Step 1c (S1c): Determine Failure Effects and/or Failure Causes for each Failure Mode.

In the ARP 4761 example, the product of step 1b and 1c is the WBS FMES table, such as the one shown in Figure 78. The WBS FMES for aircraft breaking failure effect contains failure modes related to the components within the WBS. It is identified from the FMEAs conducted on individual components within the WBS. The potential failure causes are also extracted from individual component FMEAs.



Failure Mode	Failure Rate	Potential Effect on Braking System	Potential Failure Cause *(source of failure)*	Detectability	Comments
Loss of a single BSCU command channel	8.70E-5	None	- loss of brake command from channel 1 or 2 *(see BSCU FMES)*	by BITE, failure stored in BSCU	Redundant channel commands braking
Loss of both BSCU command channels	4.0E-7	Loss of Normal Braking mode	- failure of the BSCU*(see BSCU FMES)* - failure of brake pedal transducers *(see FMEA of pedal transducers)* - loss of power supplied to BSCU *(see SSA electrical system)*	Indication on system display: "Loss of normal braking"	Alternate braking is used. Autobrakes no longer available.
Inadvert-ent brake command	2.85E-8	Brakes are applied	- failure of the BSCU *(see BSCU FMES)* - failure of brake pedal transducers *(see FMEA of pedal transducers)*	Obvious by effect	Aircraft may over run available runway at high speed
Asym-metrical brake command	3.6E-8	Brakes applied asym-metrically between two landing gear	- failure of the BSCU *(see BSCU FMES)* - failure of brake pedal transducers *(see FMEA of pedal transducers)*	Obvious by effect	Aircraft may not track runway centerline
Loss of Green Hydraulic System		Loss of Normal Braking mode			Alternate braking is used. Autobrakes no longer available
Loss of Blue Hydraulic System		Loss of Alternate Braking mode			Normal braking is used.

Figure 78. Wheel Brake System FMES of Braking Failure Effect at SSA

Step 2 (S2): Generate Safety Artefacts. Generate and document the safety artefacts and requirements that will be integrated with the larger system engineering process.

Step 3 (S3): Integrate with System Engineering Process. The generated safety artefacts will be feedback to the system engineering process for further cause of actions if they affect the system lifecycle.

6.5. Integrate T.A.G. approach into FMEA

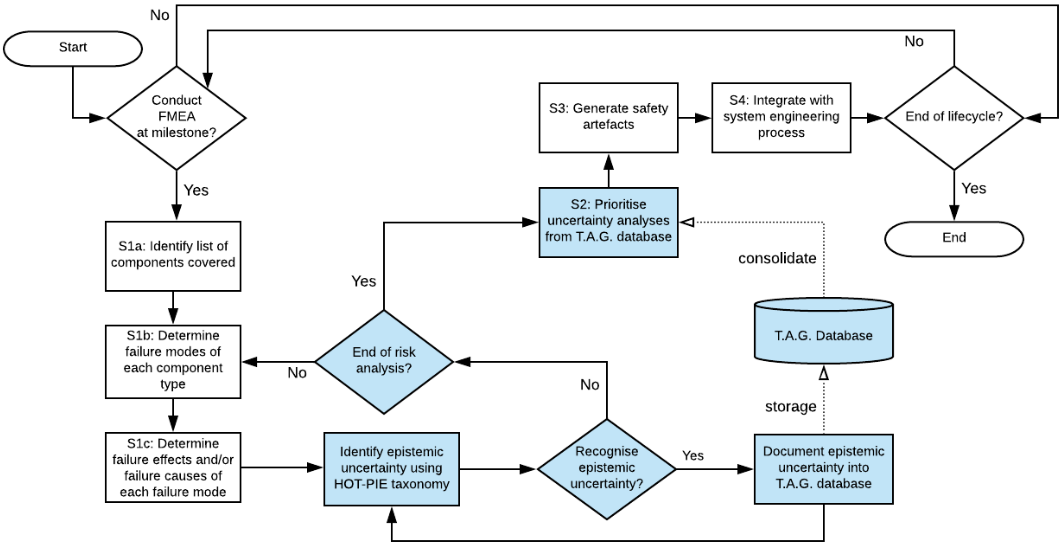
SSA is usually carried out during development phase of the system lifecycle. At this phase, system design may have been implemented and stakeholders would be involved in verifying the product against the design and safety requirements. There could still be epistemic uncertainties at this phase and the efforts to implement major changes may be massive if it involves system redesign. The T.A.G. approach remains relevant for stakeholders to identify and document such uncertainties and continue to track and address them later in the system lifecycle.

To evaluate the application of the T.A.G. approach, we use the WBS FMES constructed during the SSA in section 6.4. We will first describe the process flow of tagging uncertainties as part of the FMEA (section 6.5.1). Next, we will follow the process flow to complement FMEA with the three steps in T.A.G. approach described in chapter 4:

- Step 1. Identify uncertainties using HOT-PIE taxonomy. This risk analysis is conducted while the WBS FMES is being generated (section 6.5.2).
- Step 2. Prioritise the uncertainty analyses based on the factors introduced in the T.A.G approach (section 6.5.3),
- Step 3. Generate track and address action plans using the guided questions introduced in the T.A.G. approach (section 6.5.4).

6.5.1. Process of Tagging Uncertainties in FMEA

Using the generic process of tagging uncertainties in safety assessment from section 4.3, we construct a similar process flow to tag epistemic uncertainties in FMEA (see Figure 79). Once again, the T.A.G. approach is not a separate standalone method to conduct safety assessment but rather it introduces additional steps to complement the existing FMEA.



Start
No
Conduct FMEA at milestone?
Yes
No
S3: Generate safety artefacts
S4: Integrate with system engineering process
End of lifecycle?
Yes
End
S2: Prioritise uncertainty analyses from T.A.G. database
consolidate
S1a: Identify list of components covered
Yes
T.A.G. Database
S1b: Determine failure modes of each component type
No
End of risk analysis?
No
storage
S1c: Determine failure effects and/or failure causes of each failure mode
Identify epistemic uncertainty using HOT-PIE taxonomy
Recognise epistemic uncertainty?
Yes
Document epistemic uncertainty into T.A.G. database

Figure 79. Process Flow to Tag Epistemic Uncertainties in FMEA Safety Assessment

For the rest of this section, we will describe in details how epistemic uncertainties in the WBS FMES can be identified, documented, tracked and addressed by using the T.A.G. approach.

6.5.2. Step 1: Identify and Tag Uncertainties from FMEA

According to Figure 79, the first step of the T.A.G. approach is to tag uncertainties that are discovered while constructing the FMEA. Before uncertainties can be tagged, stakeholders need to identify the relevant components (step 1a), as well as determine the failure modes, effects and causes that constitute the FMEA (step 1b and 1c). The tagging is conducted with the help of the HOT-PIE taxonomy. In this section, we evaluate the use of the taxonomy to identify epistemic uncertainties from the WBS FMES. Uncertainty Observation 3 is about new linkage among components with existing causal relationship (see section 6.5.2.1) and Uncertainty Observation 4 is about linkage between components that do not have prior relationships with each other (see section 6.5.2.2). The differences between the two have been elaborated in section 4.2.7.

6.5.2.1. Uncertainty Observation 3: Mechanical Faults at Control Valves

From the FMES in Figure 78, we will first focus on the failure mode – "*loss of single BSCU command channel*". We use the CRMSA model to describe the causal relationships related to this failure mode between the BSCU and the hydraulic pumps in Figure 80. The original electrical failure cause due to BSCU is also shown in the figure for reference. In this example, a new causal mechanism between the BSCU and the pumps is discovered using the HOT-PIE taxonomy.

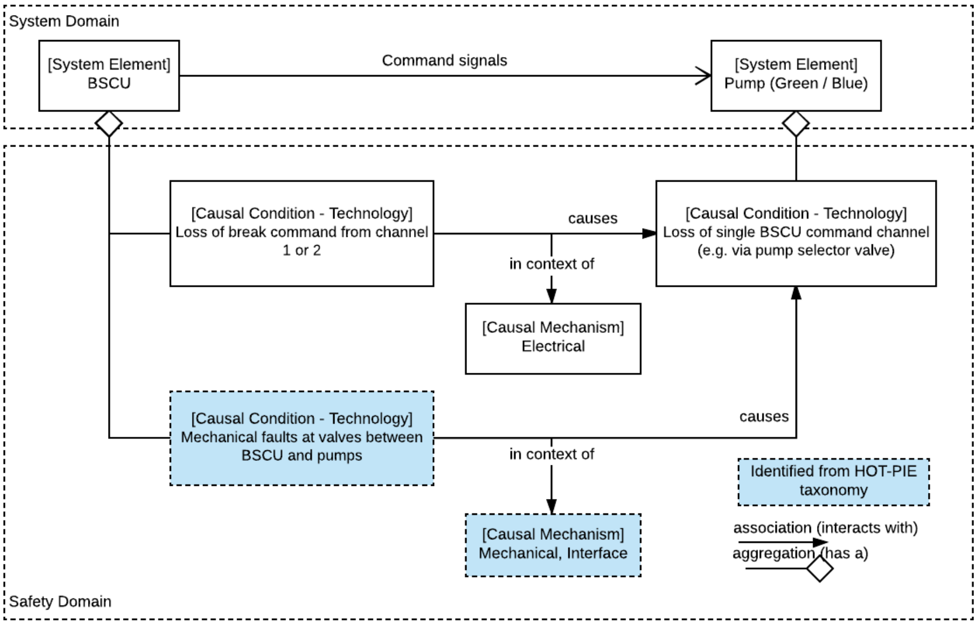

System Domain
[System Element] BSCU
Command signals
[System Element] Pump (Green / Blue)
[Causal Condition - Technology] Loss of break command from channel 1 or 2
causes
[Causal Condition - Technology] Loss of single BSCU command channel (e.g. via pump selector valve)
in context of
[Causal Mechanism] Electrical
[Causal Condition - Technology] Mechanical faults at valves between BSCU and pumps
causes
in context of
Identified from HOT-PIE taxonomy
[Causal Mechanism] Mechanical, Interface
association (interacts with)
aggregation (has a)
Safety Domain

Figure 80. Uncertainty about Mechanical Faults at Control Valves

Next, we refer to the HOT-PIE taxonomy in the user guide to recognise plausible causal relationships. Searching through the *Technology* primary condition, we found some relevant plausible-but-uncertain causal mechanisms:

- Interface (under secondary casual condition T1: Machine), and
- Mechanical (under secondary casual condition T2: Property)

Example. A plausible-but-uncertain causal relationship could be related to mechanical fault at the control valves between the BSCU and the hydraulic pumps (see the highlighted causal relationship in Figure 80). While the original focus in the FMES is on electrical failures between the BSCU and pumps, mechanical faults at the valve interfaces between the BSCU and the pumps may also result in the loss of BSCU command channel. For example, referring to Figure 77, a mechanical 'sticky' joint at the shutoff selector valve of the green pump may result in the loss of BSCU command signal to the pump. Pertinent questions to consider include:

1. Is there analysis conducted on the reliability of the mechanical valves?

2. How likely is it to lose the BSCU command signal when the mechanical joints are 'sticky'?

3. How does the mechanical consideration affect the composite failure rate calculated in the WBS FMES?

6.5.2.2. *Uncertainty Observation 4: Task Overload During Emergency*

From the FMES in Figure 78, we will next focus on the failure mode – "*inadvertent brake command*". We use the CRMSA model to describe the causal relationships related to this failure mode in Figure 81. The original two electrical failure causes due to brake pedals and BSCU are shown on the left. In this example, a plausible causal mechanism between the BSCU and a new component – pilot, is discovered using the HOT-PIE taxonomy.

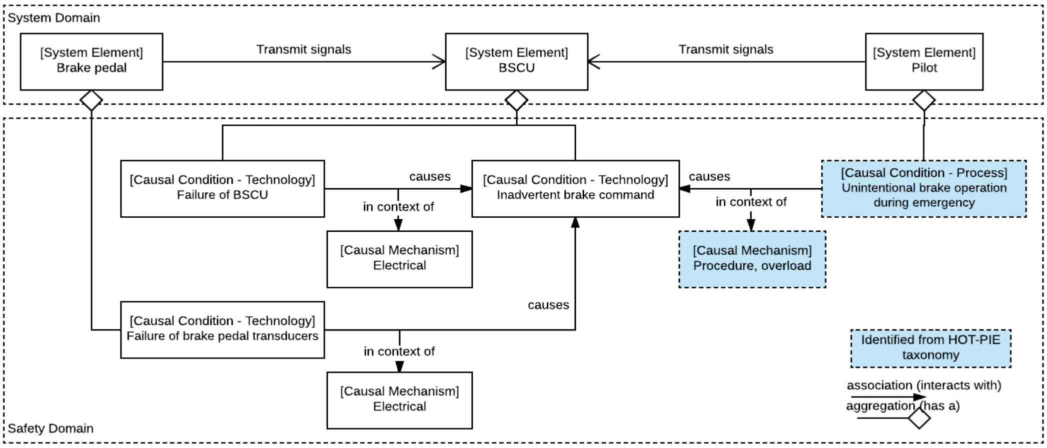

System Domain
[System Element] Brake pedal
Transmit signals
[System Element] BSCU
Transmit signals
[System Element] Pilot
[Causal Condition - Technology] Failure of BSCU
causes
in context of
[Causal Mechanism] Electrical
[Causal Condition - Technology] Inadvertent brake command
causes
in context of
[Causal Condition - Process] Unintentional brake operation during emergency
[Causal Mechanism] Procedure, overload
causes
[Causal Condition - Technology] Failure of brake pedal transducers
in context of
[Causal Mechanism] Electrical
Identified from HOT-PIE taxonomy
association (interacts with)
aggregation (has a)
Safety Domain

Figure 81. Uncertainty about Task Overload during Emergency

Next, we refer to the HOT-PIE taxonomy in the user guide to recognise plausible causal relationships. Searching through the *Process* primary condition, we found some relevant plausible-but-uncertain causal mechanisms:

- Procedure (under secondary casual condition P1: Nature), and
- Overload (under secondary casual condition P1: Nature).

Example. A plausible-but-uncertain causal relationship could be related to unintended operation of the brake system when the pilots are overloaded with multiple tasks during emergency (see the highlighted causal relationship in Figure 81). The concern is regarding possible distractions by the pilot during the deceleration process due to issues such as procedural and task overload during an emergency. With a better appreciation of the cockpit operation during system development, it could be timely to put more efforts in reviewing the human-machine interface during emergency.

1. Are there studies being conducted on the probability of the pilots being overloaded with multiple tasks during emergency?

2. How likely is it for the pilots to inadvertently activate the brake command when overloaded with multiple tasks?

3. How does the human-machine interface consideration affect the composite failure rate calculated in the WBS FMES?

6.5.2.3. Summary

To summarise, we have identified two epistemic uncertainties with the help of the HOT-PIE taxonomy (see Table 40). We have assumed that the estimated risk due to these two observations are at least tolerable such that the stakeholders can embark on managing the epistemic uncertainties. The next step is to prioritise the uncertainty analyses needed to manage these uncertainties.

Table 40. Summary of Epistemic Uncertainties Identified in WBS FMES

Observation	3	4
Uncertainty	Mechanical faults at control valves	Task overload during emergency
HOT-PIE primary causal condition (cause)	Technology	Process
HOT-PIE primary causal condition (effect)	Technology	Technology
HOT-PIE causal mechanism	T2: Property (Mechanical), T1: Machine (interface)	P1: Nature (Procedure, overload)

6.5.3. Step 2: Document and Prioritise Uncertainty Analyses

Following the process flow from Figure 79, the next step is to prioritise the uncertainty analyses for the two epistemic uncertainties found in the previous section. This is conducted with the help of the guided questions from the T.A.G. approach user guide (see Annex K). The assessment for uncertainty observations is attached in Annex H.

From the responses to the guided questions, we assess the criticality, complexity, novelty, and resource availability of managing the uncertainty observations. The expected effort is calculated from the average score of complexity, novelty and resource availability. The following Table 41 summarises the calculation of priority for the 2 observations.

Table 41. Summary of Prioritisation for the Uncertainty Observations from WBS FMES

Observation	Uncertainty	Criticality	Complexity	Novelty	Resource	Expected effort
3	Mechanical faults at control valves	0.8	0.8	0.4	0.6	0.6
4	Task overload during emergency	0.7	0.6	0.3	0.6	0..5

Using the scores for criticality and expected efforts, the 2 uncertainty observations are plotted on the prioritisation matrix chart as shown in Figure 82. Stakeholders can make use of the chart to decide which uncertainties to focus on and allocate the proportionate resources to manage them.

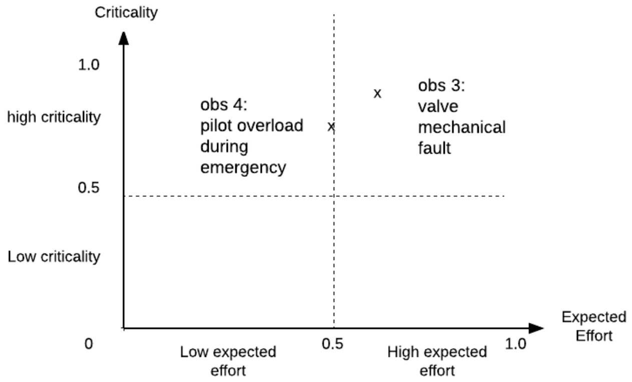


Figure 82. Uncertainty Observation 3 and 4 Plotted on Prioritisation Matrix Chart

6.5.4. Step 3: Generate Track and Address Action Plan

After prioritising the plausible-but-uncertain causal relationships, the next step is to generate the uncertainty management action plan to track and address the uncertainties after completion of the SSA. These uncertainties will be tracked throughout the system engineering lifecycle till some pre-defined points are triggered for the stakeholders to address them.

The action plan is formulated by answering a series of questions from the T.A.G. approach user guide from Annex K. In this section, we will generate the action plan to track and address Uncertainty Observation 3 regarding mechanical fault at control valves. Using the same argument as previous sections, this observation is chosen as it requires more expected effort as compared to Uncertainty Observation 4. Hence, a structured plan to track and address this uncertainty is important.

6.5.4.1. Uncertainty Observation 3: Mechanical Faults at Control Valves

The details of developing the action plan for Uncertainty Observation 3 are described in Annex I. Using the guide, we have developed the following track and address action plan:

Track Portion of the Action Plan. Tracking the uncertainty involves investigating the physical interfaces and joints at the control valves that receive the electrical signal from the BSCU. This is because these physical interfaces represent the location where electrical signals are transmitted between the BSCU and the control valves to control the valve operation. Mechanical faults such as dry or sticky joints may result in signal corruption.

To do that, dedicated mechanical engineers are needed to conduct the investigation. The mechanical engineers should be competent in analysing physical and mechanical properties of WBS operation

The mechanical engineers are responsible for investigating the mechanical contacts where the BSCU signals reach the control valves and checking on the reliability and failure rate of these mechanical contacts. This can be a paper study by reviewing the specification of the mechanical contacts. It is unsure if lab test or additional contractor test result/specification is needed to verify the design performance of the mechanical contacts. This will incur more resources and expertise.

A structured management oversight appointed by the management needs to be in place to ensure data is feedback to the safety team for analysis. Regular update is expected at the weekly safety meeting.

Address Portion of the Action Plan. The uncertainty will trigger a response when there is a discovery of mechanical conditions that can potentially compromise the electrical signal between the BSCU and the control valves and lead to signal corruption. Discovery of specific control valves that may be subjected to the above mechanical condition will also trigger a response from the investigation team.

The management safety committee shall be responsible for deciding upon the respond actions. When trigger points are met, the mechanical engineers shall escalate result directly to the management via incident report. A review committee that includes project manager and mechanical engineers shall be responsible to review any changes that will be put in place to address the uncertainty. The review committee should have knowledge about mechanical engineering and WBS operation. It should also have accessed to higher management that has the authority to review WBS operation.

If no data is observed within a month, the project manager needs to decide on a different action plan if the uncertainty is potentially still safety-critical. Depends on the findings regarding the mechanical contacts, the project manager needs to decide on the urgency to escalate the issue to higher management. There may potentially be a review of the mechanical properties of the control valve depending on the investigation.

Possible Actions when Trigger to Address Uncertainty. Taking reference from section 4.4.3, if the risk due to uncertainty about the mechanical faults at control valves is still not confidently within the tolerable or acceptable region, a decision would have to be made if one or both of the following needs to be conducted:

- **Focus on Risk through System Change**. If the assessed risk has become unacceptable, stakeholders need to derive strategies to reduce the risk. One approach could be to modify the existing mechanical valve with either a more reliable valve or a different joint with better mechanical contact between the BSCU and the control valve. Stakeholders would have to conduct the safety analysis again to conclude if the new measure or mitigation does lower the risk to a tolerable or acceptable level.

- **Focus on Knowledge through Uncertainty Clarification.** The T.A.G. tracking on the epistemic uncertainty may need to continue under certain scenarios. The following are two possible scenarios to solicit more information to clarify the epistemic uncertainties.
    - **Scenario 1 – Increasing Confidence.** Even when the assessed risk is tolerable, there may still not be enough confidence in the assessment due to the residual uncertainties. For example, the investigation team may have received assurance from the contractors providing the control valve that the reliability of the mechanical contact is within tolerance. However, the contractors are not able to provide the supporting test result in time since it has been generated by another agency. The investigation team may request for more time to solicit the test result so as to verify the reliability of the mechanical contact.
    - **Scenario 2 – Reducing Risk.** In the second scenario, clarifying certain epistemic uncertainties may directly help to reduce assessed risk. For example, after investigating the mechanical contact at the control valve, it has been discovered that dry joint may be an issue due to environmental condition (e.g. hot and humid weather). As a result, more investigation needs to be conducted to solicit information about the environmental conditions surrounding the airport. Hence, there is a need to continue and track the uncertainty as more information is needed to lower the system risk into the tolerable region.

To end the uncertainty tracking, stakeholders must be satisfied that the uncertainty due to mechanical faults at control valves being tracked by the T.A.G. approach has either been eliminated or that the risk is assessed to be confidently acceptable despite the decision to stop tracking the uncertainty.

6.6. Initial Evaluation and Summary

Like chapter 5, we have highlighted in this chapter the contributions specifically from applying the T.A.G. approach in the FTA during PSSA and FMEA during SSA. Using the T.A.G. approach, we have managed to discover four plausible-but-uncertain causal relationships across different phases of the safety assessment process and over a spectrum of primary causal conditions. Based on criticality and expected efforts to manage the uncertainties, the uncertainty observations are compared and prioritised using the prioritisation matrix chart.

Stakeholders can then make use of the chart to decide how much to focus on each uncertainty so as to allocate the proportionate level of resources to manage them. Note that the four uncertainties may not be plotted on the same prioritisation matrix at the same time since they are discovered at different phases of the safety assessment process. After selecting the uncertainty to be tracked, an action plan to manage the uncertainty has been developed using the guided questions in the user guide.

Chapter 7 - Evaluation

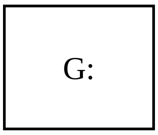
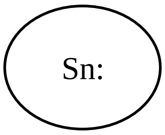
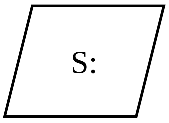
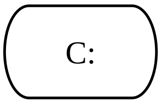
7.1. Introduction

For this thesis, the proposition was defined in section 1.3 as:

> *Epistemic uncertainties in the underlying models of safety assessments for safety-critical systems can be feasibly and systematically identified, documented and tracked through-life in order to enable intervention to address potential risk.*

In this chapter, we have systematically evaluated the above proposition as a claim supported by evidence using the Goal Structuring Notation (GSN) [133]. The GSN is a notation to represent arguments used specially in the safety domain. We have made use of four common elements under the notation (see Table 42).

Table 42. GSN Elements used in the Evaluation

	Goal or claim (e.g. the system is safe to operate in the stipulated conditions)		Solution or evidence referenced to support the goal (e.g. analysis and test results)
	Strategy or argument to support a goal (e.g. argument by showing that all identified hazards have been mitigated)		Reference to a contextual information or statement (e.g. policy where a strategy takes reference from)

7.1.1. Structured Argument

Using the GSN, a structured argument has been developed to describe how the above thesis preposition can be shown to be true based on evidence observed in this evaluation (see Figure 83).

7.1.2. Goal 1 (G1): Thesis Proposition

As shown in Figure 83, the top-level claim (G1) is the thesis proposition as stated in the introduction of this chapter. This claim is based on the principles that epistemic uncertainties should be identified (as explained in section 3.2.1), documented (as explained in section 3.2.2), tracked and addressed (as explained in section 3.2.3). These three principles (context C1, C2 and C3) serve to argue for the thesis proposition to manage epistemic uncertainties feasibly and systematically through-out the system lifecycle.

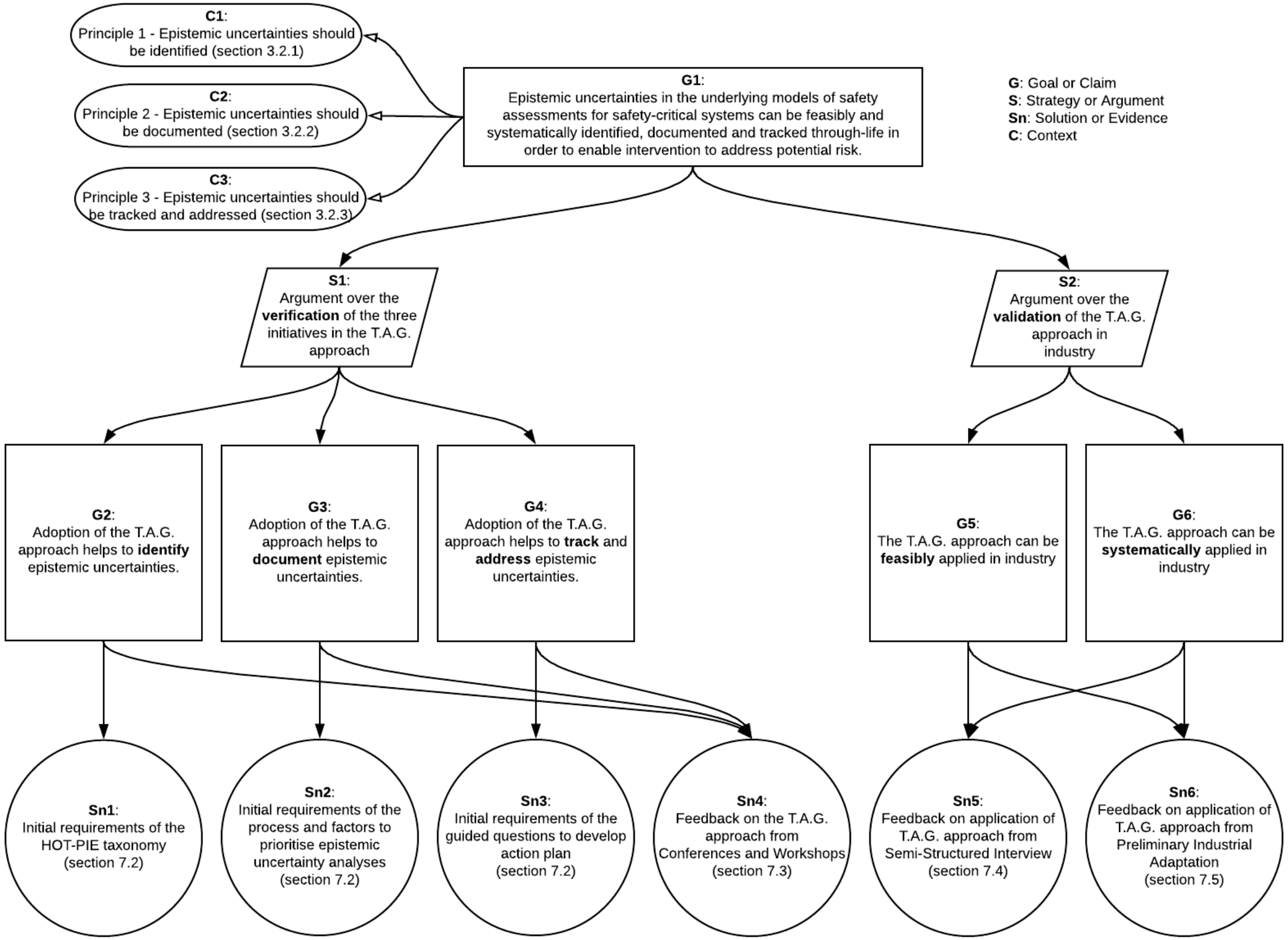

C1:
Principle 1 - Epistemic uncertainties should be identified (section 3.2.1)
C2:
Principle 2 - Epistemic uncertainties should be documented (section 3.2.2)
C3:
Principle 3 - Epistemic uncertainties should be tracked and addressed (section 3.2.3)
G1:
Epistemic uncertainties in the underlying models of safety assessments for safety-critical systems can be feasibly and systematically identified, documented and tracked through-life in order to enable intervention to address potential risk.
G: Goal or Claim
S: Strategy or Argument
Sn: Solution or Evidence
C: Context
S1:
Argument over the **verification** of the three initiatives in the T.A.G. approach
S2:
Argument over the **validation** of the T.A.G. approach in industry
G2:
Adoption of the T.A.G. approach helps to **identify** epistemic uncertainties.
G3:
Adoption of the T.A.G. approach helps to **document** epistemic uncertainties.
G4:
Adoption of the T.A.G. approach helps to **track** and **address** epistemic uncertainties.
G5:
The T.A.G. approach can be **feasibly** applied in industry
G6:
The T.A.G. approach can be **systematically** applied in industry
Sn1:
Initial requirements of the HOT-PIE taxonomy (section 7.2)
Sn2:
Initial requirements of the process and factors to prioritise epistemic uncertainty analyses (section 7.2)
Sn3:
Initial requirements of the guided questions to develop action plan (section 7.2)
Sn4:
Feedback on the T.A.G. approach from Conferences and Workshops (section 7.3)
Sn5:
Feedback on application of T.A.G. approach from Semi-Structured Interview (section 7.4)
Sn6:
Feedback on application of T.A.G. approach from Preliminary Industrial Adaptation (section 7.5)

Figure 83. Structured Argument to Evaluate Thesis Contributions using the GSN

Our evaluations shall focus on two types of activity: verification and validation. We refer to the definitions in ISO 9001:2015 [183] to explain the differences between the two activities. ISO 9001 is chosen as it is an international standard used by organisations to demonstrate that a product or service has met specific requirements. We would next explain the details of both activities in support of G1.

7.1.3. Strategy 1 (S1): Verification

According to the standard, verification activities are conducted to "*ensure that the design & development output meets the input requirements (functional requirements & specifications)*". In other words, verification is about asking the question "are we building the thing right?" Our primarily focuses are on the concepts that have been defined and the T.A.G. approach that has been developed to help safety analysts to manage epistemic uncertainties. Based on the top-level goal G1, we have derived three 2nd level goals to verify the T.A.G. approach:

1. **G2**: Adoption of the T.A.G. approach helps to identify epistemic uncertainties
2. **G3**: Adoption of the T.A.G. approach helps to document epistemic uncertainties
3. **G4**: Adoption of the T.A.G. approach helps to track and address epistemic uncertainties

We use two types of evidence to evaluate strategy 1:

1. **Sn1-3**: **Evaluation against Initial Requirements**. For this argument, we use the requirements of the three initiatives under the T.A.G. approach to evaluate if the 2nd level goals have been achieved. Each of the 2nd level goals (G2, G3 and G4) has a corresponding evidence statement (Sn1, Sn2 and Sn3 respectively). We have reproduced the three requirements in Table 43. This will be discussed in section 7.2.

Table 43. 2nd level Goals under Verification

Sn.	Initial Requirements (extracted from section 4.1)
1	Taxonomy of causal mechanisms for stakeholders to recognise epistemic uncertainties not covered in original assessment
2	Process and factors to prioritise epistemic uncertainties analysis so as to document the findings in existing safety assessment technique
3	Method to develop actionable goals to track epistemic uncertainties through-life and address when some thresholds are met

2. **Sn4**: **Peer Review from Conferences and Workshops**. Besides the theoretical assessment, we have also solicited feedback through peer reviews from internal and

external sharing of the T.A.G. approach as a mean to verify the 2nd level goals (Sn4). This will be discussed in section 7.3.

7.1.4. Strategy 2 (S2): Validation

According to ISO 9001:2015, validation activities are conducted to "*ensure that the resulting products and services meet the requirements for the specified application or intended use (customer needs)*". In other words, validation is about asking the question "are we building the right thing?" For validation, our focus is about soliciting feedback on the actual application to introduce the T.A.G. approach from this research into real-life usage. Based on the top-level goal G1, we have derived two 2nd level goals to validate the T.A.G. approach:

4. **G5**: The T.A.G. approach can be **feasibly** applied in industry. In other words, this goal considers the viability of applying the T.A.G. approach in the context of existing industrial practice.

5. **G6**: The T.A.G. approach can be **systematically** applied in industry. In other words, this goal considers if the T.A.G. approach is implemented according to a certain fixed plan (based on certain principles or concepts) that is repeatable in the industry.

We use two types of evidence to evaluate the above 2nd level goals:

1. **Sn5**: **Peer Review from Semi-Structured Interviews**. We have solicited feedback through semi-structured interviews with safety analysts in the Singapore defence to validate both of the 2nd level goals. Semi-structured interview comprises a pre-determined set of open-ended questions that guides the discussion and yet maintains the flexibility for the discussion with the participants to branch out into other areas related to the management of epistemic uncertainties. Through an open-ended interview, the participants have been able to share their perspectives on the feasibility and systematic application of the T.A.G. approach. This will be discussed in section 7.4.

2. **Sn6**: **Preliminary Industrial Application**. Besides the semi-structured interview, we have also solicited feedback on the initial adaptation of the T.A.G. approach in the Singapore Air Force. This has provided more insights if the T.A.G. approach can be feasibly and systematically applied in the Singapore defence industry. This will be discussed in section 7.5.

The four types of evaluating approach against the thesis proposition in this chapter is summarised in Table 44.

Table 44. Types of Evaluating Approach

Section	Evaluating Approach	Verification	Validation
7.2	Evaluation against Initial Requirements	x	
7.3	Peer-review: Conferences / Seminars	x	
7.4	Peer-review: Semi-structured interview		x
7.5	Preliminary Industrial Adaptation		x

7.2. Evaluation against Initial Requirements

In this section, we shall verify if the theoretical concepts and structured approach that have been developed in this research are able to support the three 2nd level goals. These three goals are:

1. **G2**: Adoption of the T.A.G. approach helps to identify epistemic uncertainties This would be discussed in section 7.2.1.

2. **G3**: Adoption of the T.A.G. approach helps to document epistemic uncertainties. This would be discussed in section 7.2.2.

3. **G4**: Adoption of the T.A.G. approach helps to track and address epistemic uncertainties. This would be discussed in section 7.2.3.

7.2.1. Verify Contributions to Identify Epistemic Uncertainty

The HOT-PIE taxonomy has been developed under the first initiative of the T.A.G. approach to better identify epistemic uncertainty. As elaborated in section 4.2, this is possible because of the following reasons:

1. **HOT-PIE taxonomy helps to search more effectively**. The taxonomy of causal mechanisms that we have created harness prior knowledge about credible causal mechanisms, especially those that may reveal certain type of unsafe situations related to the system of interest. It specifies a reasonable coverage of diverse issues to help stakeholders recognise a wide range of safety-critical concerns and causal relationships. To create a credible taxonomy, literature review of subjects that are related to safety was conducted (see Annex A for the range of literature surveyed). These causal mechanisms covered a wide range of topics such as system safety, human factor ergonomic, project uncertainty, taxonomy of safety-related subjects and situational awareness. The use of the taxonomy encourages stakeholders to shift the boundary of knowledge from not knowing to knowing about epistemic uncertainty, through surfacing and modelling of previously unknown uncertainty.

2. **HOT-PIE taxonomy helps to search more efficiently**. Without a systematic way of searching, the process may end up being laborious and inefficient. Our taxonomy helps the stakeholders to narrow their search by grouping the causal mechanisms into primary and secondary causal conditions for more efficient searching. This is explained with greater details in section 4.2. The list of causal mechanisms is classified into six primary causal conditions: Human, Organisation, Technology, Process, Information and Environment.

The approach of using the HOT-PIE taxonomy has been demonstrated using two case studies focusing on system viewpoints in chapter 5 and component viewpoints in chapter 6 respectively.

In chapter 5, we have highlighted the use of the HOT-PIE taxonomy with the STPA safety assessment techniques under the Yongwen railway accident case study. With the help of the taxonomy, we could recognise four plausible-but-uncertainty causal relationships that were not discovered by the original safety assessment. The newly identified uncertainties are listed in Table 32. Two of these are known uncertainties that form new linkage among nodes with existing causal relationship and two are unknown uncertainties that form new linkages between nodes that do not have prior relationships with each other (see section 5.4.2 and section 5.4.3). Observation 3 on the uncertainty associated with software algorithm can even plausibly lead to the discovery of the error in the track signalling system that causes the accident. In chapter 6, we have highlighted the use of the HOT-PIE taxonomy with the FTA and FMEA safety assessment techniques under the ARP 4761 aircraft design case study. With the help of the taxonomy, we could recognise four plausible-but-uncertainty causal relationships that were not discovered by the original safety assessment. The newly identified uncertainties are listed in Table 38 and Table 40. Three of these are known uncertainties that form new linkage among nodes with existing causal relationship and one is an unknown uncertainty that form new linkages between nodes that do not have prior relationships with each other (see section 6.4.2 and section 6.6.2).

In conclusion, we have demonstrated that the HOT-PIE taxonomy in the T.A.G. approach can recognise known and unknown uncertainties not covered in the original safety assessment. This verifies that the first initiative under T.A.G. approach can help to identify epistemic uncertainties.

7.2.2. Verify Contributions to Document Epistemic Uncertainty

Process and factors to prioritise epistemic uncertainties have been developed under the second initiative of the T.A.G. approach to better document epistemic uncertainty. As elaborated in section 4.3, this is possible because of the following two reasons:

1. **Generic process helps integrate T.A.G. approach into any safety assessment techniques**. Section 3.2.2.1. mentioned that there is no established approach of documenting uncertainties in current safety assessment techniques. With the generic process flow developed in section 4.3.1, we can counter this challenge by complementing any safety assessment techniques with the T.A.G. approach so as to document the uncertainties and the corresponding uncertainty analysis.

2. **Prioritisation Factors helps to improve traceability when prioritising uncertainty analyses**. With limited resources, stakeholders may discard the less certain causal relationships during safety assessment (see section 3.2.2.2) and not document the reasons they choose or discard causal relationships for analysis during safety assessment (see section 3.2.2.3). To be explicit about such selections, we have defined two major factors: Criticality and Expected Effort, that represent what stakeholders can consider when prioritising which uncertainty analyses they would manage first. These factors have been defined in section 4.3.2.

Like the previous section, the approach of using the process and prioritisation factors has been demonstrated using two case studies focusing on system viewpoints in chapter 5 and component viewpoints in chapter 6 respectively.

In chapter 5, we use the process and prioritisation factors to document and prioritise the uncertainties analyses under the Yongwen railway accident case study. Using the generic process of tagging uncertainties from section 4.3, we have constructed a similar process flow to tag epistemic uncertainties in STPA in section 5.3. This process flow is shown in Figure 61. We develop this process flow to show that the T.A.G. approach is not a separate standalone method to conduct safety assessment but rather it introduces additional steps to complement the existing STPA techniques.

With the help of the prioritisation factors, we could prioritise the 4 plausible-but-uncertainty causal relationships that were not discovered from the original safety assessment. This is conducted with the help of the guided questions from the T.A.G. approach User Guide (see Annex K).

In chapter 6, we use the process and prioritisation factors to document and prioritise the uncertainties analyses under the ARP 4761 aircraft design case study. Using the generic process of tagging uncertainties from section 4.3, we have constructed similar process flows to tag epistemic uncertainties in FTA (see section 6.3) and FMEA (see section 6.5). These process flows are shown in Figure 70 and Figure 76. Again, it shows that the T.A.G. approach can be integrated into existing safety assessment techniques.

With the help of the prioritisation factors, we could prioritise the 4 plausible-but-uncertainty causal relationships that were not discovered from the original safety assessment. Again, this is conducted with the help of the guided questions from the T.A.G. approach User Guide (see Annex K).

In conclusion, we have demonstrated that the process and factors to prioritise epistemic uncertainties analyses in the T.A.G. approach can be integrated in existing safety assessment techniques. This verifies that the second initiative under T.A.G. approach can help to document epistemic uncertainties.

7.2.3. Verify Contributions to Track and Address Epistemic Uncertainty

The guided questions have been developed under the third initiative of the T.A.G. approach to track and address epistemic uncertainty. As elaborated in section 4.4, this is possible because of the following two reasons:

1. **Guided questions help to develop actionable goals to track uncertainty**. To do that, we have developed guided questions based on specific factors in section 4.4.2. to help track the uncertainty. This is based on the GQTA approach as explained in section 4.4.1. This ensures that tracking of epistemic uncertainties is not random, and it follows a structured approach that spell out the activities to be conducted as part of the tracking.

2. **Guided questions help to develop actionable goals to address uncertainty**. We have also developed guided questions based on specific factors to address the uncertainty in section 4.4.2. These would help safety analysts to response to changes in the epistemic uncertainties when certain thresholds with regards to the uncertainties are triggered.

Like the previous section, the approach of using the guided questions has been demonstrated using two case studies focusing on system viewpoints in chapter 5 and component viewpoints in chapter 6 respectively.

In chapter 5, we use the guided questions to develop action plan so as to track and address the epistemic uncertainties under the Yongwen railway accident case study. The action plan is formulated by responding to a series of questions using the T.A.G. approach user guide (see Annex K). An example of the responses to the guided question is attached in Table 36 of chapter 5. The subsequent track and address action plan derived from the responses is elaborated in section 5.3.5.

In chapter 6, we use the guided questions to develop action plan so as to track and address the epistemic uncertainties under the ARP 4761 aircraft design case study. Like the Yongwen railway accident case study, the action plan is formulated by responding to the same set of questions in the T.A.G. approach user guide. An example of the responses to the guided question is attached in Annex G. The track and address action plans derived from the responses are elaborated in section 6.4.4. and 6.6.4

In conclusion, we have demonstrated that the guided questions in the T.A.G. approach can develop actionable goals. This verifies that the third initiative under T.A.G. approach can help to track and address epistemic uncertainties through-life.

7.3. Peer Review – Sharing at Conferences and Workshops

The research from this thesis was presented to scholars, researchers and engineers both from academia and industry. In the University of York, the research was presented internally at the High Integrity Systems Engineering (HISE) Group workshops where valuable comments and suggestions were solicited to refine the T.A.G. approach. The HISE Group is involved with research and teaching in systems and software engineering, especially in safety and security-critical applications.

Externally, we have published papers and presented at peer-reviewed international conferences, namely:

- **Causal Reasoning for Embedded and Safety-critical Systems Technologies – CREST 2017** (29 Apr 2017, Sweden). We presented the research at the 2nd International Workshop on Causal Reasoning for Embedded and Safety-critical Systems Technologies (CREST 2017) at Uppsala, Sweden on 29 April 2017. This is part of the European Joint Conferences on Theory and Practice of Software (ETAPS) for international academic and industrial researchers working on topics relating to software science in safety-critical systems. Since the focus was on causal reasoning, the concept of representing hazards as causal relationships generated interesting discussions, especially on the graphic representations of cause-and-effect relationships. The CRMSA was discussed and refined based on the comments from the workshop regarding the ways to better represent causal reasoning in safety.

- **35th International System Safety Conference** (21-25 Aug 2017, New Mexico). This conference is among the premiere international conferences on system safety with the theme “Pushing the Boundaries of System Safety”. The conference had attracted participants from both academia and industry focusing on challenges in safety assessment for systems of system. The T.A.G. approach was presented during the conference and the participants provided valuable suggestions to help concretise the approach. For example, the idea of providing a quantitative scoring to the prioritisation factor was generated during a discussion at the conference. There were differing views among the participants if higher priority should be given to critical uncertainty that requires more efforts or less critical uncertainty that requires less efforts. At the end of the discussion, the participants agreed that this can be context dependent.
- **12th International Conference on System Safety and Cyber Security** (30 Oct- 1 Nov 2017, London). The conference was organised by the Institution of Engineering and Technology (IET) as a unique conference in the UK where both safety and security engineers can network and share insights between them. Our research was presented during the conference. One of the interesting discussions was on the possibility of applying the T.A.G. approach in the security domain to identify active agents that intend to do harm to the system. While the participants see value in porting the T.A.G. approach into the security domain, the taxonomy would have to be customised from the security perspective.
- **System Safety Society (Singapore Chapter)** (19 Jan 2018, Singapore). A presentation of our research was conducted at the Singapore chapter of the International System Safety Society to share with local academics and industry system safety practitioners. While the participants acknowledge the usefulness of the T.A.G. approach, the complexity of the approach generated much concern, especially if the approach is implemented at a larger scale or with greater heterogeneity across an organisation. This is also a common observation in the semi-structured interview (see section 7.4) that we have conducted with system safety analysts working in the Singapore defence industry.
- **Singapore Aerospace Technology and Engineering Conference – SATEC 2018** (7 Feb 2018, Singapore). Singapore Aerospace Technology and Engineering Conference (SATEC) is a biannual aviation conference that is held in conjunction with the Singapore Air show. It is where researchers, aviation operators and engineers present and discuss key developments and advancements in aerospace technology and

engineering. The key focus for this year's conference is on the innovative use of technology and engineering to meet new and evolving challenges in the aerospace industry. Like the IET conference in London, there were suggestions to port the T.A.G. approach into cyber security. Again, the participants acknowledge that the approach of managing uncertainties can possibly be applied to cyber security but the list of causal conditions in the HOT-PIE taxonomy needs to be reviewed. While the six primary categories in the taxonomy are relevant for cyber security, the lower level causal conditions would need to be reviewed as it was developed based on literature surveyed. A review of the literature concerning cyber security is needed to make the taxonomy more encompassing.

7.4. Peer Review – Semi-Structured Interview

This section describes the result of the interview with industrial participants to validate the T.A.G. approach so as to complement the verification activities conducted in the previous sections. We first describe the methodology that shapes the interview in section 7.4.1, before we present the interview analysis in section 7.4.2.

7.4.1. Interview Methodology

As part of the interview methodology, we present the objectives, structure, approach of collecting and analysing the data and the profile of the participants in this section. This methodology guides us in soliciting trustworthy data that is systematically collected and ensure that it comes from participants who are subject matter experts in our area of research.

7.4.1.1. Objectives

The interview supports the validation activities with the following objectives:

- Solicit feedback if the T.A.G approach can be feasibly applied in industry
- Solicit feedback if the T.A.G approach can be systematically applied in industry

7.4.1.2. Structure of Questionnaire

The interview is guided by a questionnaire (see Annex J) with five sections that comprises both quantitative and qualitative questions. The methodology of preparing the questionnaire is based on two key questions: deciding what to ask and designing how to ask.

In terms of deciding what to ask, we have structured the questionnaire into five sections: participants' details, uncertainty in safety assessment, HOT-PIE taxonomy, prioritisation factors and guided action plans. The participants' details are to help us in finding possible correlation between the eventual feedback to the profile of the participants. The uncertainty in safety assessment portion is to solicit feedback on the current situation regarding managing uncertainty during safety assessment. The remaining three sections are aligned to the three steps under the T.A.G. approach: identify, document, track and address. As for designing how to ask, we have taken recommendations from Oppenheim [184] to design our questionnaire to be exploratory and in depth. The intent for each section is as follows:

Section 1: Participants' Details. This section records the profile of the participant in terms of the job responsibility and the level of experience in safety assessment.

Section 2: Uncertainty in Safety Assessment. This section investigates the impact of epistemic uncertainty on existing safety assessment and how well it is being managed. This would provide the demand to manage epistemic uncertainty better in safety assessment.

Section 3: HOT-PIE Taxonomy. This section solicits feedback on step 1 of the T.A.G. approach to recognise known and unknown uncertainties not covered in the original safety assessment. The three quantitative questions are scoped to focus on the possibility, usefulness and feasibility of integrating the HOT-PIE taxonomy with existing safety assessment techniques. The participants are also provided with a message box to provide qualitative feedback on the HOT-PIE taxonomy.

Section 4: Prioritisation Factors. This section solicits feedback on step 2 of the T.A.G. approach that helps to document and prioritise uncertainty analyses in existing safety assessment. The three quantitative questions are scoped to focus on the importance, usefulness and feasibility of incorporating the prioritisation with existing safety assessment techniques. The participants are also provided with a message box to provide qualitative feedback on the prioritisation factors.

Section 5: Action Plan. This section solicits feedback on step 3 of the T.A.G. approach that provides guided questions to develop goals to track and address epistemic uncertainty in existing safety assessment. The two quantitative questions are scoped to focus on the usefulness and feasibility of implementing the guided questions with existing safety assessment techniques. The participants are also provided with a message box to provide qualitative feedback on the guided questions.

7.4.1.3. Profile of Participants

The semi-structured interview targets individuals who are experience in safety analysis under the domain of aviation and defence systems management. It is hope that, by having the experiences in utilising safety assessment techniques and having been through the hazard analysis process, the participants would be able to provide insightful comments and suggestions to improve the T.A.G. approach.

Since the participants can only be selected from a niche pool of individuals, we set our target to engage at least 5 participants. In the Singapore air force, the complicated safety-critical system that we are concerned with is known as system-of-systems. Participants that have experiences in safety assessment with system-of-systems would either be military personnel from the air force or civilian engineers from the defence agency supporting the acquisition and operationalisation of the system. In total, seven participants were interviewed using the semi-structured interview approach. Their profiles are summarised in Table 45. All of them have at least 5 to 10 years of experience in safety.

Table 45. Summary of Participants

No.	Job Requirement related to Safety	Experience in Safety (Years)	Industry
1	SoS Safety Governance	10-15	Air Force
2	System Safety Engineer	10-15	Defence Agency
3	Senior Engineer (System Safety)	5-10	Defence Agency
4	System Safety (Airborne Systems)	5-10	Defence Agency
5	System Safety Engineering Lead	>15	Defence Agency
6	Safety Assessment for SoS	>15	Air Force
7	Networked Safety and Engineering	10-15	Air Force

7.4.1.4. Data Collection Approach

This section discusses the approach to collecting data from the interview so as to meet the objectives stated in section 7.4.1.1.

Inclination towards Qualitative Data. According to Punch [185], quantitative data is about numbers (or measurements) and qualitative data is about data not in the form of numbers (usually in the form of words). Qualitative research is targeted to gain insights into a problem and look for underlying trends from the opinions of a selected group of participants. Quantitative research tends to generalise numerical data using some statistical calculation collected from a large sample.

Since the interview involves seven participants, we do not have sufficient sampling size to make conclusive deduction from the quantitative data. In order to appreciate if the T.A.G. approach can be feasibly and systematically complementing existing safety assessment techniques, it is important to solicit the opinions from the participants based on their experiences. Hence, a qualitative approach is more appropriate in this case where we will be highlighting considerations, observations and challenges from the participants. While there are quantitative data collected, the interview has mostly been focusing on soliciting qualitative data.

Semi-Structured Interview to collect Qualitative Data. The one-to-one interview is a type of semi-structured interview as it comprises a pre-determined set of open-ended questions that prompt the discussion and yet maintains the flexibility for the discussion with the participants to branch out into areas related to the management of epistemic uncertainties. The participants had been provided with a case study and asked a set of predefined questions related to the case study. The case study is based on a research by Song et al [24] from Beihang University, Beijing. It consists of two parts. In part 1, we provide the background to the Yongwen railway accident and the STPA hazard analysis technique. The hazard analysis has been conducted on the Yongwen railway system by Song. In part 2, we introduce the T.A.G. approach and apply it on the STPA hazard analysis from part 1 to help manage epistemic uncertainties.

Refinement of Questions in the Semi-Structured Interview. The questions in the evaluation form is design based on internal discussions among the researchers and in accordance with the recommendations for questionnaire design described by Oppenheim [184]. The questions can be divided into four main areas. The first area focuses on the impacts of uncertainty on safety assessment; while the remaining three areas are dedicated to each of the three initiatives under the T.A.G. approach. The number of questions is optimised such that there are not too many to make the interview too laborious and yet about to sieve out critical opinions from the participants.

User Guide as a Reference. To complement the interview, a concise user guide that explains the concepts and methods under the T.A.G. approach is produced and attached in Annex K. The user guide can be used during the interview as a quick reference to retrieve information such as the HOT-PIE taxonomy and guided questions to develop the action plan. Moving forward, the user guide would be refined to serve as a document for safety analysts to refer to when they apply the T.A.G. approach during safety assessments.

Length of Interview. Being a qualitative approach, the semi-structured interview is expected to take a longer time compared to a quantitative approach. Summarising the interview has also been harder, compared to a quantitative one. In our semi-structured interview, we have specially used a thematic analysis approach to summarise the findings. This is discussed in detail under section 7.4.1.5. While we have scheduled an hour for each interview, most of the interviews lasted between 1.5hr to 2hrs. The participants were passionate to share on the topics of safety assessment and candidly provide their opinions in managing hazards and uncertainties during safety assessment. Since the semi-structured interview is open-ended, some discussions tend to take some time depending on the flow of the conversations. As some of the discussions are sensitive due to the nature of the projects that the participants were involved in, voice recorder was not used. Instead, notes and insights relevant to the research were hand written. This inadvertently increase the length of the interview as compared to using a recorder.

7.4.1.5. Data Analysis Approach

This section discusses the way that the data collected from the interview is being analysed. As mentioned in section 7.4.1.4, we conduct semi-structured interview to solicit considerations, observations and challenges highlighted by the participants. From these pointers, we group the common ones together to produce themes. Such a thematic analysis approach helps to sieve out important issues that get surfaced from multiple interviewees. According to Braun and Clarke [186], a thematic analysis typically comprises the following steps (see Table 46).

Table 46. Thematic Analysis Approach

No.	Step	Descriptions (directly extracted from Braun and Clarke)
1	Familiarisation with Data	This step requires the researcher to actively transcribe the data during the interactions. The researcher needs to understand the content of the interaction and be familiarised with all aspects of the data.
2	Generating Initial Codes	Once familiar with the data, the researcher must then start identifying preliminary codes, which are the features of the data that appear interesting and meaningful. These codes are more numerous and specific than themes but provide an indication of the context of the conversation.
3	Searching for Themes	The third step in the process is the start of the interpretive analysis of the collated codes. Relevant data extracts are sorted (combined or split) according to overarching themes.
4	Reviewing Themes	A deeper review of identified themes follows where the researcher needs to question whether to combine, refine, separate, or discard initial themes. Data within themes should cohere together meaningfully, while there should be clear and identifiable distinctions between themes. A thematic 'map' can be generated from this step.
5	Defining and Naming Themes	This step involves 'refining and defining' the themes and potential subthemes within the data. Ongoing analysis is required to further enhance the identified themes. The researcher needs to provide theme names and clear working definitions that capture the essence of each theme in a concise and punchy manner. At this point, a unified story of the data needs to emerge from the themes.
6	Producing the Report	Finally, the researcher needs to transform the analysis into an interpretable piece of writing with examples that relate to the themes, research question, and literature.

From the semi-structured interview, a thematic map was developed based on the above steps under the thematic analysis approach. The insights from the thematic mapping is reported as part of the qualitative analysis in section 7.4.2.

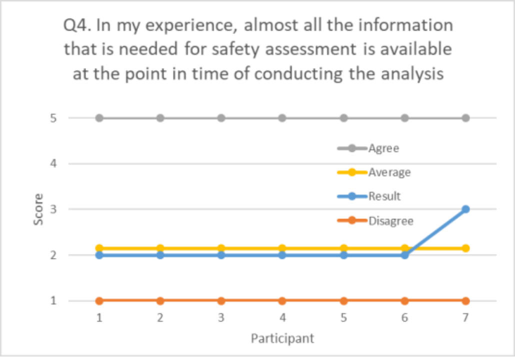
7.4.2. Interview Analysis

As explained in section 7.4.1.4 the interview focuses primarily on the qualitative data using the thematic analysis approach. In addition, we have also collected quantitative data that showcases average responses among the participants. Since these behavioural responses are subjective, we acknowledge that the quantitative result is not an exact science but more of a reference to the spread of responses among the participants. These quantitative observations are collated in section 7.4.2.1. Using the thematic analysis approach in section 7.4.1.5, the qualitative observations are presented in section 0.

7.4.2.1. Quantitative Observations

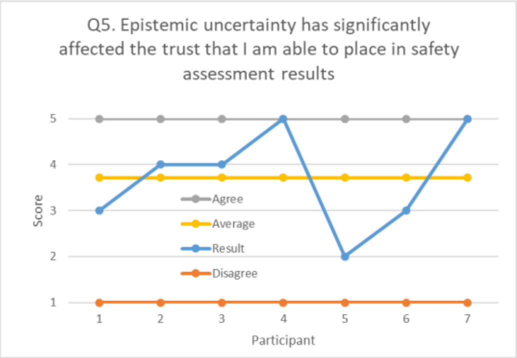
In this section, we present the quantitative data extracted from the semi-structured interview. The questions are spread across the questionnaire (see Annex J) provided to each of the participants. It is based on a 5-points Likert scale [187] to measure the participants' responses to a variety of statements according to the four focus areas defined in section 7.4.1.2. While the Likert scale is easy to construct and simple to complete for the participants, it does have its limitations. According to Bertram [188] and Subedi [189], participants of the Likert scale may be influenced by bias such as "*avoiding extreme responses at both end of the scale (central tendency bias), agreeing with statements as presented in order "please" the experimenter (acquiescence bias), and portraying themselves in a more socially favourable light rather than being honest (social desirability bias)*". While we acknowledge that the quantitative data is subjective, it serves as a reference of the participant inclinations.

<u>Focus Area 1: Uncertainty in Safety Assessment</u>

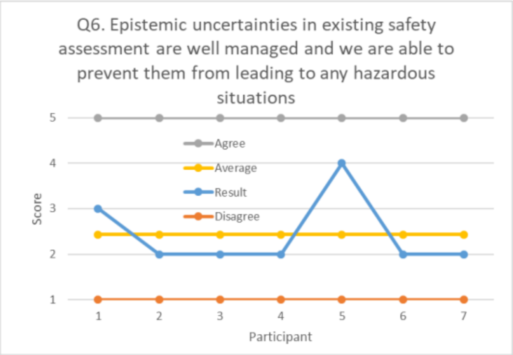


Most of the participants disagree that the information needed for safety assessment is readily available during the analysis. Hence, lack of information is common in safety assessment.

Since uncertainty is about lack of knowledge, the question should have been more precise by focusing on "*knowledge*", rather than "*information*". While information may not be equivalent to knowledge, a lack of information can potentially lead to a lack of knowledge. Hence, while it is not conclusive, the responses could imply that lack of knowledge (epistemic uncertainty) can potentially exist in safety assessment due to a lack of information.

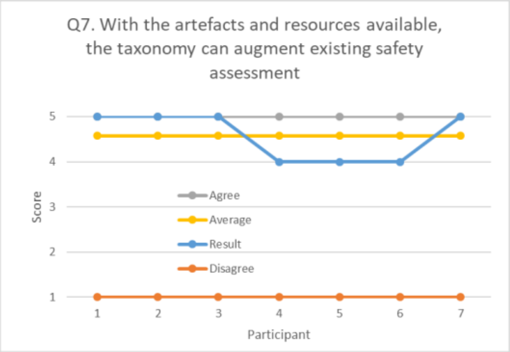

Q5. Epistemic uncertainty has significantly affected the trust that I am able to place in safety assessment results

With a mean of 3.7, the participants agree that epistemic uncertainty does affect their trust in the safety assessment results. However, there are fluctuating views among the participants as shown. Some commented that since they have '*done their best with what they know*', they are confident in trusting their assessment, even though there can be unknowns which they are not aware of.

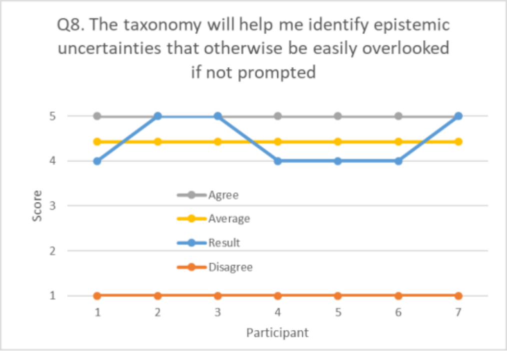

Q6. Epistemic uncertainties in existing safety assessment are well managed and we are able to prevent them from leading to any hazardous situations

With a mean of 2.4, the participants generally disagree that epistemic uncertainties in safety assessment are well managed. Most of them agree that more can be done to better manage such uncertainties during safety assessment. Some participants may feel more confident in managing epistemic uncertainties as the projects that they are involved in may not be too complicated or involved multiple systems.

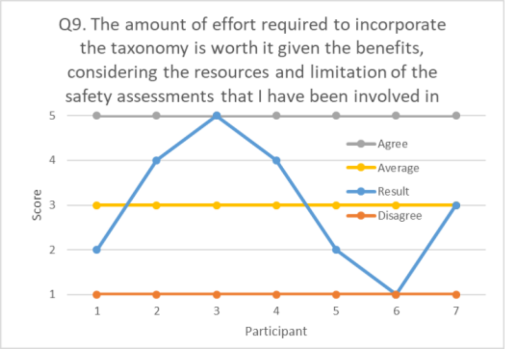
Focus Area 2: HOT-PIE Taxonomy


Q7. With the artefacts and resources available, the taxonomy can augment existing safety assessment

Most of the participants strongly agree that the HOT-PIE taxonomy can be augmented into existing way of conducting safety assessment if there is no resource constraint.

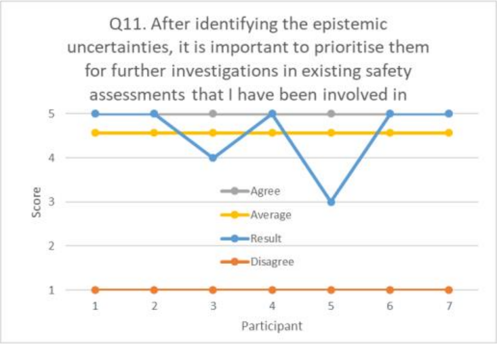


Most of the participants strongly agree that the HOT-PIE taxonomy can help to identify epistemic uncertainties which they may overlook if not prompted.

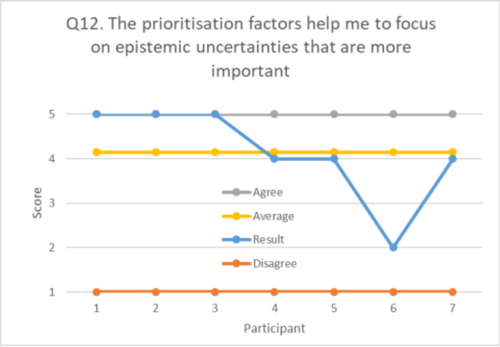


With a mean of 3, the participants have mixed opinions if the amount of efforts required to incorporate the taxonomy is practical for existing safety assessment. There are differing views among the participants on the practicality of augmenting HOT-PIE taxonomy as shown.

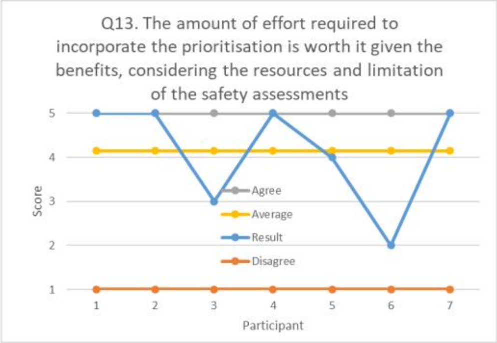
Focus Areas 3: Proritisation Factors



Most of the participants strongly agree that it is important to prioritise the further investigations needed on the epistemic uncertainties after identifying them.

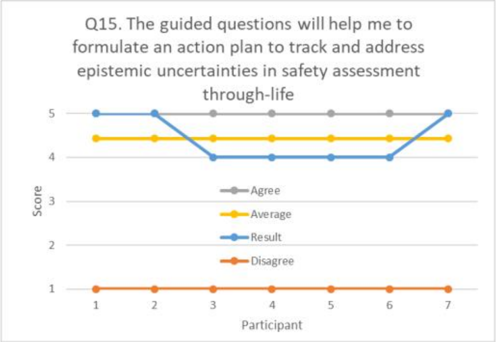


With a mean of 4.1, the participants agree that the prioritisation factors are useful to focus on epistemic uncertainties that are more important. One participant opined that the factors would only be helpful after a review by domain experts, which would depend on the context surrounding the system.

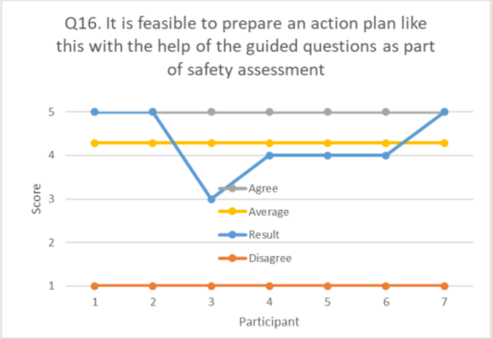


With a mean of 4.1, the participants generally agree that the amount of efforts required to incorporate the prioritisation factors is practical for existing safety assessment. Most of the participants feel that the prioritisation can be easily introduced into the current process when considering which epistemic uncertainties to analyse first.

Focus Area 4: Guided Questions for Action Plan



Most of the participants strongly agree that the guided questions can help them to formulate the action plan to track and address epistemic uncertainties through-life.



Most of the participants strongly agree that it is feasible to use the guided questions to formulate action plan as part of existing safety assessment.

In addition to the focus areas, we have also reviewed the feedback according to the experience and industry of the participants (i.e. taking reference from Table 45.).

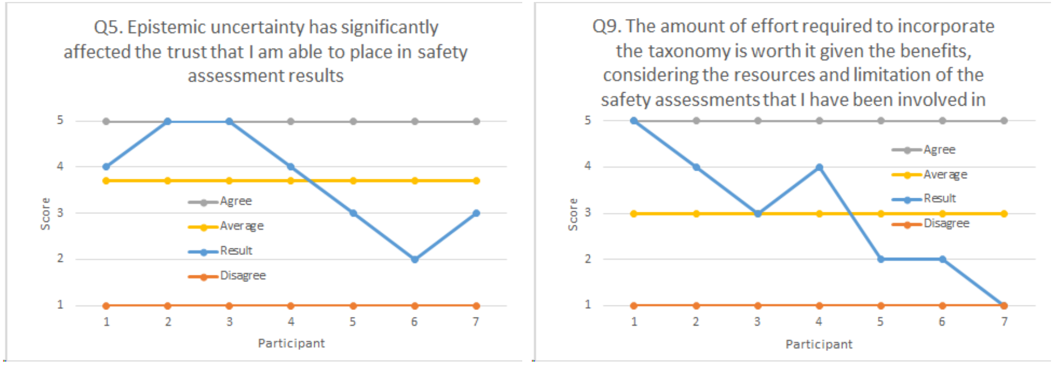


Figure 84. Trending Based on Experience

Experience. We have rearranged the 7 participants such that those with the least experience are on the left (5-10 years) and those with most experience are on the right (>15 years). Two of the questions show interesting observations: Q5 and Q9 (see Figure 84). Q5 chart indicates that more experience safety analysts seem to have less trust in the safety assessment due to the presence of epistemic uncertainty. Q9 chart indicates that more experience safety analysts are more pessimistic about incorporating the taxonomy into existing safety processes. One possible postulation is that the senior analysts may have seen many new initiatives in the past failing. Hence, they tend to be more sceptical about new approaches, even though they agree that epistemic uncertainty can affect confidence in safety assessment.

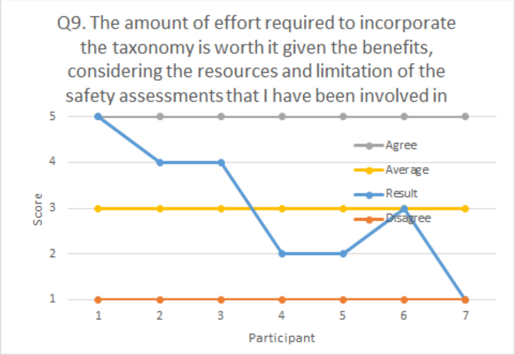


Figure 85. Trending Based on Industry

Industry. Next, we rearrange the participants by industry, with the 4 participants from Defence Agency on the left and 3 participants from the Air Force on the right. There is no obvious trending in most charts except Q9 (see Figure 85). Beside the more experience participants, it seems that the Air Force participants have less confidence in implementing the HOT-PIE taxonomy as compared to the Defence Agency. This is possibly due to the concern of balancing operational support in the air force, while still committing time and resources to carry out the uncertainty tagging using the HOT-PIE taxonomy.

7.4.2.2. *Qualitative Observations*

In this section, we present the observations based on the thematic analysis conducted after the semi-structured interview. To provide a holistic analysis, the observations have also considered the assessment from the quantitative data presented in the previous section. A total of over 30 themes and sub-themes have been identified. The thematic map constructed from the interview is shown in Figure 86. At the highest level, three top-level themes are defined to reflect the key focuses: present context, proposal and moving forward. Under these top-level themes, seven level-2 sub-themes are defined: epistemic uncertainty, trust, taxonomy, prioritisation, action plan, pilot implementation and targeted observations.

Top-level Theme 1: Present Context

In this theme, we consider the current landscape of managing epistemic uncertainty in safety assessment and how this can affect the trust in the assessment.

Epistemic Uncertainty

Uncertainty in Safety Assessment. There is a consensus among the participants that information needed for safety assessment is never complete and epistemic uncertainties are always present during their analysis. While participants are confident with the current safety assessment results, they feedback that this confidence is based on the belief that they have done their best, within their means and knowledge, to identify hazards based on existing techniques. However, the participants acknowledged that unknown uncertainty can still create that blind spot and make them unable to identify hazard. Such uncertainties do affect the trust in the safety assessment results and the participants are looking forward to managing these uncertainties better.

Lack of Problem Definition in Current Safety Assessment. The issue of epistemic uncertainty is especially pronounced in the current system level hazard analysis known as scenario-based hazard analysis. There is often a lack of credible scenarios to describe the ways constituent systems are interacting and operating with each other. This is because in early design phase, different systems have different level of maturity (some may not even be developed yet, while some are legacy systems); how different systems operate together as a complicated safety-critical system may not even be well understood yet. This can be an iterative process that takes time. Usually, the safety analysts can define the key mishap potentials based on past experiences from operating either the constituent systems independently or through references to other known mishaps. However, the key challenges are the causes of such mishaps, which are often not easily identified from the limited experiences in the scenarios that the systems will operate together.

Complementing Current Limitations. The sensing among the participants is that the T.A.G. approach is suitable for top-down hazard analysis, e.g. to complement STPA, brainstorming, system-level FTA; which is more flexible but rely heavily on experiences. It will help to prompt users to consider different scenarios so as to tackle the lack of problem definition in the safety assessment. As for more structured safety assessment techniques, e.g. FMEA, as such technique is less experience dependent, the T.A.G. approach may not be too applicable since the boundary of issues to be considered is more restricted.

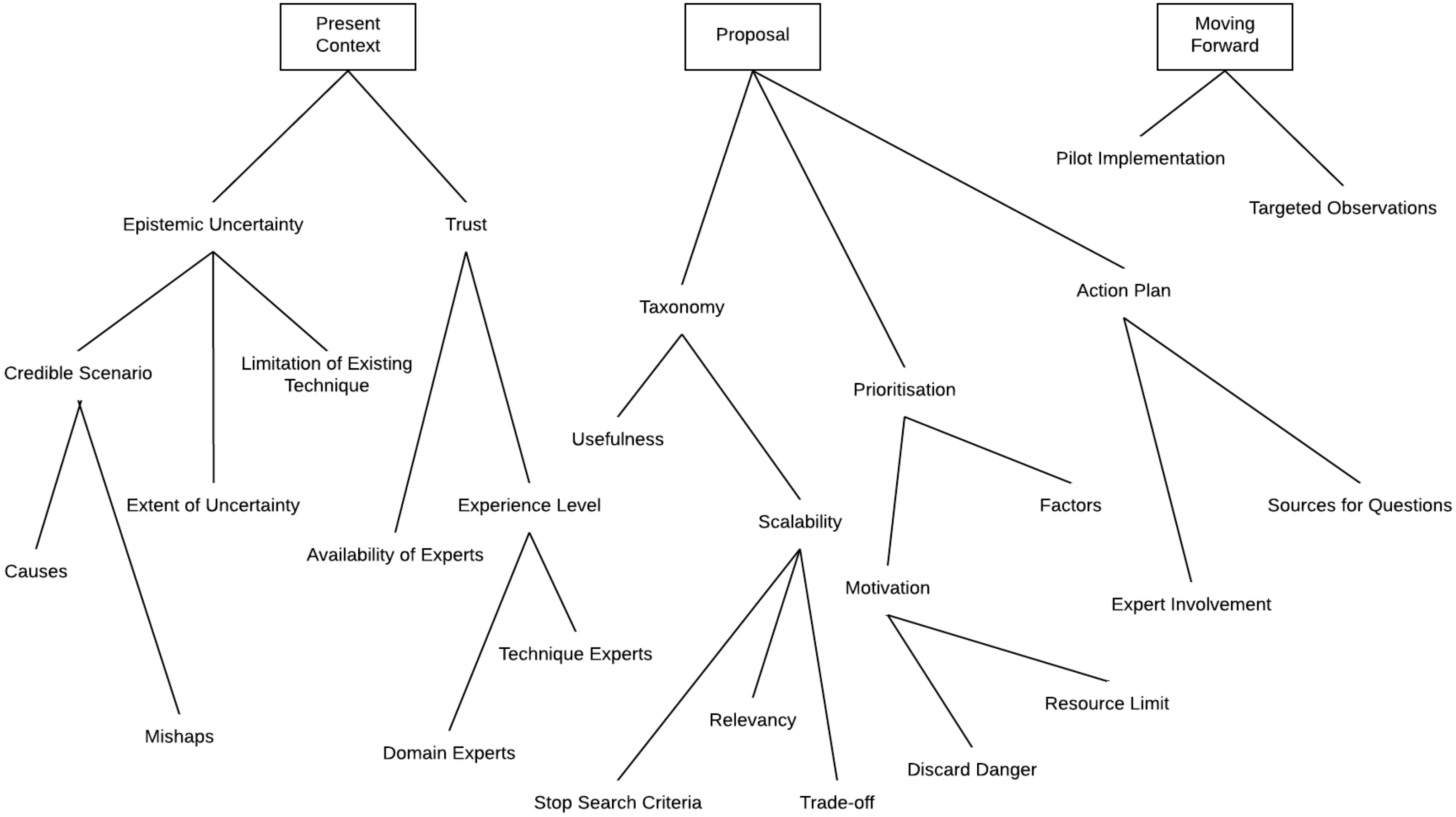


Figure 86. Thematic Map from Semi-Structured Interview

Trust

The trust in the safety assessment relies largely on the experiences of the safety analysts involved. From the interview, there are two types of experts that determine the trustworthiness of the safety assessment results. First, are the technique experts, who are the ones with experiences in using the safety assessment techniques to analyse and identify hazards. These are usually project managers and system safety engineers that can advise on **how** to use the safety assessment techniques. Second, are the domain experts, who are the ones with experiences in developing, managing and operating the system of interest. These are usually the operational manager, system engineers and system operators that can contextualise the safety assessment techniques and identify **what** are the causal conditions that can be potentially be hazardous.

Such experts are increasingly difficult to find in complicated systems as (1) we need people that not only know constituent system, but also at the integrated or system-of-systems level, (2) not easy to get all experts to be present during a safety assessment, some may send members of differing experiences. Hence, the HOT-PIE taxonomy is useful to complement the existing way of doing safety assessment when identifying hazards. Although we need to caution that, ultimately, we still need experts to decide what are hazardous, what are considered plausible with uncertainty and what are considered not plausible right from the start. While the responsibility to decide still lies with the experts, having a HOT-PIE taxonomy as a prompter to explicitly document the influence of uncertainties during safety assessment is important to help experts make a more informed safety decision.

Top-level Theme 2: Proposal

In this theme, we consider the feasibility of implementing the three initiatives under the T.A.G. approach. These are primarily the taxonomy, the prioritisation effort and the action plan.

Taxonomy

The participants felt that the HOT-PIE taxonomy is useful as it can provide a comprehensive view of potential causal mechanisms that may be hazardous. The diverse focus under the taxonomy can help to flag out issues that are not immediately apparent from existing way of conducting hazard analysis. This can help to break away from the current way of thinking about causes that can lead to the established mishaps.

While the taxonomy is systematically developed from a range of literature (see Annex A) and incorporated as a process to identify epistemic uncertainties, however, the key concern is the challenge of expanding and implementing the entire HOT-PIE taxonomy into current safety assessment process. As the taxonomy serves as a prompter (and not a compulsory checklist to follow), the participants questioned if it is necessary to refer to all the causal mechanisms listed in the taxonomy. This would determine when to stop searching through the list. Most participants are concerned that safety analysts do not have the time to consider the entire list of plausible causal mechanisms under the HOT-PIE taxonomy. Some of the participants also felt that not all causal mechanisms in the taxonomy are relevant for different domains and the list needs to be contextualised. For example, an organisation may have assumed that its people are already qualified and well trained for the tasks to operate the system of interest and so the safety concern about expertise or competency may not be considered under the taxonomy. While filtering the taxonomy in advance may be needed due to practical reason, it will limit the extent of the HOT-PIE taxonomy to prompt users about plausible safety-critical causal relationships.

Prioritisation

Some of the participants fed back that in the current stringent safety culture, there seem to be the expectations that any safety concerns identified must be analysed, and subsequently eliminated or mitigated whenever they are being discovered. Hence, there may not be the need to prioritise the analyses since all uncertainties must be managed immediately. However, this argument may not be entirely valid because:

1) In view of the pressure that ‘all identified safety concerns must be eliminated or mitigated’, safety analysts may subconsciously discard plausible causal relationships during the safety assessment as they may not have either sufficient evidence to determine the level of risk or the clarity to analyse the causal relationships due to epistemic uncertainties. As a result, safety decision may have been made without being aware of these plausible causal relationships that can be safety-critical.
2) In the context of large-scale social-technical systems with limited resources, as more and more systems begin to interoperate and integrate with each other, understanding the interactions among systems will become more and more complicated. Without a corresponding increase in resources such as manpower and time, the epistemic uncertainties that safety analysts have to deal with may eventually outweigh the resources available to analyse all such uncertainties. Such evolving landscape can eventually place a practical demand such that it is necessary to prioritise the efforts needed to analyse the epistemic uncertainties associated with the safety concerns.

The added dimension of considering expected effort is well received by most of the participants as they agreed that explicitly considering the efforts based on different factors can help to provide a systematic analysis that adds credibility for the prioritisation. As for the factor on criticality, there are concerns that it may be difficult or even less credible to assess the severity and probability of a safety concern that has epistemic uncertainty. This is especially so if the effect of a safety concern is unknown or non-deterministic during the time of assessment because of epistemic uncertainty.

Action Plan

The participants considered the guided questions to derive the track and address action plan useful. They feel that the questions can serve as a structured reference during the process of developing the action plan. However, they raised a similar concern of the need of expertise to formulate quality action plan, just like the need of expertise to make sense and contextualise the HOT-PIE taxonomy.

As a form of continuous improvement, the participants suggested that the questions can be improved by generalising the queries that are commonly asked by committee members in safety forums, as well as insights from safety assessments conducted for past projects.

<u>Top-level Theme 3: Moving Forward</u>

Pilot Implementation

Some of the participants feel that, where there are benefits, project managers may not be receptive to devote the time and efforts to conduct T.A.G. approach. This is especially so for those in the front-end design and planning, as they may not be able to appreciate how the final system may operate with other systems. The participants foresee a reluctance to embark fully on the T.A.G. approach when its benefits are not clearly demonstrated. Hence a pilot implementation may be a suitable approach moving forward. It will help to showcase a concrete example of an actual project that uses the T.A.G. approach. However, the participants acknowledge the challenge of such pilot implementation as there will be concerns such as:

- The risk of delaying the project schedule
- The security impact of sharing information about the project
- The factors to consider and areas to observe during the trial so as not to be too intrusive to the ongoing project
- The duration of the pilot implementation to obtain credible result

Targeted Observations

Using the three steps of the T.A.G. approach, many of the participants find that it can provide valuable insights into causal conditions to complement existing safety assessment techniques during hazard analysis. The T.A.G. provides a different perspective that can help to make safety assessment more credible, especially by advocating the use of HOT-PIE taxonomy to check for blind spots when identifying hazards.

One major blind spot that was brought up regularly is the safety concerns related to human. Currently, there is still this assumption in safety assessment that human will do what is intended. Especially for engineers, there may be this perception that they can't control human errors that occur when human deviates from expected operation procedure. However, the participants see value of using the T.A.G approach to discover uncertainties that may eventually influence the redesign of a system so as to mitigate human related concerns and hazards.

The participants feedback that the T.A.G. approach has the potential to make the safety assessment more comprehensive as it:

- Provides a systematic approach based on established theories and concepts
- Provides a structured process that is repeatable and traceable
- Synergises with existing safety techniques by integrating into existing process

7.5. Preliminary Industrial Adaptation

Besides conducting the semi-structured interview, discussions between the Singapore Air Force and the defence agency have also been initiated to explore the possibility of adapting the T.A.G. approach into the current way of conducting safety assessment. Through preliminary discussions with the relevant stakeholders that are responsible for the system safety methodology and safety governance in the air force, we have identified possible areas that the existing T.A.G. approach can be adapted to the existing safety assessment for complicated safety-critical systems, which is known as networked system-of-systems (NwSoS) in the Singapore context. As these are long-term ongoing efforts, we can expect more fine-tuning as we receive more feedback from the users. In this section, we will describe four key adaptations of the T.A.G. approach that have already been initiated (see Table 47). The adaptations aim to make the T.A.G. approach more feasible and systematic when integrated with existing safety assessment techniques.

Table 47. Summary of Adaptation of the T.A.G. Approach

			Enhancement to	
Section	Baseline	Adaptation	Feasibility	Systematic
1	Applying in any Hazard Analysis	Applying specifically in Scenario-Based Hazard Analysis	Yes	-
2	HOT-PIE taxonomy	Context-specific HOT-PIE taxonomy	Yes	Yes
3	Original prioritisation factors	Review prioritisation factors to better represent domain	Yes	-
4	Free text to develop action plan	Action-item table to track action plan	Yes	Yes

7.5.1. Applying T.A.G. approach in Multi-Perspective Hazard Analysis

Multi-Perspective Hazard Analysis (MPHA) is the current approach to analyse complicated NwSoS in the Singapore Air Force. MPHA consists of four types of hazard analysis, with each type catering for a specific context. The details of the MPHA are described in the paper by Chan et al that was presented at the 2015 International System Safety Conference. The following table summarises the differences between the four types of hazard analysis (see Table 48).

Table 48. Multi-Perspectives Hazard Analysis

No.	Types of HA	Purpose (Extracted from Chan et al)
1	Top Level Mishap Scenarios (TLMS)	The TLMS technique was developed to analyse NwSoS in its entirety for top level mishaps and their potential causes. The TLMS findings also serve as a hazard prompt list for subsequent techniques. A holistic top-level perspective is adopted to study the overarching concept of the NwSoS operations and capabilities, which spans across multiple constituent systems.
2	Scenario Based Hazard Analysis (SBHA)	The scenario perspective examines the safety concerns related to SoS operations and interoperability between the constituent systems. Subject matter experts and operators need to draw upon their experience and expert knowledge to raise pertinent safety concerns associated with specific operational scenarios. These scenarios typically involve multiple constituent systems, and therefore may require various experts to communicate their knowledge of the constituent systems.
3	Message Based Hazard Analysis (MBHA)	In the NwSoS, the main interaction is through electronic messages communicated between the constituent systems. The MBHA was developed to analyse the message interface between the constituent systems and how message faults can impact the recipient constituent systems, and subsequently the SoS.
4	Constituent System Hazard Analysis (CSHA)	Constituent systems come inherently with hazards and some of these hazards may be triggered by the interoperability nature of the SoS. The

No.	Types of HA	Purpose (Extracted from Chan et al)
		CSHA was developed to analyse each constituent system for potential emergent hazards and causal factors when interacting with the SoS.

After deliberations with the technique experts, it is assessed that the most relevant hazard analysis to integrate the T.A.G. approach is the SBHA. This is because SBHA depends largely on the experiences of the safety analysts during the safety assessment and the quality of the assessment depends on the information available. This is where the T.A.G. approach can provide the safety analysts with the tool to manage the epistemic uncertainties that can affect the SBHA.

By choosing the SBHA as the target of implementation, it makes the implementation of the T.A.G. approach more feasible and acceptable by the stakeholders. Moving forward, the T.A.G. approach would be incorporated into one of the constituent systems (i.e. subsystems within the NwSoS) safety assessment and monitored for a few milestones across the project lifecycle for at least 6 months to a year.

7.5.2. Context-Specific HOT-PIE Taxonomy

Raising the same concern from the semi-structure interview, both the technique and domain experts are concerned that the full list of HOT-PIE taxonomy is not practical for implementation. A session was conducted with the relevant stakeholders to trim the taxonomy to an initial list of context-specific HOT-PIE causal mechanisms (see Table 49). The list shall be used during a SBHA for a planned engineering testing for one portion of the NwSoS. The context-specific HOT-PIE taxonomy will be reviewed after three months so that the team will not be fixated only by the causal mechanisms that have been selected in the first session.

Table 49. Context-Specific HOT-PIE under Scenario Based HA for Engineering Testing

Primary	Secondary Factor	Probing questions
Human	Communication and Coordination	How was the test made known to and communicated to the users? How are the agencies responsible to carry out the mitigation measures being informed?
	Expertise	Are the personnel carrying out the testing identified and tagged individually?
	Ownership	Does the test director know his/her term of references?
Organisation	Structure	What is the composition of the testing team? Is the schedule for the testing realistic?
	Goals and Expectations	Is the intent of the testing made known explicitly?
	Audit and Control	Is there any independent party involved in the testing? What are the control mechanisms to ensure proper documentation of the testing?

Primary	Secondary Factor	Probing questions
Technology	Engineering Feedback	How are technical faults affecting the testing feedback timely?
Process	Systematic	How are outstanding issues being tracked and resolved? How are changes in the testing plan and procedures being managed?
	Level of Autonomy	How is the conduct of the testing being administered? Who can make changes to the test plan?
	Risk Management	How are users being kept aware of the risk during testing?
	Configuration Control	How are hardware and software versions being monitored during the duration of the testing?
Information	Documentation	How are the test results documented?
	Standards and Policy	What are the policies and orders governing the testing? How are incidents supposed to be reported?
	Evidence Management	How is evidence being documented and transferred for future usage under the lifecycle?
Environment	Regulatory Control	Is there any configuration control manager? How would any deviation during testing be managed?
	Security	How is the testing protected from virus?
	Operating Condition	Are there any pre-requisite conditions before conducting the testing?

By developing a context-specific HOT-PIE taxonomy, it makes the T.A.G. more feasible for the stakeholders to apply during hazard analysis. This process of coming out with probing questions based on the HOT-PIE categories introduces a systematic way of identifying epistemic uncertainties during the safety assessment. The plan is to solicit feedback on the usefulness of the context-specific HOT-PIE taxonomy after the first three months and to decide if further refinement is needed.

The evidence that we can make use of the HOT-PIE taxonomy and optimise it by choosing a subset that is practically reasonable for the stakeholders, reinforces our goal to develop an approach that is feasible and systematic. It is feasible since the taxonomy is optimised for practical usage. It is systematic since we can trace the final context-specific list to the original, more extensive HOT-PIE taxonomy that has been developed from references in Annex A.

7.5.3. Review of Prioritisation Factors

The third adaption is with regards to the prioritisation factors. After a discussion with the domain experts for the NwSoS, the following changes have been made to the prioritisation factors to better represent the local context. The changes are highlighted in Figure 87.

1. Instead of using the word 'Criticality', a more common term in the local context is 'Mishap Potential'. The concern is that 'criticality' has a different definition in software analysis as it denotes different levels of software criticality. In order not to be confused, 'Mishap Potential' is considered a more precise term to mean the probability and severity of a mishap as part of evaluating the risk due to the uncertainty.

2. 'Feasibility' replaces 'resource availability' as it was felt that 'feasibility is a word that is more commonly used in industry. According to the domain experts, the word 'feasibility' can better represent the assessment of how much resources that are available to clarify the uncertainty, as compared to resource availability.

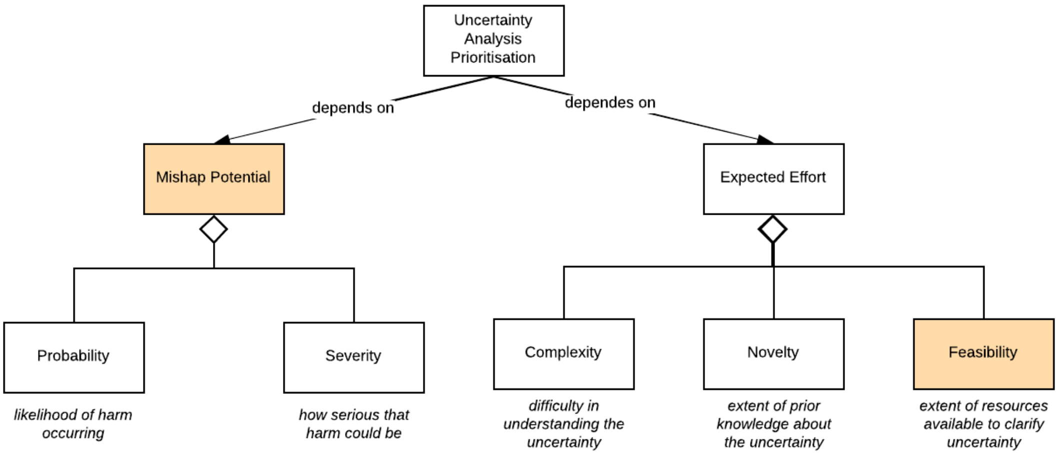

Uncertainty Analysis Prioritisation
depends on
dependes on
Mishap Potential
Expected Effort
Probability
likelihood of harm occurring
Severity
how serious that harm could be
Complexity
difficulty in understanding the uncertainty
Novelty
extent of prior knowledge about the uncertainty
Feasibility
extent of resources available to clarify uncertainty

Figure 87. Refined Prioritisation Factors for NwSoS

By changing these two terms, it is believed that stakeholders will be more receptive of using the factors to prioritise the uncertainty analyses after identifying the uncertainties during the safety assessment. The prioritisation factors can be used after identifying the uncertainties in the ongoing safety assessments of the constituent system acquisition lifecycle.

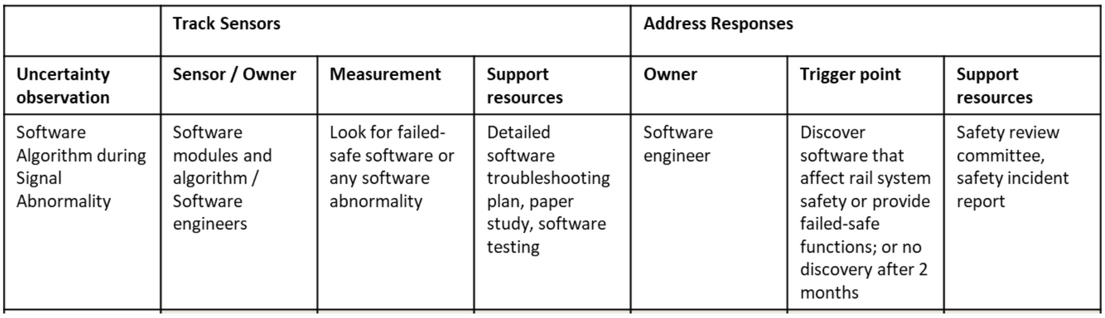
However, while it is important to customise the factors for the local context, we need to be mindful that 'criticality' and 'resource availability' have different meanings compared to 'mishap potential' and 'feasibility'. We have explicitly documented the conscious decision to adopt the later, while keeping in mind the original two words as reference. This is important as the words 'mishap potential' and 'feasibility' may not be suitable in other contexts.

A post implementation review will be initiated after one year of using the prioritisation factors.

7.5.4. Action-Item Table for Track Action Plan

The original intent is to make use of the guided questions to set goals in the form of track and address reports. Such textual report, while comprehensive, may not be easily traceable or tracked throughout the system lifecycle. During the discussion with domain and technique experts, it was suggested that key issues under the track and address action plan can be monitored in the form of a table. This can mirror the way that action items are being tracked in current way of doing business in the air force. By using a familiar format, it makes the T.A.G. approach easier to be accepted by the safety analysts. Having specific categories under the tracking table that follows the format of the guided questions (see Table 50 for an example) would also make the process more systematic.

Table 50. Example of Track and Address Table



	Track Sensors			Address Responses		
Uncertainty observation	Sensor / Owner	Measurement	Support resources	Owner	Trigger point	Support resources
Software Algorithm during Signal Abnormality	Software modules and algorithm / Software engineers	Look for failed-safe software or any software abnormality	Detailed software troubleshooting plan, paper study, software testing	Software engineer	Discover software that affect rail system safety or provide failed-safe functions; or no discovery after 2 months	Safety review committee, safety incident report

Like the prioritisation factors, a post implementation review on the action-item table will be conducted after a year of usage.

7.6. Conclusions

This chapter aims to evaluate the thesis proposition by (1) verifying if the T.A.G. approach does indeed help to identify, document, track and address epistemic uncertainties, as well as (2) validating if the T.A.G. approach can be feasibly and systematically applied in industry.

Verification. Through the evaluations against the initial requirement (section 7.2) and sharing at conferences / workshops (section 7.3), we have demonstrated it is possible to use the T.A.G. approach to identify, document, track and address epistemic uncertainties. However, we have also observed limitations of the T.A.G. approach during the verification, especially the following:

- The HOTPIE taxonomy is useful to help safety analysts to identify causal condition that have yet been considered. However, its usefulness is limited by what has been included in the taxonomy. While we have tried to be as diverse as possible, we can never be so comprehensive that the taxonomy covers all domains. We must be prepared to acknowledge this limitation and continue to keep updating the list on a regular basis. Moving forward, we will continue to conduct periodic review on the list of causal conditions identified in the HOTPIE taxonomy.

- While making the HOTPIE taxonomy comprehensive, the list may grow to a level that is not practical to use. Theoretically, we suggest that domain experts may look through the original list and discard those that may not be relevant. This may not be a simple process as we have assumed the domain experts are provided with some guidelines or experiences to carry out this filtering process. Moving forward, we are exploring if the full set of taxonomy can be categorised into the groups: compulsory and optional. The compulsory group should be included in all T.A.G. approach, while the optional ones may be discarded depending on the domain of interest.

- We have suggested factors to consider when prioritising the uncertainties to focus on and introduced a scoring that is based on individual subjective assessment. Like other safety assessment techniques, the limitation comes with this subjective assessment, which depends on who are the safety analysts present during the assessment and the experience level of these safety analysts. There is a danger that the assessment may not be consistent or reproducible since it is subjected to whoever is present. This limitation also applies to the guided questions that are introduced to help in tracking and addressing the epistemic uncertainties. The usefulness of these questions depends to a great extent on the knowledge of the safety analysts using the guide. Moving forward, we will continue to evaluate the effectiveness of the prioritisation factors and well as the approach of quantifying the various factors in a more consistent manner.

- We have provided examples in chapter 5 and 6 to show that it is possible to use the T.A.G. approach to complement existing safety assessment techniques. We have intentionally chosen STPA, FTA and FMEA techniques as these are the common techniques in safety assessment. However, we acknowledge that there are other techniques where further customisation may be needed before our T.A.G. approach can be integrated to manage epistemic uncertainties. Moving forward, we are looking at developing similar processes for other techniques such as Hazard and Operability Study (HAZOP) and Common Cause Analysis.

Validation. While we verify that it is possible to use the T.A.G. approach to complement existing safety assessment technique, it is more difficult to validate the approach in the industry. Through the semi-structured interview (section 7.4) and preliminary industrial application (section 7.5), we have solicited feedback to indirectly show that it is feasible to implement the T.A.G. approach. However, there are limitations in our validation approach and the key challenges are as follows:

- During our validation, we have chosen to focus on the domain of military aviation and defence systems management for the RSAF since it is the key domain of interest for this research. However, being such a niche domain, the participants for the interview would have to be targeted. Since the number of participants is small, any quantitative feedback would not be statistically significant, and it is used more as a reference rather than a conclusive observation. Moving forward, we intend to continue the validation beyond the RSAF, to explore either to solicit feedback from our Army and Navy counterparts; or to even share the approach with other air forces in the world.

- Ideally, to validate the T.A.G. approach, it should be integrated into all the safety assessments throughout the system lifecycle from early design concept phase till the deployment of the system. Unfortunately, this would take years and even decades for some complicated systems; which is not practical as part of our research to observe for such an extensive period. Specifically, the pilot implementation for the industry, as proposed in section 7.5, would require multiple reviews throughout the years to assess their effectiveness in better managing epistemic uncertainties. Realistically, we can only solicit anecdotal observations and feedback as much as possible, either through a one-time application of the approach in the pilot implementation or request the participants to imagine themselves using the T.A.G. approach for their safety assessments. Moving forward, we have already derived a systematic way to assess the usefulness of the pilot implementation whenever it is being used by safety analysts. Such data will be consolidated into the existing database of feedback that we have created for further trending and analysis.

Chapter 8 – Conclusions

8.1. Summary of Thesis Contributions

This thesis has developed and demonstrated an approach to managing epistemic uncertainties that are associated with the underlying models of safety assessment for safety-critical systems. This approach is based on three inter-dependent initiatives that aim to identify, document, track and address known and unknown uncertainties. Specifically, the research has contributed the following:

- Defined a taxonomy to classify causal mechanisms so as to identify epistemic uncertainties and demonstrated an approach that uses the taxonomy to identify uncertainties not covered in the original safety assessment (see section 8.1.1)
- Defined a process to document and prioritise epistemic uncertainties and demonstrated an approach that is based on this process to analyse and prioritise uncertainties that have been identified in existing safety assessments (see section 8.1.2)
- Defined guided questions to develop action plans and demonstrated an approach to develop actionable goals to track epistemic uncertainties throughout the system lifecycle and address them when some thresholds are met (see section 8.1.3)

8.1.1. Contributions to Identifying Epistemic Uncertainties

In chapter 3, we have highlighted that while epistemic uncertainties are inherent in the underlying models used in safety assessment, there is no commonly practised approach to identifying them. This is reinforced in the literature survey on uncertainty management in chapter 2. Our survey shows that uncertainty in safety management is widely acknowledged and yet there is no systematic and common industrial practice to identify such uncertainties in safety.

To set the foundation of identifying epistemic uncertainties, we have developed the *Conceptual Model of System Safety* to emphasise that uncertainties can reside in both the safety and system models, as well as the *Causal Relationships Model of Safety Assessment* to help locate the uncertainties that reside in the cause-and-effect relationship during hazard identification. Through a review of literature related to safety, we have established the HOT-PIE taxonomy of causal mechanism that comprises six primary causal conditions and fifteen secondary causal conditions. The taxonomy serves as a guide to help the user to recognise epistemic uncertainties not covered in the original assessment. Since the HOT-PIE taxonomy is based on literature research, safety analysts can harness the collective wisdom from other experts, rather than be limited by the experience and expertise of individuals in a safety assessment.

To show the application of HOT-PIE taxonomy, we have used it on safety causal models with both system and component viewpoints. In chapter 5, we applied the HOT-PIE taxonomy on the Yongwen railway accident on models from system viewpoints to discover four plausible-but-uncertain causal relationships. In chapter 6, we demonstrate the usefulness of HOT-PIE taxonomy on the ARP 4761 example by identifying four plausible-but-uncertain causal relationships across different phases of the safety assessment process and over a spectrum of primary causal conditions. These uncertainties are linkages among nodes that may or may not have prior relationships with each other.

8.1.2. Contributions to Document Epistemic Uncertainties

In chapter 3, we have highlighted that structured documentation of epistemic uncertainties reduces bias during safety assessment and reduces the chance of ignoring such uncertainties. And yet, in the literature survey, while we came across literature that attempted to define uncertainties beyond just safety, such as the Cynefin model, there is no research that focus on integrating uncertainty management into existing safety techniques to be applied in the industry.

We have introduced the T.A.G. approach in chapter 4 and deliberately integrated the approach into existing safety assessment techniques to show that it is practical to apply the approach in the industry. We further demonstrate that the T.A.G. approach can be integrated into the STPA technique to analyse the Yongwen railway accident and the FMEA and FTA techniques used to analyse the ARP 4761 example. By complementing these three existing safety assessment techniques, we have been able to systematically document the epistemic uncertainties identified in previous section. Subsequently, we prioritise these uncertainties for analysis based on criticality and expected effort, the two factors we defined in chapter 4. Participants of the semi-structured interview have fed back that the factors are indeed more useful and systematic to help in prioritising the uncertainties, compared to doing the prioritisation in an unstructured manner.

8.1.3. Contributions to Track and Address Epistemic Uncertainties

System safety engineering demands a systematic and through-life management of knowledge and this is especially applicable to uncertainty management since uncertainty can evolve with time. In our review of existing safety management processes and safety standards, there are established processes to track hazards across the system life-cycles. However, in most safety assessments, hazards are derived using known causal relationships that can be traced from root causes to failure effects.

Our proposal to manage hazards through-life is novel as we track both the hazards and the influence of epistemic uncertainties on the hazards through-life. This approach allows a timely response when information related to the epistemic uncertainties becomes available. To advocate a continuous through-life management of uncertainty, we have developed guided questions based on four factors for safety analysts to consider when they develop their action plans to monitor uncertainties: deciding on monitoring technique, forming the monitoring activity, setting trigger points and adapting to change. To demonstrate its feasibility, we have constructed the action plans in chapter 5 and 6 for the Yongwen railway accident and the ARP 4761 example respectively to manage epistemic uncertainties through-life. During the semi-structured interview, participants have fed back that the guided questions help them in formulating the action plan to track and address epistemic uncertainties through-life.

8.1.4. Overall Conclusions

In conclusion, we have established an approach that pushes the knowledge frontier during safety assessment for safety-critical systems in two specific ways. Firstly, we have created a HOT-PIE taxonomy that provides the opportunity for safety analysts to transform unknown uncertainties into known uncertainties. The risk of such known uncertainties can then be assessed as part of ongoing safety assessment. Secondly, we have introduced the T.A.G. approach to track and address known uncertainties systematically to reduce the chance of known uncertainties that may be safety-critical from being ignored, as a system is being developed throughout its lifecycle.

While we have shown that it is possible to complement existing safety assessment technique using our T.A.G. approach, there are scalability and validity concerns with applying the approach. Feedback from peer reviews and interviews raise the issue of scaling the list of causal factors in the HOTPIE taxonomy and the list of uncertainties being tracked. Both lists may grow to a level that may not be practical to manage realistically as we scale up the usage of the approach in large scale system. In terms of validity, our approach can only be evaluated comprehensively if it is incorporated into safety assessments right from system conceptual design till deployment. Realistically, this is not possible within the duration of our research due to industrial limitations and the protracted system acquisition lifecycle that may span many years.

8.2. Future Work

From the development, application and feedback on the research, the following areas have been identified for future work:

- Complementing quantitative techniques
- Complement with IT and Data Analytic Tools
- Integrating with other Safety Assurance Initiatives
- Integrating with other Safety Assessment Initiatives
- Refining through Pilot Implementation
- Applying beyond Safety

The following sub-sections provide brief descriptions of these areas.

8.2.1. Complementing Quantitative Techniques

In chapter 2, we have discussed the importance of appreciating the context from which a quantitative uncertainty analysis take reference. For example, the “model of the world” in Mensing’s approach [57] and the “background knowledge” defined by Aven in his quantitative assessment of uncertainty [61]. As this context changes, the validity of the quantitative assessment may be affected.

The T.A.G. approach can be used to complement such quantitative techniques by explicitly identifying epistemic uncertainties in these contexts using the HOT-PIE taxonomy as a prompter. Subsequently, any epistemic uncertainties associated with these contexts can be tracked through-life and reviewed wherever it is suspected that the validity of the quantitative assessment may be affected.

When prioritising the uncertainties to focus on, as highlighted in section 4.3.2.3, applying an average of the overall priority score (based on the factors of criticality and expected effort) from each safety analyst may produce skew and misleading results. One way to improve the T.A.G. approach could be to conduct qualitative discussion and sharing among the safety analysts to build consensus in selecting the more important uncertainties to focus on first.

8.2.2. Complementing with Database

To make the T.A.G. approach more pervasive, future research can consider developing a scalable database to allow epistemic uncertainties to be easily documented, stored and tracked during a safety assessment. The database should reference the HOT-PIE taxonomy to prompt users of potential epistemic uncertainties. Once documented, the epistemic uncertainties can be tracked and mobile application can be implemented to notify the user when some triggering points are met to address the uncertainties.

Another possible area of focus is to ride on the tools that are developed for data analytic to help in decision making. Data analytics uses software tools to analyse data so as to derive conclusions about the information in the data. There is a push for data analytics in the recent years as software tools are becoming more accessible and pervasive. If the T.A.G. approach can collect sufficient data about the epistemic uncertainties associated with the safety assessments across many safety-critical systems, this can potentially be a rich source of information for data analytics.

8.2.3. Integrating with other Safety Assurance Initiatives

We have mentioned in chapter 2 that the ways to manage safety assurance include either arguing that an uncertainty is not safety critical or developing an approach to reduce the uncertainty to a level that is acceptable. T.A.G. approach can be used when epistemic uncertainties that affect safety assurance need to be tracked through-life after a safety assessment. This is important as the uncertainties evolve, the level of safety assurance may change depending on how much unknown regarding the safety assessment has become available. For example, in the "assured safety argument" proposed by Kelly and Hawkin [72], the assurance case can reference the epistemic uncertainties discovered by a T.A.G. approach as part of its arguments and evidence.

8.2.4. Integrating with other Safety Assessment Initiatives

The advantage of the T.A.G. approach is that it provides a generic process flow (see section 4.3) that can be easily adopted on other safety assessment initiatives. Specifically, in the defence industry, there are many opportunities to use the T.A.G. approach to complement with other ongoing initiatives to enhance safety assessment for complicated systems. For example, the relevant authorities can consider integrating the T.A.G. approach in the following ongoing enhancement to safety assessment:

- ASEMS Proportionality Assessment (UK MoD). The intent of the study in the UK MoD is to ensure that the extent of managing and assuring the safety of a system is proportionate to the risks involved. Using the same argument, the extent of tracking an epistemic uncertainty should be proportionate to the assessed risks. Moving forward, a study can be conducted to synergise the efforts between uncertainties prioritisation under the T.A.G. approach and proportionality assessment introduced in ASEMS.

- Change Impact Analysis (Singapore Air Force). Another area where the T.A.G. approach can be applied is on the Change Impact Analysis (CIA) that is developed to complement existing Multi-Perspective Hazard Analysis (see section 7.5). The CIA is used to systematically characterise changes in complicated safety-critical systems so as to assess the extent that a change may proliferate to other subsystems. This would determine to what extent the four types of hazard analysis need to be conducted due to the change. The T.A.G. approach can be used to monitor epistemic uncertainties that concern the CIA, as such uncertainties can affect the extent of conducting the four types of hazard analysis.

8.2.5. Refining through Pilot Implementation

As mentioned in section 7.4, a pilot implementation of the T.A.G. approach may be preferred to demonstrate its benefit without causing too much disruption to the established way of conducting safety assessment. With the observations from chapter 7, the pilot implementation can be more targeted and further reviews can focus on concerns that have already been surfaced in chapter 7 (such as evaluating the context-specific HOT-PIE taxonomy and refining the prioritisation factors). Besides selecting the scope of the pilot study, proper communication with the stakeholders on the intent and content of the T.A.G. approach is necessary before applying the T.A.G. approach. Periodic progress reviews are also important so as to introduce timely refinement.

8.2.6. Applying beyond Safety

One possible area to apply the T.A.G. approach beyond safety is for security assessment. This has also been surfaced before during the sharing of T.A.G. approach in conferences and workshops. Although security assessment is also a type of risk analysis, the potential causes of security concerns can be different from safety. Hence, the HOT-PIE taxonomy needs to be reviewed to encompass security concerns. For example, further research is needed to integrate key security considerations (e.g. vulnerability, denial of service, social engineering) into the six primary causal conditions of the HOT-PIE taxonomy.

8.3. **CODA**

As technology advances and systems get more complicated, safety assessment techniques have to evolve so as to model multiple viewpoints in order to be comprehensive. For example, the STPA hazard analysis technique is a method to include multiple viewpoints (e.g. component, operation, organisation and government) in a safety assessment model. As such a complicated model is being developed within the limitation of existing resources (such as time and manpower), it can exacerbate the presence of epistemic uncertainties during safety assessment.

For example, while the STPA provides the means to consider organisational viewpoints, the safety analysts may not have the experience or necessary information to assess the safety concerns at the organisational level (e.g. concerns relating to organisation policies and standards). As a result, this may create more epistemic uncertainty during the safety assessment as the analysts attempt to make decisions based on the assessed risk taking into consideration organisational issues.

While others focus on developing safety causal models to better represent the real world, we took a more pragmatic approach by targeting the possible presence of uncertainties (i.e. both unknown unknowns and known unknowns) associated with the underlying models in safety assessment that can become safety-critical in the lifetime of the system.

Managing epistemic uncertainties in the underlying models of safety assessment will require time and resources to identify, document, track and address these uncertainties. Although effort to manage uncertainties into the future may seem not to immediately affect the risk assessment, it affects the confidence in the assessed risk. There is the unspoken danger that such uncertainties can become one of the contributing safety-critical hazards down the system lifecycle, such as the fate of the CHALLENGER space shuttle as illustrated in chapter 1. If one cannot confidently accept an uncertainty as not safety-critical, a reasonable attempt to capture and manage such uncertainty can be beneficial in keeping the system safer.

This has motivated us to review the current approach of treating epistemic uncertainties and embark on an endeavour to reduce the undesired effects of epistemic uncertainties associated with safety assessment. The T.A.G approach is a plausible product to complement existing safety assessment techniques to provide a pragmatic balance between the desire to eliminate uncertainties and the practical limitation of managing such uncertainties.

Annex A – Literature Referenced in HOT-PIE Taxonomy

No.	Title	Author	Description
1	Failure in safety-critical systems [154]	C. Johnson	Barrier analysis: assumes that hazard comes into contact with target because barriers or controls were unused or inadequate. The method analyses people, process and technology measures taken to prevent target from affected by hazard. Change analysis: Determine if abnormal work practices contribute to the cause of adverse occurrence. Deviation from normal operation may be intentional but it is difficult to predict the impact of the change.
2	An Accident Causation Analysis and Taxonomy (ACAT) model of complex industrial system from both system safety and control theory perspectives [155]	W. Li et al.	The ACAT model uses a combination of system factors and control functions to form a matrix model for accident causation analysis and classification. In the model, complex systems are decomposed into six components, namely machine, man, management, information, resources and environment. From control theory perspective, actuator, sensor, controller and communication are defined as part of the control function.
3	IS/IT project failures: a review of the extant literature for deriving a taxonomy of failure factors [156]	Y. K. Dwivedi et al.	The research collates and classifies existing literature review to understand common failure factors in Information System / Information Technology projects. It provides a list of factors and attempts to categorise them by geographical location.
4	FRAM, the functional resonance analysis method: modelling complex socio-technical systems [157]	E. Hollnagel	The Functional Resonance Analysis Model (FRAM) defines a systemic framework to model complex systems for accident analysis purposes. FRAM uses a non-linear accident model based on the assumption that accidents result from unexpected combinations (resonance) of normal performance variability. It identifies essential system functions by characterizing them using six basic parameters (Input, Output, Time, Control, Precondition, and Resource)
5	The human factors analysis and classification system—HFACS [146]	S. A. Shappell et al.	The HFACS framework is based on Reason's 'Swiss Chess' model and has been developed to provide a general human error framework around which investigative methods can be designed and accident database can be restructured. The framework has been used within the military, commercial,

No.	Title	Author	Description
			and general aviation sectors to systematically examine underlying human causal factors and to improve aviation accident investigations. It comprises four levels of failure: 1) Unsafe Acts, 2) Preconditions for Unsafe Acts, 3) Unsafe Supervision, and 4) Organisational Influences.
6	Handbook of human systems integration [158]	H. R. Booher	Human Systems Integration (HSI) is an integrating discipline designed to help move business and engineering cultures toward a more people-technology orientation. A wide range of tools, techniques, and technologies have been developed to integrate human factors into engineering systems. For this to be effective, HSI outlines the principles and methods that can be used to help integrate people, technology, and organisations with a common objective toward designing, developing, and operating systems effectively and efficiently.
7	Patient safety: latent risk factors [159]	M. V. Beuzekom et al.	Factors that promote errors may not be directly visible in the working environment. The underlying latent causes are categorised into a limited number of classes known as the Latent Risk Factors (LRFs). Understanding how LRFs affect safety can help to design more effective control measures
8	Risk management in a dynamic society: a modelling problem [160]	J. Rasmussen	The research argues that risk management must be modelled by cross-disciplinary studies, considering risk management to be a control problem and serving to represent the control structure involving all levels of society for each particular hazard category. The socio-technical system includes several levels ranging from legislators, over managers and work planners, to system operators. This system is stressed by a fast pace of technological change, by an increasingly aggressive, competitive environment, and by changing regulatory practices and public pressure.
9	Causal mechanisms in the social sciences [142]	Hedstrom & Ylikoski	Provides an account of the important philosophical and social science concepts to the mechanism approach. It focuses on how the idea of causal mechanism has been applied in the social sciences such as political science and criminology. The literature covers factors from the human, organisation and environment perspectives.
10	Addressing enablers in layers of protection analysis [161]	P. Baybutt	Enablers are events and conditions that do not directly cause the hazard scenarios but are required to be present or active for the scenario to proceed. Such enablers can be classified

No.	Title	Author	Description
			broadly into human actions, equipment failures and external events.
11	Near miss reporting in the chemical process industry [162]	Schaaf	Eindhoven Classification Model: Use to classify causal factors into three main categories of failures (technical, organisational and human) for detailed analysis
12	A taxonomy of network challenges [163]	Cetinkaya & Sterbenz	Taxonomy of network challenges focus on adverse events triggering faults that eventually result in service failures. Understanding these challenges accordingly is essential for the improvement of the current networks and for designing future internet architectures. The classification of network challenges includes natural, human-made, dependency and technology factors.
13	Cognitive reliability and error analysis method (CREAM) [164]	E. Hollnagel	CREAM consists of a method, a classification scheme and a model to analyse human reliability and erroneous actions. The underlying model makes a distinction between causes, known as genotypes, and effects, known as phenotypes. There are 3 categories of causes, namely human psychology, technology and organisation. The effects would be different error modes.
14	The balance theory and the work system model [165]	P. Carayon	The work system model in the balance theory provides a way of describing all the elements of work that affect workers and outcomes. According to the model, tasks are performed by a person who uses tools and technologies; the tasks are performed in a physical environment and under organizational conditions. Hence, the five elements of work are individual, tasks, tools and technologies, physical environment and organisational conditions.
15	Job strain, effort-reward imbalance and employee well-being: a large-scale cross-sectional study [166]	Jonge et al	This study investigated the effects of the Job Demand-Control (JD-C) Model and the Effort-Reward Imbalance (ERI) Model on employee well-being. Part of a larger scientific program that aims at understanding the contribution of social and psychological factors to human health and disease. It primarily focuses on the human and environment conditions.
16	Man and machine - Systems for safety [167]	E. Edwards	In the S.H.E.L.L. model, each person (centre Liveware) is applied to and interacted with the other four components (Hardware, Software, Environment and Liveware). It is believed that a mismatch between the centre Liveware and

No.	Title	Author	Description
			any other four components will potentially lead to a cause of human error.
17	MIS problems and failures: a socio-technical perspective, part II: the application of socio-technical theory [168]	Bostrom & Heinen	The research introduces socio-technical interacting classes to counter the issue of not having a systematic view to analyse socio-technical issues. These classes include technology and tasks in the technical system and the structure and people in the social system.
18	Engineering a safer world: Systems thinking applied to safety [66]	N. G. Leveson	The STAMP model is an accident causality model that is based on system and control theories. The model is constructed from three basic concepts: constraints, hierarchical levels of control, and process models. The method focuses on issues related to technical, human and organisation factors in complex socio-technical system.
19	Fundamentals of ergonomics in theory and practice [169]	J. R. Wilson	Ergonomic refers to the theoretical and fundamental understanding of human behaviour and performance in purposeful interacting socio-technical systems, and the application of that understanding to design of interactions in the context of real settings. It considers factors that interact with an individual, including organisation, logistics, environment, task, technology interface, context, temporal and cooperation.
20	Work system theory: overview of core concepts, extensions, and challenges for the future [170]	S. Alter	The theory provides a static view of a current (or proposed) system in operation and a dynamic view of how a system evolves over time through planned change and unplanned adaptations. The work system framework represents a work system in terms of nine elements: processes and activities, participants, information, technologies, products/services, customers, environment, infrastructure and strategies.
21	Sociotechnical attributes of safe and unsafe work systems [171]	Kleiner et al	Theoretical and practical approaches to safety based on socio-technical system principles focus on the intersection between social, organisational, technical and work process factors. The literature discusses three approaches to analyse and design complex STS, namely human-systems integration, macro-ergonomics and safety climate.

No.	Title	Author	Description
22	A taxonomy of situation awareness errors [80]	M. R. Endsley	The research is conducted on aircrew situation awareness. For aircrew, while some of these incidents may represent failures in actual decision making (action selection), a high percentage of these errors are actually errors in situation awareness. These errors are categorised according to the ability to perceive, comprehend and project situation. The taxonomy focuses on both information and human factors that affect the situational awareness of an individual.
23	The job demands-resources model: State of the art [172]	Bakker & Demerouti	The model specifies risk factors associated with job stress. These factors can be classified in two general categories (i.e. job demands and job resources) and can be applied to various occupational settings. It primarily focuses on human factors.
24	The centrality of work [173]	Dejours & Deranty	The "centrality of work" is based on the "psychodynamic" approach to work. It distinguishes between four separate but related ways in which work can be said to be central: psychologically, in terms of gender relations, social-politically and epistemically. It considers factors like work engagement, psychic defence and worker behaviour.
25	Modelling, metamodeling, and taxonomy of system failures [174]	J. P. van Gigch	Taxonomy of failures is important so that system analysts, engineers, designers, and managers may one day agree upon a standard system of labelling, coding, counting, and measuring failures. Classification and measurement are prelude to the discovery of patterns and leads to understanding, prediction, and avoidance. The types of failure in the taxonomy include structure, technology, regulation, rationality, behaviour and evolution.
26	Reason's Human error theory [175]	J. Reason	In the "Swiss cheese" model of human error, Reason describes four levels of human failure, each influencing the next: Organisational influences, unsafe supervision, precondition for unsafe acts and unsafe acts. For the human errors, the focus is on slip, lapse and mistake potentially committed by an individual.
27	Promoting safety in the oil industry [176]	Wagenaar et al	The TRIPOD framework is used to analyse disturbances to safe operation based on an underlying model of causation. Assume incident is caused by local triggering factors known as general failure types: hardware (HW): design, maintenance management, operating procedures, error-enforcing

No.	Title	Author	Description
			conditions, housekeeping, incompatible goals, communication, organisation, training and defence planning.
28	Macroergonomics: analysis and design of work systems [177]	B. M. Kleiner	Macroergonomics is the design of work systems which focuses on organization-system interaction. A work system comprises two or more people working together (i.e. personnel sub-system), interacting with technology (i.e. technological sub-system) within an organizational system that is characterized by an internal environment (both physical and cultural) and external environment.
29	Towards standardisation of no fault found taxonomy [178]	Khan et al	The research attempts to standardise the taxonomy surrounding the phenomena commonly known as No Fault Found. It classifies NFF conditions under organisational, built-in tests, integration and technical intermittent.
30	Management Oversight and Risk Tree-MORT [179]	W.G. Johnson	MORT comprises a fault tree that serves as a checklist to analyse potential causes either due to management failure or a failure in the technical control. The fault tree considers factors like technology, process and human.
31	A failure modes and mechanisms naming taxonomy [147]	O'Halloran et al	The taxonomy is developed to improve the accuracy of reliability analyses during the early stages of design. It aims to help analyses techniques such as FMECA and FTA by providing a hierarchical failure mode and mechanism taxonomy and the ability to correctly classify failures analyses in reliability engineering. The taxonomy focuses primarily on technical properties and takes reference from Failure Mode/Mechanism Distributions (FMD) documentation, such as the FMD-2016.
32	System of systems acquisition trade-offs [149]	Burton, Paige, Poulding and Smith	The Defence Lines of Development (DLoD) components (i.e. Training, Equipment, Personnel, Infrastructure, Concepts & Doctrine, Organisation, Information and Logistics) are fundamental components to guide military through life capability management for the UK Ministry of Defence
33	Defence capability development manual [150]	D. J. Hurley	The Fundamental Inputs to Capability from the Department of Defence in Australia are Organisation, Command and Management, Personnel, Collective Training, Major Systems, Facilities and Training, Supplies, Support and Industry.

No.	Title	Author	Description
34	Joint publication 1-02 - Department of Defence dictionary of military and associated terms [151]	S. A. Fry	The DOTLMPF from the United States Department of Defence refers to Doctrine, Organization, Training, Leadership, Materiel, Personnel, Facilities and Policy.
35	Toward a Capability Engineering Process [152]	M. Lizotte et al	The PRICIE from the Canada Department of Defence refers to Personnel and Leadership; Research and Development, Infrastructure, Environment and Organization; Concepts and Doctrine; Information Management and Technology; Equipment and Support;

Annex B – Goal Setting Guide

Target Uncertainty: < Describe the targeted uncertainty here >		**Structural**: < e.g. platform, process, routines, reviews, milestones >	**Behavioural**: < e.g. attitude, mindset, soft-skill, experience-based, anticipation, acknowledgement >	**Relational**: < e.g. communication, dialogue, collaboration, visibility, metaphors >
1: **Track Sensors**	1a: **Decide Monitoring Technique** – What shall we track?	Track by process: e.g. periodic review Track by incident: e.g. monitor threshold Determine technique to track causal conditions dependency	Track by chance (e.g. like hazard logging) Track by proper planning (e.g. like hazard tracking system)	Track by separate observer Track by involved actors
	1b: **Form Monitoring Activity** – How do we track?	Create sensing process and review gates Create structure of reporting observations Create measurements to quantify and/or qualify uncertainty Create resources (e.g. time, infrastructure, funding) Scrutinise iterative deviation, prototype, demo (e.g. trial, test)	Use of soft skills (e.g. flexibility, tenacity, resilience. optimism) Create opportunities to be sensitive and caution about specific conditions Create positive tension among specialists with diverse perspectives	Create opportunities for collaboration (e.g. workshop, scenario building) Create engagement opportunities to communicate among stakeholders Established channels to escalate observations (e.g. proactive identification vs fire-fighting) Establish dedicated roles and responsibilities

Target Uncertainty: < Describe the targeted uncertainty here >		**Structural**: < e.g. platform, process, routines, reviews, milestones >	**Behavioural**: < e.g. attitude, mindset, soft-skill, experience-based, anticipation, acknowledgement >	**Relational**: < e.g. communication, dialogue, collaboration, visibility, metaphors >
		Establish areas of certainty separately from uncertainty		
2: **Address Responses**	2a: **Set Trigger Points** – When and how shall we decide to respond?	Establish threshold and triggering limit Establish decision making protocol and governance Decide on quality of evidence (e.g. data, subjective judgement)	Develop decision-making skills Create atmosphere of trust and cooperation to build consensus Establish risk appetite (e.g. higher threshold treats observation as one-off or lower threshold treats it as critical)	Use metaphor to conceptualise and orientate decision making Manage stakeholders' awareness and expectations (e.g. operators, end users, contractors, manufacturer) Determine composition of decision-makers (e.g. individual vs collective)
	2b: **Adapt to Change** – What are the possible responses?	Decide on reach and depth (e.g. localised, system-wide) Decide on duration of response (e.g. one-time, longer term) Decide on the additional resources required	Decide on respond urgency (e.g. reactive in nature, proactive plan already in place)	Decide on response agility (e.g. iterative changes that move in the right direction of safety, less responsive but absolute changes by getting full consensus)

Note: The three categories of structural, behavioural and relational are complementary such that an initiative can encompass characteristics from one or more of the categories.

Annex C – Sequence of Events for Yongwen Railway Accident

Time	Event
19:30	Lightning strike causes fuse to blow in the power supply circuit of the data acquisition drive unit at the Wenzhou South Station Train Control Center (TCC). Due to a design flaw, the output signal at the TCC will continue to report the status of the track prior to the fused being blown. Hence, even when the rails were occupied after the fuse was blown, the TCC output continued to indicate that the track was not occupied (which is the state prior to the fuse being blown). Lightning also caused sporadic communication failure between track circuit 5829AG and TCC, which resulted in corrupted signal being sent from 5829AG. This resulted in a red warning on the computer terminal in the TCC. The red warning implied that the track was either occupied or in a failure state.
19:39	Watchman in TCC noticed the red warning and reported the observation to the train dispatcher at Shanghai railway bureau. Watchman also informed the signalling engineer to carry out inspection and maintenance.
19:45	Signalling engineers replaced some transmitters that communicate with 5829AG in TCC without putting the equipment out of service.
19:54	According to the regulations, train dispatcher turned TCC into the emergency control mode. In the emergency control mode, the autonomous control by each train shall be deactivated and control is centralised managed by the TCC.
20:09	Train dispatcher informed D3115 driver about the red warning observed at the TCC. Train dispatcher also advised D3115 driver to switch to visual driving mode and continue driving if the train stops because of missing signals. The driver acknowledged the advice from the train dispatcher.
20:12	D301 stopped at Yongwen Station (the station before Wenzhou South Station) waiting for the signals to proceed. The train had been delayed by 36 minutes. At 20:14, D3115 left Yongwen Station.
20:17	Train dispatcher informed D3115 driver to switch to visual driving mode and drive at a speed below 20 km/h when the passing signal was red.
20:21	Due to the track circuit failure, the Automatic Train Protection (ATP) system on D3115 activated the automatic braking function. D3115 stopped in the faulted 5829AG track section. From 20:21 to 20:28, the D3115 driver had failed 3 times to start the train due to the automatic braking function in place.
20:22-27	Due to communication failure from the lightning strike, D3115 driver had failed 6 time in his attempt to contact the train dispatcher and the watchman in TCC had failed 3 times to contact D3115 driver.

Time	Event
20:24	D301 left Yongwen station for Wenzhou South Station.
20:26	Train dispatcher asked the watchman in Wenzhou South Station about D3115 and the watchman replied, "*D3115 is close to the faulted track section but the driver is out of reach, I will continue to contact him*."
20:27	Watchman reached the driver of D3115, and the driver reported, "*the train is 3 block sections to the Wenzhou South Station, but I failed to switch to visual driving mode due to abnormal track signals. I cannot reach the train dispatcher because the communication system has no signal and I will try again*."
20:28	Driver of D3115 failed twice to contact the dispatcher.
20:29	D3115 succeeded in starting the train by switching to the visual driving mode.
20:29	Engineer in Wenzhou South Station tried to warn the driver of D301 about the presence of D3115. At the same time, D301 entered the faulted track section behind D3115. The driver of D301 saw the slowly moving D3115 and launched emergency brake.
20:30	D301 travelling at the speed of 99km/h collided with D3115 travelling at the speed of 16km/h at track circuit 5829AG.

Annex D – Chapter 5 STPA-TAG: Prioritisation of Observations 2, 3 & 4

D.1. Uncertainty Observation 2: Infrastructure Support by Shanghai Railway Bureau

The CRMSA model for uncertainty observation 2 is reproduced in Figure 88. The responses to the guided questions from section 5 of the T.A.G. approach user guide is shown in Table 51.

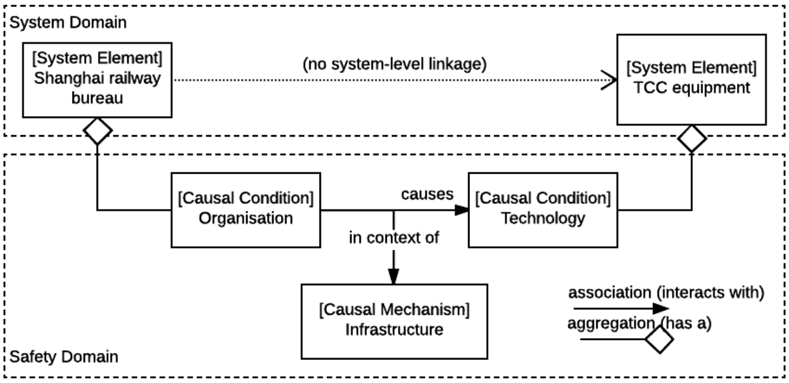

System Domain
[System Element] Shanghai railway bureau
(no system-level linkage)
[System Element] TCC equipment
[Causal Condition] Organisation
causes
[Causal Condition] Technology
in context of
[Causal Mechanism] Infrastructure
association (interacts with)
aggregation (has a)
Safety Domain

Figure 88 Uncertainty in the Infrastructure Support by Shanghai Railway Bureau

Table 51. Responses to Guided Questions for Uncertainty Observation 2

Factors	Questions
Criticality: Assess the criticality of the target to affect system safety from the stakeholders' perspectives.	
Probability	Q1. How likely will this target cause harm to the system? May not be directly. But poor infrastructure support can reduce the quality of the TCC equipment and indirectly lead to equipment failure or not operating safely as intended.
Severity	Q2. How serious are the consequences if this target occurs? Not having reliable communication in the TCC equipment can be dangerous to operation.
Complexity: Assess the complexity of the system related to the target from the stakeholders' perspectives.	
Safety Causal Model	Q3. How complex are the safety causal models from where the target is identified? (e.g. no. of related causal conditions, hazard interdependency, residual risk interdependency) No safety causal model involved
System Model	Q4: How complex are the system models from where target is identified? (e.g. system structure / product size / product design, process / task interdependency, process design, technology interdependency, project management difficulties)

Factors	Questions
	May potentially be complicated to understand the infrastructure support policy and actual implementation between shanghai railway bureau and TCC
Model Type	Q5: How complex are the model types from where target is identified? (e.g. reach and depth of patterns, templates and conventions used by the safety and system models) Unknown infrastructure support model. Need to find out more.
Novelty: Assess the novelty of knowledge related to the target from the stakeholders’ perspectives.	
Safety Novelty	Q6: How much of past information was used to define the safety causal model from where the target is identified? (e.g. causal factors, hazard list) No safety causal model involved.
System Novelty	Q7: How much of past information was used to define the system model from where the target is identified? (e.g. system legacy, process, technology, objective) Past infrastructure support document should be available but may not be easily obtainable.
Model Type Novelty	Q8: How much of past information was used to define the model type from where the target is identified? (e.g. patterns, templates, conventions) Types of infrastructure support should be following standard policies and orders.
Experience	Q9: Are there other prior experiences related to the target? This is a new system link between shanghai railway bureau and TCC to be investigated.
Resource Availability: Assess the resource availability of knowledge related to the target from the stakeholders’ perspectives.	
Capability	Q10: Do we have the capability to learn the knowledge needed? (e.g. manpower, skills, infrastructure and support, environment, other resources) May need to gain access to information regarding infrastructure support under the Yongwen railway system. No information immediately available.
Urgency	Q11: Do we have the time required to learn the knowledge needed? (e.g. project deadline, lifecycle milestones) Depends on how much time needed for the request information to be available.

From the responses to the guided questions, we assess the criticality, complexity, novelty, and resource availability of managing the epistemic uncertainty about level of technical infrastructure support to provide reliable communication for the TCC equipment by Shanghai railway bureau. The expected effort is calculated from the average score of complexity, novelty and resource availability. Based on criticality and expected effort, we plotted the uncertainty action on the prioritisation matrix chart (see Figure 89).

Factors	Score
Criticality	0.4
Complexity	0.2
Novelty	0.2
Resource	0.7
Expected effort	0.37

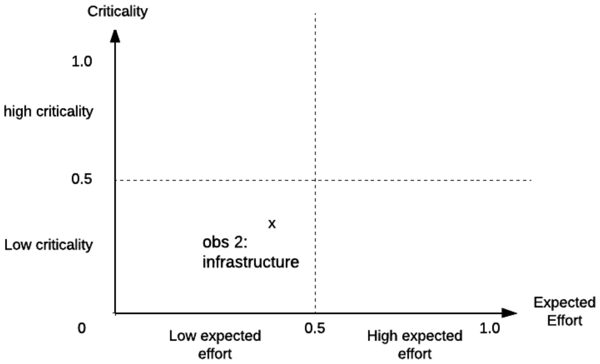


Figure 89 Uncertainty Observation 2 Plotted on Prioritisation Matrix Chart

We indicate on the bottom left quadrant the point that corresponds to the score of 0.4 criticality and 0.37 expected effort. This is the '*low hanging fruit*' quadrant, which means that stakeholders should consider allocating resources on the uncertainty actions regarding the infrastructure support by Shanghai railway bureau on TCC equipment only when there are still resources available after focusing on the high criticality issues.

D.2. Uncertainty Observation 3: Software Algorithm during Signal Abnormality

The CRMSA model for uncertainty observation 3 is reproduced in Figure 90. The responses to the guided questions from section 5 of the T.A.G. approach user guide is shown in Table 52.

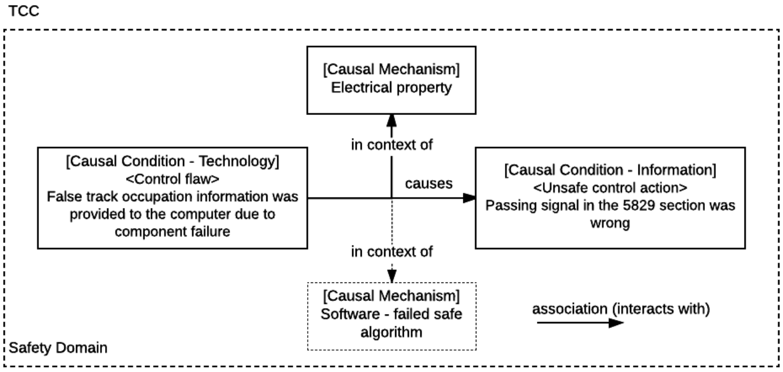

TCC
[Causal Mechanism]
Electrical property
in context of
[Causal Condition - Technology]
<Control flaw>
False track occupation information was provided to the computer due to component failure
causes
[Causal Condition - Information]
<Unsafe control action>
Passing signal in the 5829 section was wrong
in context of
[Causal Mechanism]
Software - failed safe algorithm
association (interacts with)
Safety Domain

Figure 90 Uncertainty in the Role of Software during Signal Abnormality

Table 52. Responses to Guided Questions for Uncertainty Observation 3

Factors	Questions
Criticality: Assess the criticality of the target to affect system safety from the stakeholders' perspectives.	

Factors	Questions
Probability	Q1. How likely will this target cause harm to the system? Very likely, since component failure will most likely lead to this fault.
Severity	Q2. How serious are the consequences if this target occurs? Passing wrong signal can lead to serious harm, especially collision
Complexity: Assess the complexity of the system related to the target from the stakeholders' perspectives.	
Safety Causal Model	Q3. How complex are the safety causal models from where the target is identified? (e.g. no. of related causal conditions, hazard interdependency, residual risk interdependency) Fairly complex, unsure how the software algorithm that controls the train movement can affect other software or hardware components.
System Model	Q4: How complex are the system models from where target is identified? (e.g. system structure / product size / product design, process / task interdependency, process design, technology interdependency, project management difficulties) Fairly complex, unsure if the software can affect other system components when malfunction.
Model Type	Q5: How complex are the model types from where target is identified? (e.g. reach and depth of patterns, templates and conventions used by the safety and system models) Unknown software algorithm. Unknown if any fail-safe algorithm has been implemented. Need to find out more.
Novelty: Assess the novelty of knowledge related to the target from the stakeholders' perspectives.	
Safety Novelty	Q6: How much of past information was used to define the safety causal model from where the target is identified? (e.g. causal factors, hazard list) No in-depth knowledge about the software.
System Novelty	Q7: How much of past information was used to define the system model from where the target is identified? (e.g. system legacy, process, technology, objective) Past software support document should be available but may not be easily obtainable.
Model Type Novelty	Q8: How much of past information was used to define the model type from where the target is identified? (e.g. patterns, templates, conventions) Software algorithm and support should be following standard documentation
Experience	Q9: Are there other prior experiences related to the target? This is a fairly new investigation to focus on the software algorithm.
Resource Availability: Assess the resource availability of knowledge related to the target from the stakeholders' perspectives.	
Capability	Q10: Do we have the capability to learn the knowledge needed? (e.g. manpower, skills, infrastructure and support, environment, other resources)

Factors		Questions
	Urgency	May need software experts to support the investigation. No information immediately available. Q11: Do we have the time required to learn the knowledge needed? (e.g. project deadline, lifecycle milestones) Depends on how much time needed for the request information to be available.

Factors	Score
Criticality	0.9
Complexity	0.7
Novelty	0.6
Resource	0.7
Expected effort	0.67

From the responses to the guided questions, we assess the criticality, complexity, novelty, and resource availability of managing the epistemic uncertainty about the ability to use software to create failed-safe algorithm or highlight potential abnormality when false signals are received. The expected effort is calculated from the average score of complexity, novelty and resource availability. Based on criticality and expected effort, we plotted the uncertainty action on the prioritisation matrix chart (see Figure 91).

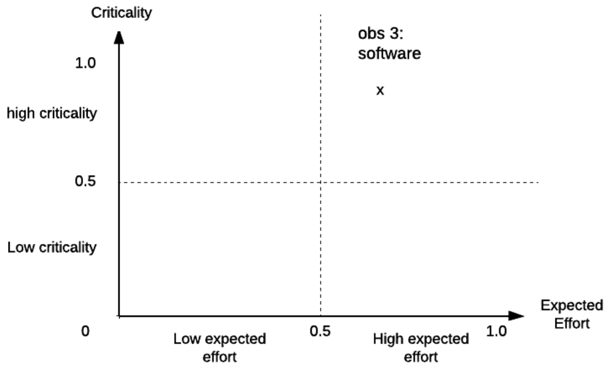


Figure 91 Uncertainty Observation 3 Plotted on Prioritisation Matrix Chart

We indicate on the top right quadrant the point that corresponds to the score of 0.9 criticality and 0.67 expected effort. This is the '*pay more attention*' quadrant, which means that stakeholders should consider allocating more resources on the uncertainty actions regarding the plausible use of software to create failed-safe algorithm or highlight potential abnormality when false signals are received at the TCC.

D.3. Uncertainty Observation 4: Human Intervention during Signal Abnormality

The CRMSA model for uncertainty observation 5 is reproduced in Figure 92. The responses to the guided questions from section 5 of the T.A.G. approach user guide is shown in Table 53.

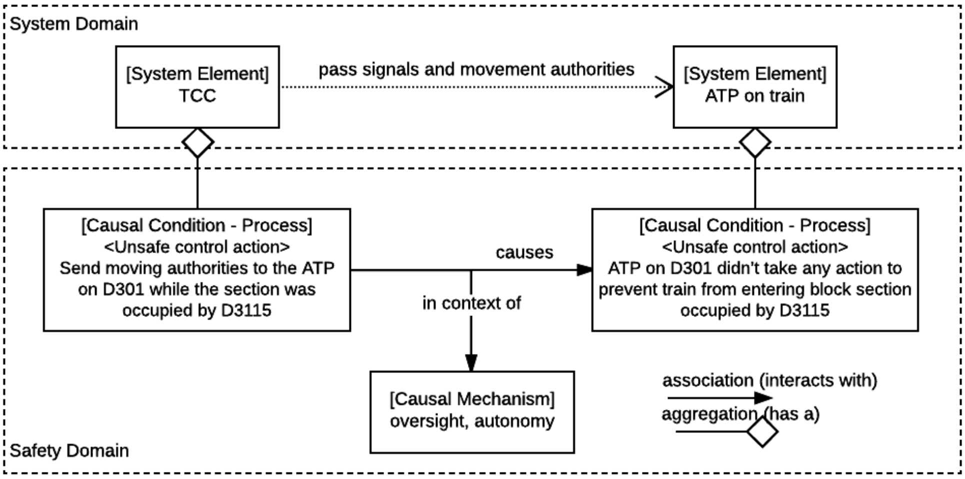

System Domain
[System Element] TCC
pass signals and movement authorities
[System Element] ATP on train
[Causal Condition - Process] <Unsafe control action> Send moving authorities to the ATP on D301 while the section was occupied by D3115
causes
in context of
[Causal Condition - Process] <Unsafe control action> ATP on D301 didn't take any action to prevent train from entering block section occupied by D3115
[Causal Mechanism] oversight, autonomy
association (interacts with)
aggregation (has a)
Safety Domain

Figure 92 Uncertainty about Human Interrupt during Signal Abnormality

Table 53. Responses to Guided Questions for Uncertainty Observation 4

Factors	**Questions**
Criticality: Assess the criticality of the target to affect system safety from the stakeholders' perspectives.	
Probability	Q1. How likely will this target cause harm to the system? Very likely, if the level of oversight and autonomy are not set to a safe level and not validated with all foreseeable scenarios.
Severity	Q2. How serious are the consequences if this target occurs? Wrong signals and movement authorities can lead to catastrophic collisions.
Complexity: Assess the complexity of the system related to the target from the stakeholders' perspectives.	
Safety Causal Model	Q3. How complex are the safety causal models from where the target is identified? (e.g. no. of related causal conditions, hazard interdependency, residual risk interdependency) Need to have an in-depth appreciation about the human-machine interface and integration between the TCC and the train to find out the underlying causal relationships.
System Model	Q4: How complex are the system models from where target is identified? (e.g. system structure / product size / product design, process / task interdependency, process design, technology interdependency, project management difficulties) Need to have an in-depth appreciation about the human-machine interface and integration between the two system components: TCC and the train
Model Type	Q5: How complex are the model types from where target is identified? (e.g. reach and depth of patterns, templates and conventions used by the safety and system models) Documented MMI should be available but the actual operational practice would need to be discovered. Need to find out more.
Novelty: Assess the novelty of knowledge related to the target from the stakeholders' perspectives.	
Safety Novelty	Q6: How much of past information was used to define the safety causal model from where the target is identified? (e.g. causal factors, hazard list)

Factors	Questions
	Machine operation is based on past experiences, but not the human-machine interactions.
System Novelty	Q7: How much of past information was used to define the system model from where the target is identified? (e.g. system legacy, process, technology, objective) Need to find out more about the human-machine interactions.
Model Type Novelty	Q8: How much of past information was used to define the model type from where the target is identified? (e.g. patterns, templates, conventions) Potential to take reference from available literature about human-machine interactions
Experience	Q9: Are there other prior experiences related to the target? Interviews may not to be conducted with the operators.
Resource: Assess the resource availability related to the target from the stakeholders' perspectives.	
Capability	Q10: Do we have the capability to learn the knowledge needed? (e.g. manpower, skills, infrastructure and support, environment, other resources) No special skills needed, focus on governance and processes. Need manpower to conduct interviews.
Urgency	Q11: Do we have the time required to learn the knowledge needed? (e.g. project deadline, lifecycle milestones) Depends on how much time needed for the request information to be available. May need to conduct interview and understand more about actual operation practice.

From the responses to the guided questions, we assess the criticality, complexity, novelty, and resource availability of managing the epistemic uncertainty about the extent of human interrupt during autonomous operation and the possibility of human oversight in the process of handling abnormality, such as sending moving authorities when it is unsafe. The expected effort is calculated from the average score of complexity, novelty and resource availability. Based on criticality and expected effort, we plotted the uncertainty action on the prioritisation matrix chart (see Figure 93).

Factors	Score
Criticality	0.8
Complexity	0.6
Novelty	0.5
Resource	0.6
Expected effort	0.57

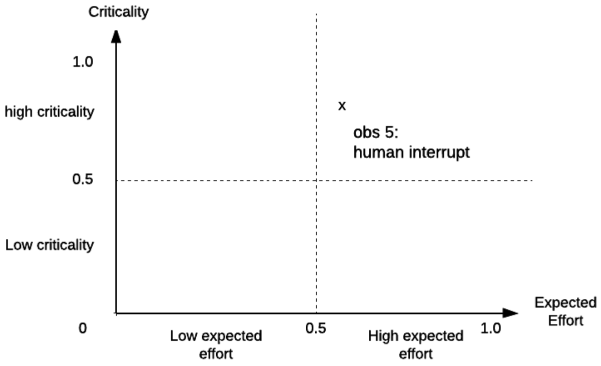


Figure 93 Uncertainty Observation 5 Plotted on Prioritisation Matrix Chart

We indicate on the top right quadrant the point that corresponds to the score of 0.8 criticality and 0.57 expected effort. This is the '*pay more attention*' quadrant, which means that stakeholders should consider allocating more resources on the uncertainty actions regarding the extent of human interrupt during autonomous operation and the possibility of human oversight in the process of handling abnormality, such as sending moving authorities when it is unsafe.

Annex E – Chapter 6 FTA-TAG: Prioritisation of Observations 2

The epistemic uncertainty is about the logic error in the BSCU command messages. The answers to the guided questions from the T.A.G. approach user guide is shown in Table 54.

Table 54. Answers to Guided Questions for Uncertainty Observation 2

Factors	Questions
Criticality: Assess the criticality of the target to affect system safety from the stakeholders' perspectives.	
Probability	Q1. How likely will this target cause harm to the system? May not be often as it happens only when abnormal messages are sent. Need information on other unintended messages.
Severity	Q2. How serious are the consequences if this target occurs? Can potentially lead to the loss of brake functions.
Complexity: Assess the complexity of the system related to the target from the stakeholders' perspectives.	
Safety Causal Model	Q3. How complex are the safety causal models from where the target is identified? (e.g. no. of related causal conditions, hazard interdependency, residual risk interdependency) No safety causal model involved.
System Model	Q4: How complex are the system models from where target is identified? (e.g. system structure / product size / product design, process / task interdependency, process design, technology interdependency, project management difficulties) Complex understanding of the messages that can be communicate with the BSCU from other systems.
Model Type	Q5: How complex are the model types from where target is identified? (e.g. reach and depth of patterns, templates and conventions used by the safety and system models) It follows conventional appreciation of communication protocols.
Novelty: Assess the novelty of knowledge related to the target from the stakeholders' perspectives.	
Safety Novelty	Q6: How much of past information was used to define the safety causal model from where the target is identified? (e.g. causal factors, hazard list) No safety causal model involved
System Novelty	Q7: How much of past information was used to define the system model from where the target is identified? (e.g. system legacy, process, technology, objective) Not new but need to research on the possible messages communicated.
Model Type Novelty	Q8: How much of past information was used to define the model type from where the target is identified? (e.g. patterns, templates, conventions) Not new model type.
Experience	Q9: Are there other prior experiences related to the target?

Factors	Questions
	Not sure if there are individuals with experience in data communication.
Resource: Assess the resource availability related to the target from the stakeholders' perspectives.	
Capability	Q10: Do we have the capability to learn the knowledge needed? (e.g. manpower, skills, infrastructure and support, environment, other resources) No specialised skill set required but need the interconnect specification for the messages.
Urgency	Q11: Do we have the time required to learn the knowledge needed? (e.g. project deadline, lifecycle milestones) Yes, there should be sufficient time to find out the information.

From the answers to the guided questions, we assess the criticality, complexity, novelty, and resource availability of managing the epistemic uncertainty about the logic error in the BSCU command messages. The expected effort is calculated from the average score of complexity, novelty and resource availability. Based on criticality and expected effort, we plotted the uncertainty action on the prioritisation matrix chart (see Figure 94).

Factors	Score
Criticality	0.7
Complexity	0.8
Novelty	0.7
Resource	0.7
Expected effort	0.7

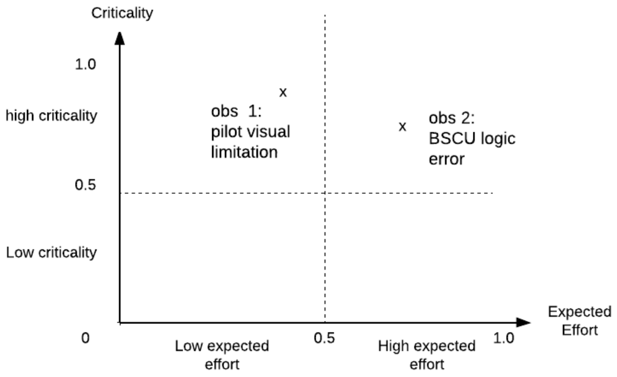


Figure 94 Uncertainty Observations Plotted on Prioritisation Matrix Chart

We indicate on the top right quadrant the point that corresponds to the score of 0.7 criticality and 0.7 expected effort. For this quadrant, stakeholders should consider putting in resources to manage this uncertainty since it is critical and the expected effort is high.

Annex F – Wheel Brake System Description

The Wheel Brake System is installed on the two main landing gears. Braking on the main gear wheels is used to provide safe retardation of the aircraft during taxiing and landing phases, and in the event of a rejected take-off. The wheel brake system is shown in Figure 95.

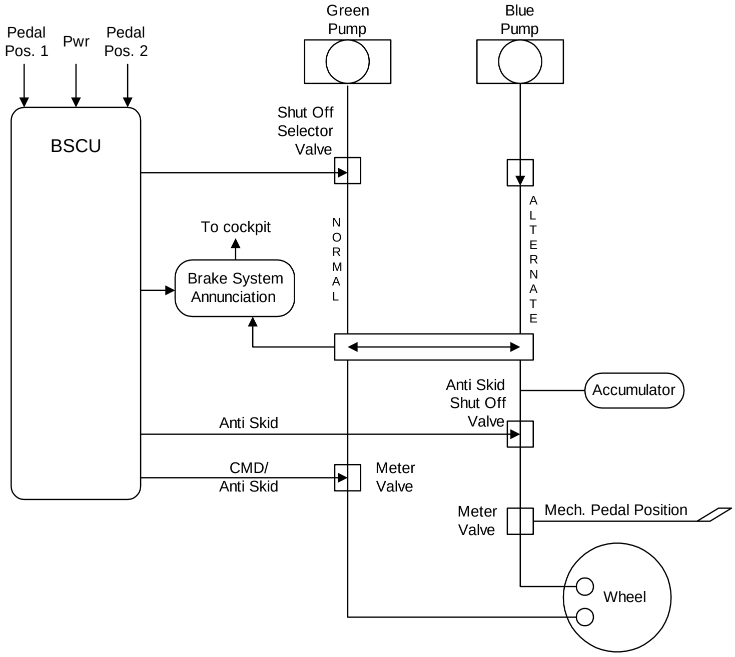


Figure 95. Wheel Break System (WBS) Components Breakdown

The wheel brakes also prevent unintended aircraft motion when parked and may be used to provide differential braking for aircraft directional control. A secondary function of the wheel brake system is to stop main gear wheel rotation upon gear retraction.

Braking on the ground is commanded either manually, via brake pedals, or automatically (autobrake) without the need for pedal application. The Autobrake function allows the pilot to prearm the deceleration rate prior to takeoff or landing. Autobrake is only available with the NORMAL braking system.

The eight main gear wheels have multi-disc carbon brakes. Based on the requirement that loss of all wheel braking is less probable than 5E-7 per flight, a design decision was made that each wheel has a brake assembly operated by two independent sets of hydraulic pistons. One set is operated from the GREEN hydraulic supply and is used in the NORMAL braking mode. The Alternate Mode is on standby and is selected automatically when the NORMAL system fails. It is operated independently using the BLUE hydraulic power supply and is backed by an accumulator which is also used to drive the parking brake. The accumulator supplies the ALTERNATE system in the EMERGENCY braking mode, when the BLUE supply is lost and the NORMAL mode is not available. Switch-over is automatic under various failure conditions or can be manually selected. Reduction of GREEN pressure below a threshold value, either from loss of GREEN supply itself or from its removal by the BSCU due to the presence of faults, causes an automatic selector to connect the BLUE supply to the ALTERNATE brake system. An anti-skid facility is available in both the NORMAL and ALTERNATE modes and operates at all speeds greater than 2 meters per second.

In the NORMAL braking mode, all eight wheels are individually braked from their own servo valves, which are also used to apply anti-skid. In the ALTERNATE mode, a dual metering valve provides a low-pressure hydraulic braking input via four servo valves which provide the antiskid function to four pairs of wheels. Operation of the ALTERNATE system is precluded when the NORMAL system is in use.

In the NORMAL mode, the brake pedal position is electrically fed to a braking computer. This in turn produces corresponding control signals to the brakes. In addition, this computer monitors various signals which denote certain critical aircraft and system states, to provide correct brake functions and improve system fault tolerance, and generates warnings, indications and maintenance information to other systems. This computer is accordingly named the Braking System Control Unit (BSCU). It automatically provides the following functions.

a. Takeover of manual braking (brake pedals), or automatic controls (engagement of Autobrake, autopilot commands during CAT IIIb landing)

b. Control of interfaces with other aircraft systems (Editor’s Note: Interfaces with other systems may include the hydraulic system, the brake temperature monitoring system, etc.)

c. Generation of braking commands, according to commands received and the status of the system

d. Braking regulation in order to avoid skidding of the main wheels

e. Transmission of information (indications, lights, warnings, etc.) to the flight deck and to the various aircraft computers concerning the BSCU status

Annex G – Chapter 6 FTA-TAG: Action Plan of Observation 2

The action plan is formulated by answering a series of questions using the T.A.G. approach user guide. The details are reproduced here in Table 55.

Table 55. Guide to Develop Action Plan for Uncertainty 2

Target Uncertainty: Uncertainty about the logic error in command messages	
1: **Track Sensors**	
1a: **Decide Monitoring Technique** – What shall we track?	Q1. What are the monitoring techniques needed to track the epistemic uncertainties? • Interfaces between BSCU and other WBS components such as Interface Control Document (ICD) • Separate manpower (communication engineers) are needed to conduct the investigation. Q2. Why are the monitoring techniques able to meet the goal of managing the uncertainty condition? • Having the interfaces and ICD will allow the engineers to investigate the command messages that are passed between the BSCU and the other WBS components. This is needed to sieve out any potential unintended logic error in the command messages that are communicated between the components during operations.
1b: **Form Monitoring Activity** – How do we track?	Q3. What types of monitoring activity are to be taken? • Investigate the communication messages that are transmitted between the BSCU and the other WBS components, check if there are any scenarios where correct messages are transmitted based on the technical specifications but the messages lead to unintended WBS operation. Q4. When should these monitoring activities be collected from the techniques? • Can be a paper study by reviewing the ICD and interface documentation. • Unsure if actual test on the physical system is necessary to validate the design performance. This will incur more resources and expertise. Q5. Who are responsible to do the tracking? • Communication engineers need to be assigned to capture and analyse the command messages. Q6. What are the structures and supporting resources to put in place for the tracking? • A structured management oversight appointed by the management needs to be in place to ensure data is feedback to the safety team for analysis. • Regular update to be presented at the weekly safety meeting. Q7. What are the skills, experiences and attitudes needed to do the tracking? • Require communication engineers that are competent in analysing electronics signals in WBS operation.
2: **Address Responses**	

Target Uncertainty: Uncertainty about the logic error in command messages	
2a: **Set Trigger Points** – When and how shall we decide to respond?	Q8: What are the trigger points from the monitoring activity that require proactive response? • Discovery of command message that has a logic error. • Discovery of command message that affect the safety of WBS operation. Q9. Who are responsible to decide on the respond actions? Management safety committee Q10. What are the governance and supporting structure to put in place to make the decision? • Results and observations to be presented in the weekly safety meeting. • If trigger points are met, communication engineers should escalate result directly to the management via incident report. Q11. What are the skills, experiences and attitudes needed to make the decision? • Knowledge about communication engineering • Knowledge about WBS operation. • Having the access to higher management that has the authority to review WBS operation.
2b: **Adapt to Change** – What are the possible responses?	Q12. Who is responsible to review the changes that will be put in place? • Project manager, together with communication engineers. Q13. How prepared and responsive should the system be in addressing the uncertainty? • If no data is observed within a month, project manager needs to decide on different action plan if the uncertainty is potentially still safety-critical. • Depends on the findings regarding the interfaces, project manager needs to decide on the urgency to escalate the issue to higher management. Q14. What are the structure and support resources to put in place to address the uncertainty? Process to escalate safety concerns will follow the existing safety report channels. Incident report to be submitted when trigger points are met. Potentially a review of the software-hardware interface components may be needed depending on the investigation. Q15. What are the skills, experiences and attitudes needed to address the uncertainty? Knowledge about communication engineering.

Using the guide, we have developed the following track and address action plan:

Track Portion of the Action Plan. Tracking the uncertainty involves investigating the interfaces between the BSC and other WBS components, such as the Interface Control Document (ICD). This is because having the ICD will allow the engineers to investigate the command messages that are passed between the BSCU and the other WBS components. This is needed to sieve out any potential unintended logic error in the command messages that are communicated between the components during operations.

To do that, dedicated communication engineers are needed to conduct the investigation. The communication engineers should be competent in analysing electronic signals in WBS operation.

The communication engineers are responsible to investigate the communication messages that are transmitted between the BSCU and the other WBS components, check if there are any scenarios where correct messages are transmitted based on the technical specifications but the messages lead to unintended WBS operation. This can be a paper study by reviewing the ICD and interface documentation. It is unsure if actual test on the physical system is necessary to validate the design performance. This will incur more resources and expertise.

A structured management oversight appointed by the management needs to be in place to ensure data is feedback to the safety team for analysis. Regular update is expected at the weekly safety meeting.

Address Portion of the Action Plan. The uncertainty will trigger a response when there is a discovery of command message that either has logic error or affect the safety of the WBS operation.

The management safety committee shall be responsible to decide on the respond actions. When trigger points are met, the communication engineers shall escalate result directly to the management via incident report.

A review committee that includes project manager and communication engineers shall be responsible to review any changes that will be put in place to address the uncertainty. The review committee should have knowledge about communication engineering and WBS operation. It should also have accessed to higher management that has the authority to review WBS operation.

If no data is observed within a month, the project manager needs to decide on a different action plan if the uncertainty is potentially still safety-critical. Depends on the findings regarding the software modules, the project manager needs to decide on the urgency to escalate the issue to higher management. There may potentially be a review of the software-hardware interface components depending on the investigation.

Annex H – Chapter 6 FMES-TAG: Prioritisation of Observations 3

The epistemic uncertainty is about the mechanical fault at the control valves between the BSCU and the hydraulic pumps. The answers to the guided questions from the T.A.G. approach user guide is shown in Table 56.

Table 56. Answers to Guided Questions for Uncertainty Observation 3

Factors	Questions
Criticality: Assess the criticality of the target to affect system safety from the stakeholders’ perspectives.	
Probability	Q1. How likely will this target cause harm to the system? May not be often as a valve typically should be reliable. Need more information about the failure rate of the control valve
Severity	Q2. How serious are the consequences if this target occurs? A failure in the mechanical valve can lead to catastrophic failure to the break system.
Complexity: Assess the complexity of the system related to the target from the stakeholders’ perspectives.	
Safety Causal Model	Q3. How complex are the safety causal models from where the target is identified? (e.g. no. of related causal conditions, hazard interdependency, residual risk interdependency) No safety causal model involved.
System Model	Q4: How complex are the system models from where target is identified? (e.g. system structure / product size / product design, process / task interdependency, process design, technology interdependency, project management difficulties) System model of the control valves is needed
Model Type	Q5: How complex are the model types from where target is identified? (e.g. reach and depth of patterns, templates and conventions used by the safety and system models) Since the team is mostly electrical engineers, additional knowledge is needed to appreciate mechanical models and theories.
Novelty: Assess the novelty of knowledge related to the target from the stakeholders’ perspectives.	
Safety Novelty	Q6: How much of past information was used to define the safety causal model from where the target is identified? (e.g. causal factors, hazard list) No safety causal model involved
System Novelty	Q7: How much of past information was used to define the system model from where the target is identified? (e.g. system legacy, process, technology, objective) Mechanical models about the control valves should be available
Model Type Novelty	Q8: How much of past information was used to define the model type from where the target is identified? (e.g. patterns, templates, conventions)

Factors	Questions
	Mechanical model types about the control valves should not be new concept.
Experience	Q9: Are there other prior experiences related to the target? There should be maintenance engineers with experience about control valves
Resource: Assess the resource availability related to the target from the stakeholders' perspectives.	
Capability	Q10: Do we have the capability to learn the knowledge needed? (e.g. manpower, skills, infrastructure and support, environment, other resources) Yes, is about gathering the necessary theory and failure modes about the control valves.
Urgency	Q11: Do we have the time required to learn the knowledge needed? (e.g. project deadline, lifecycle milestones) Yes, there should be sufficient time to find out the information.

From the answers to the guided questions, we assess the criticality, complexity, novelty, and resource availability of managing the epistemic uncertainty about the mechanical fault at the control valves between the BSCU and the hydraulic pumps. The expected effort is calculated from the average score of complexity, novelty and resource availability. Based on criticality and expected effort, we plotted the uncertainty action on the prioritisation matrix chart (see Figure 96).

Factors	Score
Criticality	0.8
Complexity	0.8
Novelty	0.4
Resource	0.6
Expected effort	0.6

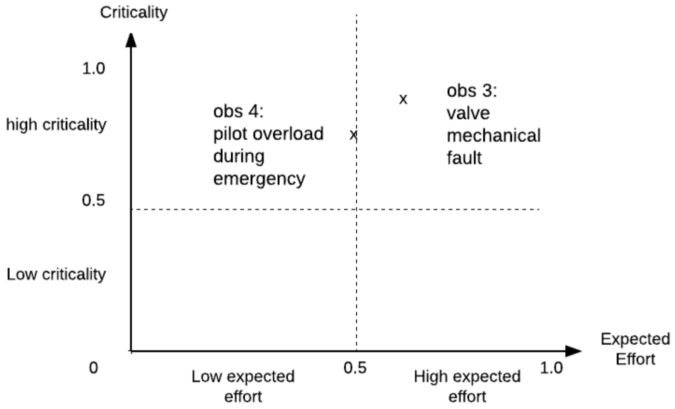


Figure 96 Uncertainty Observation 1 Plotted on Prioritisation Matrix Chart

We indicate on the top right quadrant the point that corresponds to the score of 0.8 criticality and 0.6 expected effort. For this quadrant, stakeholders should consider putting in resources to manage this uncertainty since it is critical and the expected effort is high.

Annex I – Chapter 6 FMES-TAG: Action Plan of Observation 3

The action plan is formulated by answering a series of questions using the T.A.G. approach user guide. The details are reproduced here in Table 57.

Table 57. Guide to Develop Action Plan for Uncertainty 3

Target Uncertainty: Uncertainty about the mechanical faults at control valves	
1: **Track Sensors**	
1a: **Decide Monitoring Technique** – What shall we track?	Q1. What are the monitoring techniques needed to track the epistemic uncertainties? • Physical interfaces and joints at the control valves that receive the electrical signal from the BSCU. • Separate manpower (mechanical engineers) are needed to conduct the investigation. Q2. Why are the monitoring techniques able to meet the goal of managing the uncertainty condition? • These physical interfaces represent the location where electrical signals are transmitted between the BSCU and the control valves to control the valve operation. Mechanical faults such as dry or sticky joints may result in signal corruption.
1b: **Form Monitoring Activity** – How do we track?	Q3. What types of monitoring activity are to be taken? • Investigate the mechanical contacts where the BSCU signals reach the control valves, check on the reliability and failure rate of these mechanical contacts. Q4. When should these monitoring activities be collected from the techniques? • Can be a paper study by reviewing the specification of the mechanical contacts. • Unsure if lab test or additional contractor test result/specification is needed to verify the design performance of the mechanical contacts. This will incur more resources and expertise. Q5. Who are responsible to do the tracking? • Mechanical engineers need to be assigned to analyse the mechanical performance. Q6. What are the structures and supporting resources to put in place for the tracking? • A structured management oversight appointed by the management needs to be in place to ensure data is feedback to the safety team for analysis. • Regular update to be presented at the weekly safety meeting. Q7. What are the skills, experiences and attitudes needed to do the tracking? • Require mechanical engineers that are competent in analysing physical and mechanical properties of WBS operation.
2: **Address Responses**	

Target Uncertainty: Uncertainty about the mechanical faults at control valves	
2a: **Set Trigger Points** – When and how shall we decide to respond?	Q8: What are the trigger points from the monitoring activity that require proactive response? • Discovery of mechanical conditions that can potentially compromise the electrical signal between the BSCU and the control valves and lead to signal corruption. • Discovery of specific control valves that may be subjected to the above mechanical condition that can cause mechanical faults and lead to signal corruption. Q9. Who are responsible to decide on the respond actions? • Management safety committee Q10. What are the governance and supporting structure to put in place to make the decision? • Results and observations to be presented in the weekly safety meeting. • If trigger points are met, mechanical engineers should escalate result directly to the management via incident report. Q11. What are the skills, experiences and attitudes needed to make the decision? • Knowledge about mechanical engineering • Knowledge about WBS operation. • Having the access to higher management that has the authority to review WBS operation.
2b: **Adapt to Change** – What are the possible responses?	Q12. Who is responsible to review the changes that will be put in place? • Project manager, together with mechanical engineers. Q13. How prepared and responsive should the system be in addressing the uncertainty? • If no data is observed within a month, project manager needs to decide on different action plan if the uncertainty is potentially still safety-critical. • Depends on the findings regarding the mechanical contacts, project manager needs to decide on the urgency to escalate the issue to higher management. Q14. What are the structure and support resources to put in place to address the uncertainty? • Process to escalate safety concerns will follow the existing safety report channels. • Incident report to be submitted when trigger points are met. • Potentially a review of the mechanical properties of the control valve may be needed depending on the investigation. Q15. What are the skills, experiences and attitudes needed to address the uncertainty? • Knowledge about mechanical engineering.

Using the guide, we have developed the following track and address action plan:

Track Portion of the Action Plan. Tracking the uncertainty involves investigating the physical interfaces and joints at the control valves that receive the electrical signal from the BSCU. This is because these physical interfaces represent the location where electrical signals are transmitted between the BSCU and the control valves to control the valve operation. Mechanical faults such as dry or sticky joints may result in signal corruption.

To do that, dedicated mechanical engineers are needed to conduct the investigation. The communication engineers should be competent in analysing physical and mechanical properties of WBS operation

The communication engineers are responsible to investigate the mechanical contacts where the BSCU signals reach the control valves, check on the reliability and failure rate of these mechanical contacts. This can be a paper study by reviewing the specification of the mechanical contacts. It is unsure if lab test or additional contractor test result/specification is needed to verify the design performance of the mechanical contacts. This will incur more resources and expertise.

A structured management oversight appointed by the management needs to be in place to ensure data is feedback to the safety team for analysis. Regular update is expected at the weekly safety meeting.

Address Portion of the Action Plan. The uncertainty will trigger a response when there is a discovery of mechanical conditions that can potentially compromise the electrical signal between the BSCU and the control valves and lead to signal corruption. Discovery of specific control valves that may be subjected to the above mechanical condition will also trigger a response from the investigation team.

The management safety committee shall be responsible to decide on the respond actions. When trigger points are met, the mechanical engineers shall escalate result directly to the management via incident report.

A review committee that includes project manager and mechanical engineers shall be responsible to review any changes that will be put in place to address the uncertainty. The review committee should have knowledge about mechanical engineering and WBS operation. It should also have accessed to higher management that has the authority to review WBS operation.

If no data is observed within a month, the project manager needs to decide on a different action plan if the uncertainty is potentially still safety-critical. Depends on the findings regarding the mechanical contacts, the project manager needs to decide on the urgency to escalate the issue to higher management. There may potentially be a review of the mechanical properties of the control valve depending on the investigation.

Annex J – Evaluation Questionnaire

1. Name:

2. Job Responsibility:

3. Can you please rate your level of experience in safety assessments?

Beginner 1 2 3 4 5 Expert ( ]

For the following questions, rate how strongly you agree or disagree with the statements.

With reference to the safety assessments that you have been involved in …

4. "In my experience, almost all the information that is needed for safety assessment is available at the point in time of conducting the analysis."

Disagree 1 2 3 4 5 Agree ( ]

5. "Epistemic uncertainty (i.e. uncertainty due to the lack of information) has significantly affected the trust that I am able to place in safety assessment results."

Disagree 1 2 3 4 5 Agree ( ]

6. "Epistemic uncertainties in existing safety assessment are well managed and we are able to prevent them from leading to any hazardous situations."

Disagree 1 2 3 4 5 Agree ( ]

HOT-PIE Taxonomy

In Step 1 of the T.A.G. approach, the HOT-PIE taxonomy is established to recognise known and unknown uncertainties not covered in the original safety assessment.

With reference to the HOT-PIE taxonomy …

7. “With the artefacts and resources available, the taxonomy can augment existing safety assessment.”

Disagree 1 2 3 4 5 Agree ( ]

8. “The taxonomy will help me identify epistemic uncertainties that otherwise be easily overlooked if not prompted”

Disagree 1 2 3 4 5 Agree ( ]

9. “The amount of effort required to incorporate the taxonomy is worth it given the benefits, considering the resources and limitation of the safety assessments that I have been involved in”

Disagree 1 2 3 4 5 Agree ( ]

10. I have the following comments to support my views regarding the taxonomy:

Prioritisation Factors

In Step 2 of the T.A.G. approach, prioritisation factors are established to prioritise uncertainty analyses in existing safety assessment.

With reference to the prioritisation factors …

11. "After identifying the epistemic uncertainties that are present in the safety assessment, it is important to prioritise them for further investigations in existing safety assessments that I have been involved in."

Disagree 1 2 3 4 5 Agree ( ]

12. "The prioritisation factors help me to focus on epistemic uncertainties that are more important"

Disagree 1 2 3 4 5 Agree ( ]

13. "The amount of effort required to incorporate the prioritisation is worth it given the benefits, considering the resources and limitation of the safety assessments that I have been involved in"

Disagree 1 2 3 4 5 Agree ( ]

14. I have the following comments to support my views regarding the prioritisation factors:

Action Plan

In Step 3 of the T.A.G. approach, an action plan is established using guided questions to develop goals to track and address epistemic uncertainties through-life in anticipation that the uncertainties may change with time.

With reference to the guided questions…

15. “The guided questions will help me to formulate an action plan to track and address epistemic uncertainties in safety assessment through-life”

Disagree 1 2 3 4 5 Agree ( ]

16. “It is feasible to prepare an action plan like this with the help of the guided questions as part of safety assessment”

Disagree 1 2 3 4 5 Agree ( ]

17. I have the following comments to support my views regarding the action plan:

18. I have the following further comments about the T.A.G. approach:

Thank you for participating in the interview.

Annex K – T.A.G. Approach Users Guide

Generic Process of Tagging Uncertainties

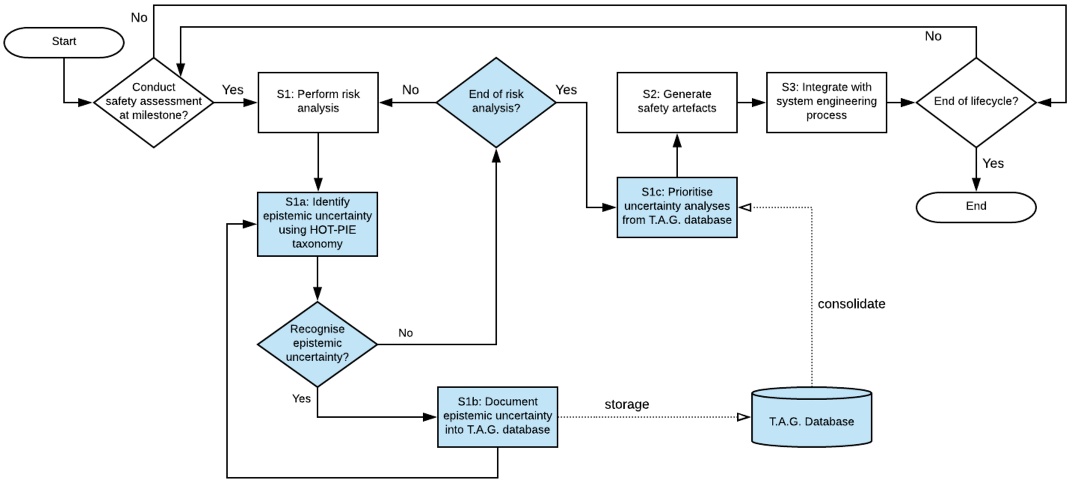

Start
Conduct safety assessment at milestone?
No
Yes
S1: Perform risk analysis
No
End of risk analysis?
Yes
S2: Generate safety artefacts
S3: Integrate with system engineering process
End of lifecycle?
No
Yes
End
S1a: Identify epistemic uncertainty using HOT-PIE taxonomy
S1c: Prioritise uncertainty analyses from T.A.G. database
consolidate
Recognise epistemic uncertainty?
No
Yes
S1b: Document epistemic uncertainty into T.A.G. database
storage
T.A.G. Database

Figure 97. Process Flow to Integrate T.A.G. Approach in Safety Assessment

Step 1 (S1): Perform Risk Analysis. The usual risk analysis per existing safety assessment technique. In addition, T.A.G. approach is integrated here by inserting the following sub steps:

Step 1a (S1a): Identify Epistemic Uncertainty. This is Step 1 of the T.A.G. approach. Stakeholders can use the HOT-PIE taxonomy as a reference to identify both known and unknown epistemic uncertainties. Stakeholders can also make use of the CMSS and CRMSA models to help in finding uncertainties in the safety causal models, system models and the model types in the safety assessment.

Step 1b (S1b): Document Epistemic Uncertainty. This is part of Step 2 of the T.A.G. approach. Any epistemic uncertainty that has been recognised, regardless if it is assessed to be safety critical, should be tagged and stored in a T.A.G. database. This database is needed in the next step for prioritisation.

Step 1c (S1c): Prioritise Uncertainty Analyses. The T.A.G. database will document causal relationships with epistemic uncertainties that may potentially be safety critical. We refer to them as plausible-but-uncertain causal relationships. Stakeholders are not able to commit if such plausible-but-uncertain causal relationships are safety critical due to a lack of information at the point of conducting the risk analysis. They would also have to assess if the risk is at least tolerable for these uncertainties to be tracked further. Due to potential lack of resources, stakeholders may need to prioritise these uncertainties for further tracking.

Step 2 (S2): Generate Safety Artefacts. Besides generating safety artefacts from safety analyses (e.g. list of hazards, risk mitigation plans, outstanding action items and residual risk report), there will also be a database of prioritised uncertainties to be analysed from the T.A.G database. All this information should be integrated into the same existing reporting channel and transferred to the system engineering process in the next step.

Step 3 (S3): Integrate with System Engineering Process. Besides the usual course of actions due to the risk analysis, the T.A.G. approach would also generate action plans to track and address epistemic uncertainties. The requirements from such action plans would be feedback to the system engineering process for further actions.

T.A.G. Approach Step 1: Identify Uncertainties using HOT-PIE taxonomy

In step 1 of the T.A.G. approach (see Figure 98), we develop a taxonomy of causal mechanisms (see Table 58) to help users identify plausible uncertainties during safety assessment.

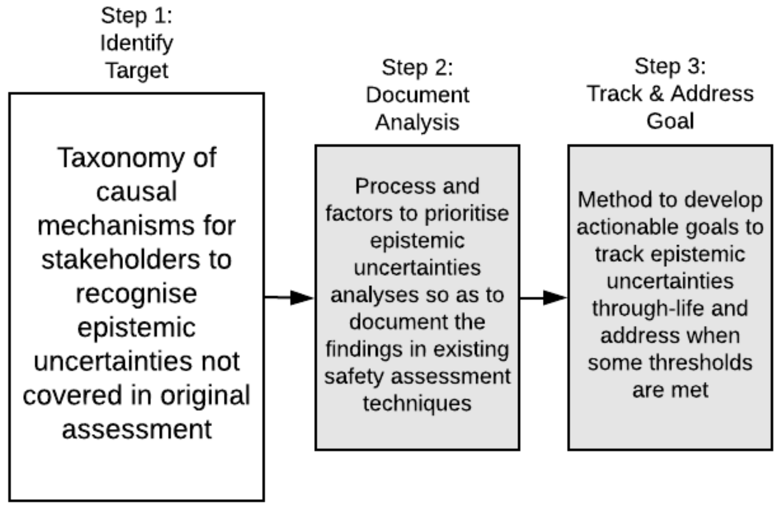

Step 1: Identify Target
Taxonomy of causal mechanisms for stakeholders to recognise epistemic uncertainties not covered in original assessment
Step 2: Document Analysis
Process and factors to prioritise epistemic uncertainties analyses so as to document the findings in existing safety assessment techniques
Step 3: Track & Address Goal
Method to develop actionable goals to track epistemic uncertainties through-life and address when some thresholds are met

Figure 98. T.A.G. Approach First Step: Identify Target

Table 58. The HOT-PIE Taxonomy of Causal Mechanisms

Primary Causal Conditions	Secondary Causal Conditions
Human	**H1: Manpower –** expertise[154-156] staffing[146, 154, 156-160] mix[158] ownership[154] experience[154, 158] leadership[142, 146] skill[146, 156, 158, 159, 161-163] ability[158] characters[164] individualistic[165] demographic[165] cultural[165] obligation[166] survivable[158] stakeholders[66, 156, 167-169] user[170] turnover[156] education[156]
	H2: Mental state – escalation[142] brokerage[142] free rider[142] convention[142] norm[142] selective benefit[142] morale and motivation [142, 156, 158] social[163, 171] deliberate[161] esteem[166] complacency[146] stress[146] overconfidence[146] fatigue[146] distraction[80, 146] confusion[146] health[158] comfort[158] visual limitation[146] illness[146] injury[158] disability[158] hearing limitation[146] cognitive[158] physical[158, 172] sensory[158] team dynamic[158, 159] aptitude[158] emotional[172] psychic[173] conflict of interest[156] lack purpose[156] perception[80, 174] memory fail[80] poor mental model[80] incorrect mental model[80] reliance on default[80]
	H3: Action – operation[155] network[142] broadcast[142] rumour[142] communication[146, 154, 156, 159] open[159] interrelation[159] atmosphere[159] engagement[173] coordination[146] omission[80, 161] commission[161] extraneous act[161] observation[164] interpretation[164] overcommit[166] performance slip[175] specification slip[175] lapse-forgot[175] lapse-overlook[175] rest[146]

Primary Causal Conditions	Secondary Causal Conditions
	preparation[146] intentional violation[159, 163, 176] behaviour[173] lack involvement[156] influence[174]
Organisation	**O1: Management –** supervision[142, 146, 155] audit[142] communication[164] structure [146, 164, 168, 174, 177] levels of domain[174] role ambiguity and conflict[165] schedule[165] demand[166] feedback and refine[146] company[160] project size[156] project uniqueness[156] project density[156]
	O2: Policy – regulation and control[142, 160, 167, 174] job future and security[165, 166], culture and climate[146, 156, 162, 165, 177, 178] reward and recognition[165, 166] incompatible goals[156, 159, 176] trade-off[159] ambiguous goal[156] narrow goal[156] expectation[156] customer satisfaction[170]
	O3: Resource – training facility[142, 155, 164, 170, 176] material[154, 155, 162] supplier management[142, 156, 169] support facility[146, 156, 161, 170, 172] time phase[157, 161] time step[157, 161] project urgency[156] allocation[146] monetary[146, 156] instructional[158] unrealistic time frame[156] outsource management[156] interdependent infrastructure[163, 170] test equipment[178] test procedure[178]
Technology	**T1: Machine –** hardware capability[155, 157, 163, 167, 169, 174, 176, 177] hardware compatibility[178] technical[168, 171, 179] equipment [146, 161, 164] interface[146, 164] link[163] node[163] display[146] construction[162] software[147, 157, 163, 167, 169, 174, 177, 178] communication[147, 170, 176] engineering[66] mobility[163] traffic[163] area coverage[163] services[170] tool[170] technique[170] abstraction[154] working range[154] tech change[154, 156] innovation[154] complexity[146] availability[159] function[159] standardisation[159] features[156] customisation[156] interdependency[163]
	T2: Property – energy[157] kinetic[154] biological[154] acoustical[154] chemical[154] electrical[154] mechanical[154] electro-magnetic[154] thermal[154] radiation[147, 154] bonding[147] buckling[147] change in property[147] corrosion[147] cracking[147] deformation[147] fatigue[147] seizure[147] impact[147] rupture[147] voiding[147] wear[147, 178] breakdown[147] contamination[147] diffusion[147] degradation[147] incorrect current[147] punch through[147] leak[178] loose[178] drift[178] synchronisation[178]
	T3: Support – system design[162, 176] tool design[165] tool usability[165] work area design[165] task design[146, 176] medium[163]
Process	**P1: Nature –** segregation[154] systematic[154] oversight[146, 154] procedure [146, 154, 157, 159, 161, 162, 164, 167, 176-178] practice[154, 167] overload[80, 165] control[157, 165] autonomy[165, 172] repetitiveness[165, 174] feedback[165, 172] ability to learn[165] input[157] output[157] lower level failure[163] cascade failure[163] delay[163]
	P2: Phase – design and plan[164, 179] validation[154] verification[154] manufacturing[66] operation[66] risk management[154, 156, 158, 176] review[154] maintenance[159, 176, 178, 179] housekeeping[176] inspection[179] supervision[179] work[160, 170, 171, 177] training[159, 161] execution and operation[146, 161, 170, 178] mis-operation[161] task[165, 168, 169] sense-making[170] decision making[170] thinking[170]
Information	**I1: Knowledge –** procedure[155] standard[155] method[155] assumption[161] policy[146, 169] rule[162, 167] guideline[157] precondition[157] type of info[164] manual and checklist[146] protocol[159, 163] roles and responsibilities[156] best practice[156] data[156] concept[156] no fault found[178] rationalities[174] evidence[174] values[174] fluctuation[174] customer requirements[170] codified information[170]
	I2: Error – application error[175] assumption error[175] syntax error[175] requirement error[175] lack of distinction[175] lack of awareness[175] insufficient knowledge[175] situational awareness error[159] incomplete specification[156] conflicting requirements[156] info processing problem[156, 170] data unavailable[80] data not detected[80] decisional error[174] executional error[174]
Environment	**E1: Physical –** transport network[142] ambient condition[161, 164] weather[146, 161] orientation[161] size[161] location[161] elevation[161] operating condition[158, 164] noise[146, 165] lighting[146, 165] vibration[146, 165] pollution[165] heat[146] terrestrial[163] meteorological[163] cosmological[163] **E2: Non-physical –** cultural[155, 170, 177] social[167] attitude[155] economic[142, 156, 163, 167, 177] competitiveness[170] political[142, 156, 163, 167, 169, 177] regulatory[170, 177] legal[156, 167] contract[142] propaganda[142] duration[161] delayed[161] alternative[166] strategic interest[166] government[160] complexity[156] security[163]

Note: The referenced materials explaining each casual mechanism are listed in the reference.

T.A.G. Approach Step 2: Prioritise Uncertainty Analyses

In step 2 of the T.A.G. approach (see Figure 99), we develop a questionnaire and prioritisation factors to help users document the way they prioritise uncertainties to be analysed (see Table 59).

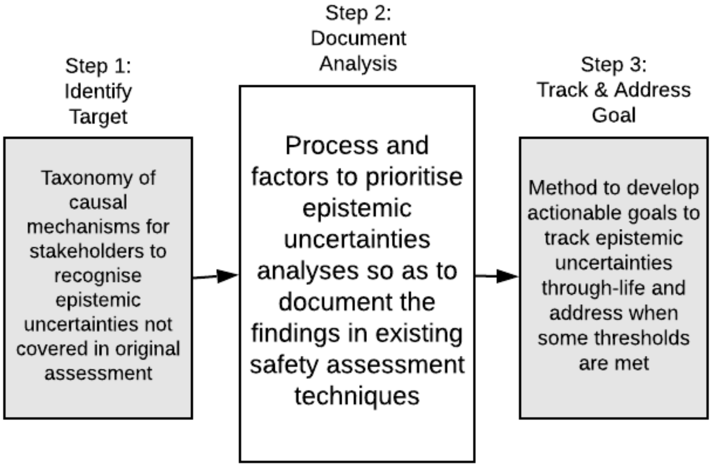

Step 1: Identify Target
Taxonomy of causal mechanisms for stakeholders to recognise epistemic uncertainties not covered in original assessment
Step 2: Document Analysis
Process and factors to prioritise epistemic uncertainties analyses so as to document the findings in existing safety assessment techniques
Step 3: Track & Address Goal
Method to develop actionable goals to track epistemic uncertainties through-life and address when some thresholds are met

Figure 99. T.A.G. Approach Second Step: Document Analysis

Instructions to prioritise uncertainty analyses:

1. Identify the uncertainty in the targeted causal relationship.

2. Using the help of the questionnaire, analyse the criticality of the target based on the guided questions (see Table 59). Score the criticality from 0 (low criticality) to 1 (high criticality).

3. Analyse the complexity of the target based on the guided questions. Score the complexity from 0 (low complexity) to 1 (high complexity).

4. Analyse the novelty of the target based on the guided questions. Score the novelty from 0 (low novelty) to 1 (high novelty).

5. Analyse the resource availability of the target based on the guided questions. Score the resource availability from 0 (high resource availability) to 1 (low resource availability).

6. Calculate the overall expected effort by average out the sum of the three scores under complexity, novelty and resource availability (assuming the factors have equal weightage).

7. Locate the quadrant on the prioritisation matrix that corresponds to the score for criticality and expected effort (see Table 60). Prioritise based on the location on the matrix.

Table 59. Guided Questions to Prioritise Uncertainty Analyses

Target Uncertainty to be analysed: < Describe the targeted uncertainty here >	
Factors	**Questions**
Criticality: Assess the criticality of the target to affect system safety from the stakeholders' perspectives.	
Probability	Q1. How likely will this target cause harm to the system?
Severity	Q2. How serious are the consequences if this target occurs?
Complexity: Assess the complexity of the system related to the target from the stakeholders' perspectives.	
Safety Causal Model	Q3. How complex are the safety causal models from where the target is identified? (e.g. no. of related causal conditions, hazard interdependency, residual risk interdependency)
System Model	Q4: How complex are the system models from where target is identified? (e.g. system structure / product size / product design, process / task interdependency, process design, technology interdependency, project management difficulties)
Model Type	Q5: How complex are the model types from where target is identified? (e.g. reach and depth of patterns, templates and conventions used by the safety and system models)
Novelty: Assess the novelty of knowledge related to the target from the stakeholders' perspectives.	
Safety Novelty	Q6: How much of past information was used to define the safety causal model from where the target is identified? (e.g. causal factors, hazard list)
System Novelty	Q7: How much of past information was used to define the system model from where the target is identified? (e.g. system legacy, process, technology, objective)
Model Type Novelty	Q8: How much of past information was used to define the model type from where the target is identified? (e.g. patterns, templates, conventions)
Experience	Q9: Are there other prior experiences related to the target?
Resource : Assess the resource availability related to the target from the stakeholders' perspectives.	
Capability	Q10: Do we have the capability to learn the knowledge needed? (e.g. manpower, skills, infrastructure and support, environment, other resources)
Urgency	Q11: Do we have the time required to learn the knowledge needed? (e.g. project deadline, lifecycle milestones)

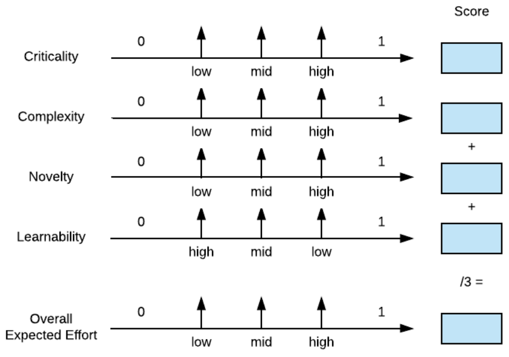
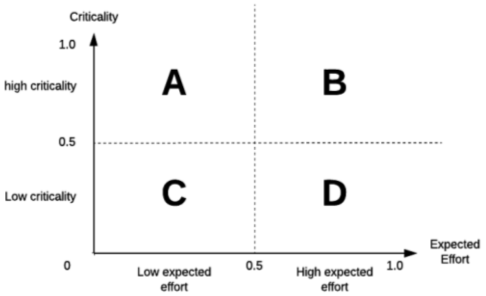


Table 60. Prioritisation Table for Uncertainty Action

Quadrant	Criticality	Expected Effort
A	Low	Low
B	Low	High
C	High	Low
D	High	High



As an example, users may decide to prioritise the uncertainty analyses based first on the criticality of the uncertainties and secondly by the amount of efforts needed (see Figure 100). With this strategy, user would decide to put high priorities to analyses in quadrants A and B, rather than C and D. Take note that this is but one of the many ways that users can prioritise the uncertainty analyses. The final choice of where to focus on depends on the context facing the users during the analysis.

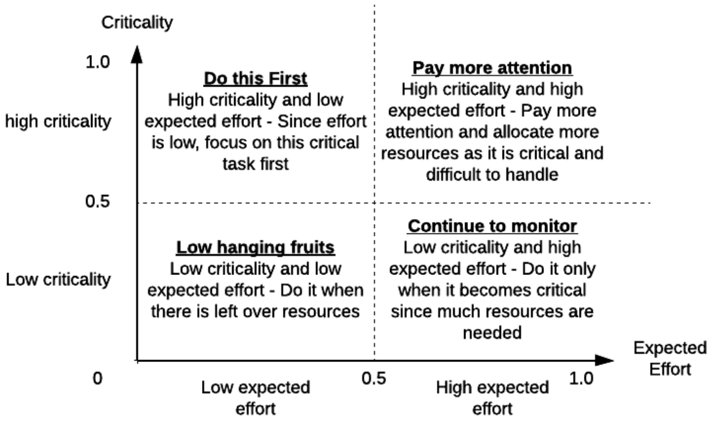

Criticality
1.0
high criticality
0.5
Low criticality
0
Do this First
High criticality and low expected effort - Since effort is low, focus on this critical task first
Pay more attention
High criticality and high expected effort - Pay more attention and allocate more resources as it is critical and difficult to handle
Low hanging fruits
Low criticality and low expected effort - Do it when there is left over resources
Continue to monitor
Low criticality and high expected effort - Do it only when it becomes critical since much resources are needed
Expected Effort
Low expected effort
0.5
High expected effort
1.0

Figure 100 An Example of Prioritisation for Uncertainty Analysis

T.A.G. Approach Step 3: Set Goals for Action Plan to Track and Address Uncertainties

In step 3 of the T.A.G. approach (see Figure 101), we develop a questionnaire (see Table 61) to help users set goals in the action plan to track the uncertainty through-life and address the uncertainty when some thresholds are met.

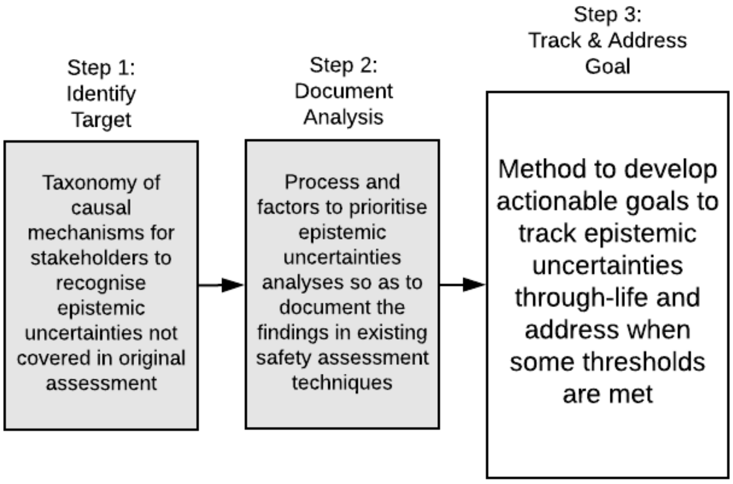

Step 1:
Identify
Target
Taxonomy of causal mechanisms for stakeholders to recognise epistemic uncertainties not covered in original assessment
Step 2:
Document
Analysis
Process and factors to prioritise epistemic uncertainties analyses so as to document the findings in existing safety assessment techniques
Step 3:
Track & Address
Goal
Method to develop actionable goals to track epistemic uncertainties through-life and address when some thresholds are met

Figure 101. T.A.G. Approach Third Step: Track and Address Goal

Table 61. Guide to Set Goals in Action Plan

Target Uncertainty: < Describe the targeted uncertainty here >		**Guided Questions to set Goals for Action Plan**
1: **Track Sensors**	1a: **Decide Monitoring Technique** – What shall we track?	Q1. What are the monitoring techniques needed to track the epistemic uncertainties? Q2. Why are the monitoring techniques able to meet the goal of managing the uncertainty condition?
	1b: **Form Monitoring Activity** – How do we track?	Q3. What types of monitoring activity are to be taken? Q4. When should these monitoring activities be collected from the techniques? Q5. Who are responsible to do the tracking? Q6. What are the structures and supporting resources to put in place for the tracking? Q7. What are the skills, experiences and attitudes needed to do the tracking?
2: **Address Responses**	2a: **Set Trigger Points** – When and how shall we decide to respond?	Q8: What are the trigger points from the monitoring activity that require proactive response? Q9. Who are responsible to decide on the respond actions? Q10. What are the governance and supporting structure to put in place to make the decision?

Target Uncertainty: < Describe the targeted uncertainty here >	**Guided Questions to set Goals for Action Plan**
2b: **Adapt to Change** – What are the possible responses?	Q11. What are the skills, experiences and attitudes needed to make the decision?
	Q12. Who is responsible to review the changes that will be put in place?
	Q13. How prepared and responsive should the system be in addressing the uncertainty?
	Q14. What are the structure and support resources to put in place to address the uncertainty?
	Q15. What are the skills, experiences and attitudes needed to address the uncertainty?

Note: A sample of considerations for the track and address plan is provided in the reference.

The following is a process flow (see Figure 102) to guide user during the follow-up actions to address the epistemic uncertainty. Note that the greater the epistemic uncertainties surround a risk analysis, the lower the confidence in the analysis.

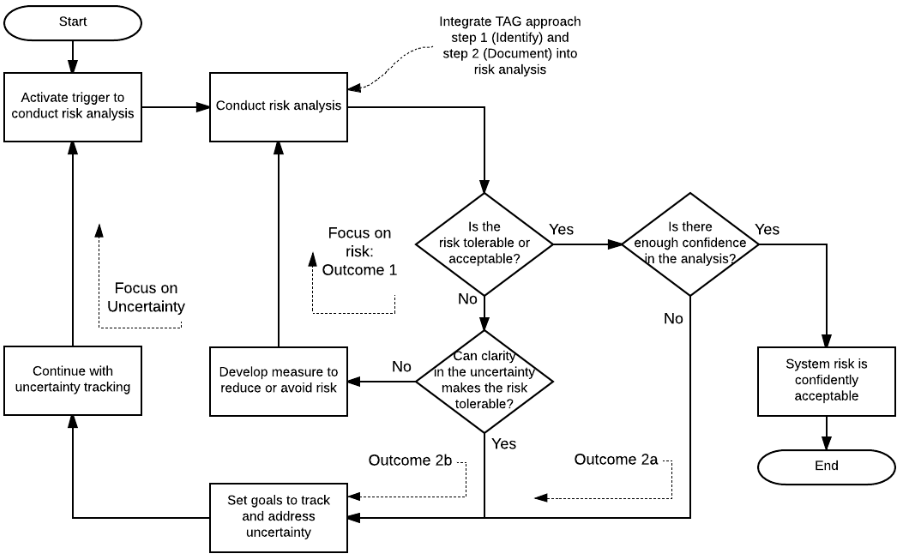

Start
Integrate TAG approach step 1 (Identify) and step 2 (Document) into risk analysis
Activate trigger to conduct risk analysis
Conduct risk analysis
Is the risk tolerable or acceptable?
Yes
Is there enough confidence in the analysis?
Yes
Focus on risk: Outcome 1
Focus on Uncertainty
No
No
System risk is confidently acceptable
End
Continue with uncertainty tracking
Develop measure to reduce or avoid risk
No
Can clarity in the uncertainty makes the risk tolerable?
Yes
Outcome 2b
Outcome 2a
Set goals to track and address uncertainty

Figure 102. Possible Outcomes when Addressing the Uncertainty

When being triggered to address the uncertainty, stakeholders would conduct a risk analysis to assess the risk that is still facing the system, as well as the assessing the confidence in the analysis. At this moment, stakeholders would have to decide if the risk is confidently within the tolerable or acceptable region. If it is not, a decision would have to be made to either focus on managing the risk or clarifying the uncertainty.

Referenced Models

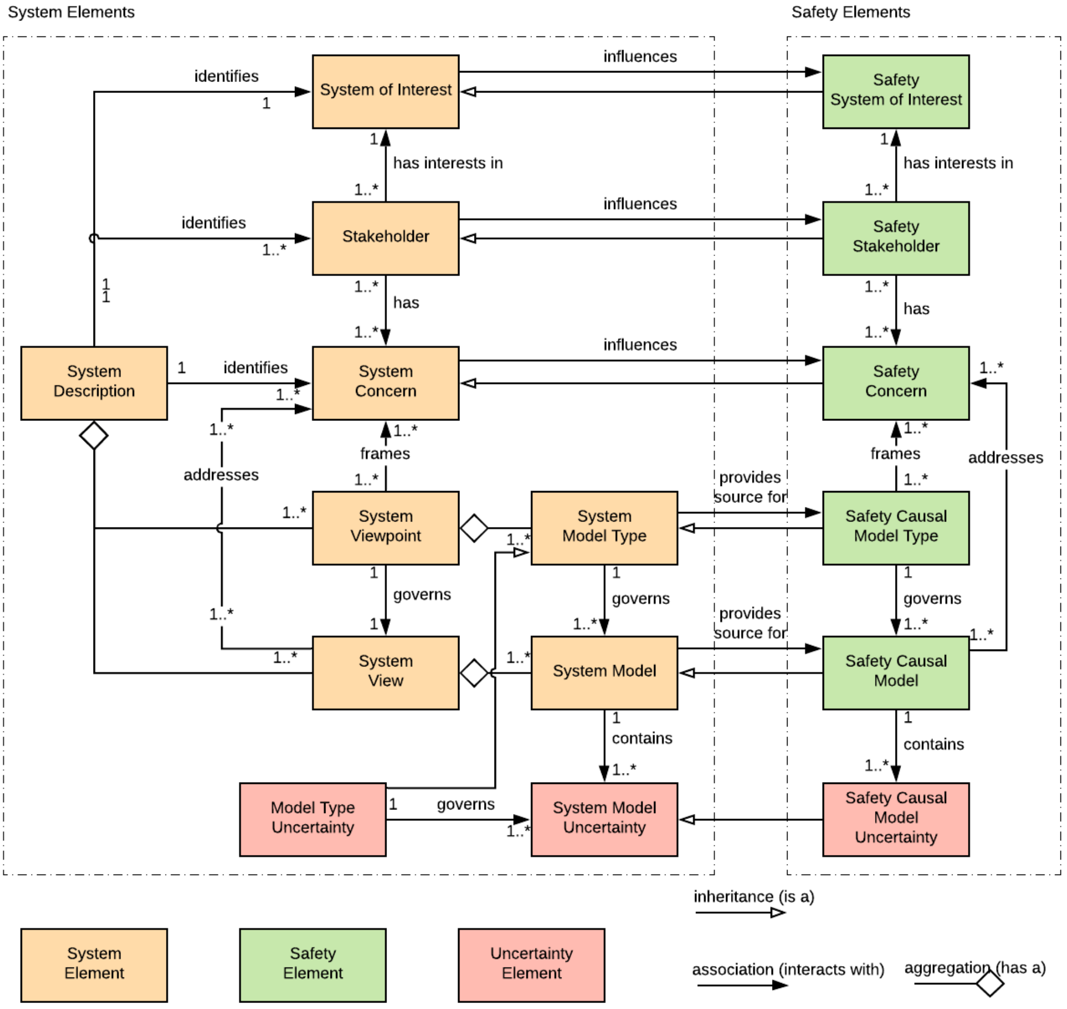


Figure 103 Conceptual Model of System Safety (CMSS)

Conceptual Model of System Safety (CMSS). The CMSS comprises three types of elements: system, safety and uncertainty elements, as shown in Figure 103. The CMSS represents important relationships between system and safety-related elements. This is adopted from the IEEE 42010 standard that describe an architectural description for system elements. The original standard has been modified to better represent the influence of safety elements and the associated epistemic uncertainties.

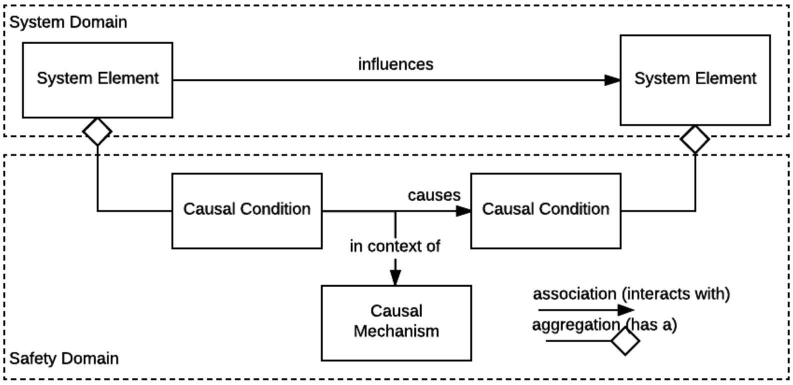


Figure 104 Causal Relationship Model of Safety Assessment (CRMSA)

Causal Relationship Model of Safety Assessment (CRMSA). The CRMSA can be used to represent the causal relationships during safety assessment when identifying epistemic uncertainties (see Figure 104).

Safety assessment is carried out with the aim of identifying causal relationships that could be hazardous. The CRMSA shows the causal conditions in a causal relationship when identifying hazards. Examples of causal conditions in the safety domain include cause, effect, hazard and accident. A system element can have one or more causal conditions. A causal condition can exhibit causal relationship with potentially many other conditions. The causal mechanism provides the narrative or context of how causal conditions can influence each other. An example of a CRMSA is shown in Figure 105.

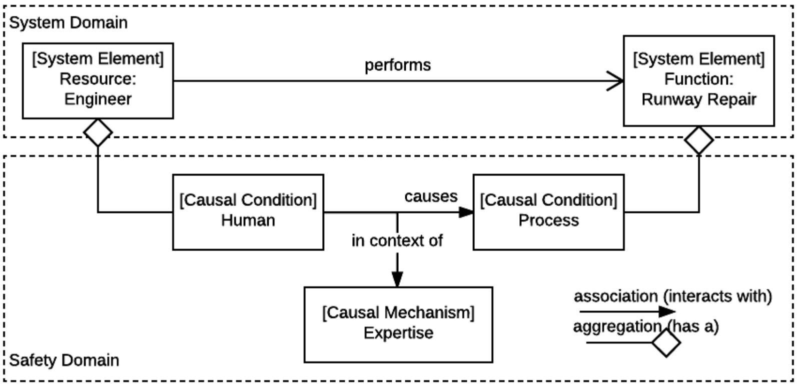

System Domain
[System Element] Resource: Engineer
performs
[System Element] Function: Runway Repair
[Causal Condition] Human
causes
[Causal Condition] Process
in context of
[Causal Mechanism] Expertise
association (interacts with)
aggregation (has a)
Safety Domain

Figure 105 Illustration of CRMSA

<u>Literature Referenced in the HOT-PIE Taxonomy (same as Annex A)</u>

<u>Sampled Track and Address Considerations (same as Annex B)</u>

Glossary

Term	**Definition**
Aleatory Uncertainty	Uncertainty that is due to the random nature of the subject of interest [10]
Assumption	A clearly stated but insufficiently supported proposition that is expected to be true [133]
Causal Condition	An abstraction for capturing some ‘state of affairs’ – event or state, in the system or its environment. Example of safety conditions include hazards, failure modes and faults [21]
Causal Mechanism	A phenomenon that provides the mechanism for a causal relationship, that may either be known or unknown [142, 143]
Causal Relationship	An associated pair of a cause condition that leads to an effect condition [21]
Confidence	Degree of trust one can have in the truth of a position [8]
Epistemic Uncertainty	Uncertainty that is due to a lack of knowledge [10]
Harm	Death, physical injury or damage to the health of people, or damage to property or the environment [63, 64]
Hazard	A situation with the potential to cause harm [63, 64]
Incompleteness	The state of not having all the necessary parts [190]
Information	Systematically organised data [26]
Knowledge	Actionable information being used in the context of a system being investigated [26]
Model	Abstraction of the real world for the purpose of analysis (section 3.2.1.1)

Risk	Combination of the likelihood of harm and the severity of that harm [63, 64]
Risk Assessment Uncertainty (ra-uncertainty)	Uncertainty associated with risk assessment (section 1.2.3)
Risk Assessment Models Uncertainty (ram-uncertainty)	A type of ra-uncertainty due to the lack of knowledge associated with the models that are used to predict safety risk (section 1.2.3)
Safe	Freedom from unacceptable or intolerable levels of harm [63, 64]
Safety Assessment	The processes and techniques to conduct safety analysis (section 2.2)
Safety Assurance	Degree of confidence over the results obtained from the safety assessment (section 2.2)
Safety-Critical System	A system whose failure or malfunction may result in one (or more) outcomes that cause harm [76]
System View	A representation (or model) of one or more aspects of a system that illustrates how the system addresses one or more concerns held by one or more of its stakeholders [20]
System Viewpoint	A collection of patterns, templates, and conventions for constructing one type of view [20]
Taxonomy	A system for naming and organizing things into groups that share similar qualities [191]
Uncertainty	A situation where something is not known [10]

References

[1] D. Vaughan, *The Challenger launch decision: Risky technology, culture, and deviance at NASA*. University of Chicago Press, 1997.

[2] R. P. Feynman, "Rogers Commission Report, Volume 2 Appendix F-Personal Observations on Reliability of Shuttle," ed: NASA, 1986.

[3] L. W. D. Cullen, "The public inquiry into the Piper Alpha disaster," *Drilling Contractor;(United States)*, vol. 49, no. 4, 1993.

[4] UK_Ministry_of_Defence. "CADMID acquisition lifecycle." https://www.aof.mod.uk/aofcontent/general/lifecycles/sg_cadmid.htm?zoom_highlight= (accessed May 29, 2017).

[5] DoD, "MIL-STD-882E System Safety," *DoD System Safety*, May 2012.

[6] F. Redmill, "ALARP Explored," 2010.

[7] *ARP 4761: Guidelines and Methods for Conducting the Safety Assessment Process on Civil Airborne Systems and Equipment*, SAE, 1996.

[8] G. F. Jelen and J. R. Williams, "A practical approach to measuring assurance," in *Computer Security Applications Conference, 1998. Proceedings. 14th Annual*, 1998: IEEE, pp. 333-343.

[9] F. A. Administration. "Safety management system components." (accessed May 2015).

[10] T. Aven and E. Zio, "Some considerations on the treatment of uncertainties in risk assessment for practical decision making," *Reliability Engineering & System Safety*, vol. 96, no. 1, pp. 64-74, 2011, doi: 10.1016/j.ress.2010.06.001.

[11] A. Kerwin, "None too solid medical ignorance," *Science Communication*, vol. 15, no. 2, pp. 166-185, 1993, doi: 10.1177/107554709301500204.

[12] M. B. Spangler, "The Role of Interdisciplinary Analysis in Bridging the Gap Between the Technical and Human Sides of Risk Assessment 1," *Risk Analysis*, vol. 2, no. 2, pp. 101-114, 1982.

[13] F. H. Knight, *Risk, uncertainty and profit*. Courier Corporation, 2012.

[14] D. Hencke. "Chinook blunders cost MoD £500m." @guardian. http://www.theguardian.com/uk/2008/jun/04/military.defence (accessed 17 June, 2019).

[15] C. R. Kothari, *Research methodology: Methods and techniques*. New Age International, 2004.

[16] G. E. Aviña *et al.*, "The Art of Research: A Divergent/Convergent Thinking Framework and Opportunities for Science-Based Approaches," in *Engineering a Better Future*: Springer, 2018, pp. 167-186.

[17] A. Rozan, M. Zaidi, and M. Yoshiki, "The presence of beneficial knowledge in web forum: Analysis by Kipling's framework," in *Knowledge Management*

International Conference & Exhibition (KMICE 2006), Kuala Lumpur, Malaysia, 2006, pp. 153-160.

[18] K. E. Weick, K. M. Sutcliffe, and D. Obstfeld, "Organizing and the process of sensemaking," *Organization science,* vol. 16, no. 4, pp. 409-421, 2005.

[19] D. Frank, "Uncertainty Drives the need to Sense and Respond," vol. 2017, ed, 2017.

[20] I. ISO, "IEEE: 42010: 2011 systems and software engineering, architecture description," *International Standard,* 2011.

[21] S. P. Wilson and J. A. McDermid, "Integrated analysis of complex safety critical systems," *The Computer Journal,* vol. 38, no. 10, pp. 765-776, 1995.

[22] C. James, "Foundations of social theory," *Cambridge, MA: Belknap,* 1990.

[23] V. R. B. G. Caldiera and H. D. Rombach, "The goal question metric approach," *Encyclopedia of software engineering,* pp. 528-532, 1994.

[24] T. Song, D. Zhong, and H. Zhong, "A STAMP analysis on the China-Yongwen railway accident," in *Computer Safety, Reliability, and Security*: Springer, 2012, pp. 376-387.

[25] W. E. Walker *et al.*, "Defining uncertainty: a conceptual basis for uncertainty management in model-based decision support," *Integrated assessment,* vol. 4, no. 1, pp. 5-17, 2003.

[26] A. Jashapara, *Knowledge management: An integrated approach*. Pearson Education, 2004.

[27] H. Dogan, M. J. d. C. Henshaw, and G. Ragsdell, "The risk of information management without knowledge management: a case study," *Journal of Information & Knowledge Management,* vol. 10, no. 04, pp. 393-408, 2011.

[28] C. J. Roy and W. L. Oberkampf, "A comprehensive framework for verification, validation, and uncertainty quantification in scientific computing," *Computer methods in applied mechanics and engineering,* vol. 200, no. 25-28, pp. 2131-2144, 2011.

[29] A. Der Kiureghian and O. Ditlevsen, "Aleatory or epistemic? Does it matter?," *Structural Safety,* vol. 31, no. 2, pp. 105-112, 2009.

[30] D. J. Skinner, S. A. Rocks, S. J. Pollard, and G. H. Drew, "Identifying uncertainty in environmental risk assessments: The development of a novel typology and its implications for risk characterization," *Human and Ecological Risk Assessment: An International Journal,* vol. 20, no. 3, pp. 607-640, 2014.

[31] D. C. Logan, "Known knowns, known unknowns, unknown unknowns and the propagation of scientific enquiry," *Journal of experimental botany,* vol. 60, no. 3, pp. 712-714, 2009.

[32] C. Daase and O. Kessler, "Knowns and unknowns in thewar on terror': uncertainty and the political construction of danger," *Security Dialogue,* vol. 38, no. 4, pp. 411-434, 2007.

[33] C. C. Chow and R. K. Sarin, "Known, unknown, and unknowable uncertainties," *Theory and Decision,* vol. 52, no. 2, pp. 127-138, 2002.

[34] D. Rumsfeld, *Known and unknown: a memoir*. Penguin, 2011.

[35] S. Žižek, "Philosophy, the “unknown knowns,” and the public use of reason," *Topoi,* vol. 25, no. 1-2, pp. 137-142, 2006.

[36] J. Luft and H. Ingham, "The johari window," *Human Relations Training News,* vol. 5, no. 1, pp. 6-7, 1961.

[37] D. H. McQuiston, "Novelty, complexity, and importance as causal determinants of industrial buyer behavior," *The Journal of Marketing,* pp. 66-79, 1989.

[38] W. J. Johnston and T. V. Bonoma, "The buying center: structure and interaction patterns," *The Journal of Marketing,* pp. 143-156, 1981.

[39] D. X. Peng, G. R. Heim, and D. N. Mallick, "Collaborative product development: The effect of project complexity on the use of information technology tools and new product development practices," *Production and Operations Management,* vol. 23, no. 8, pp. 1421-1438, 2014.

[40] S. Novak and S. D. Eppinger, "Sourcing by design: Product complexity and the supply chain," *Management science,* vol. 47, no. 1, pp. 189-204, 2001.

[41] M. Bensaou and N. Venkatraman, "Configurations of interorganizational relationships: A comparison between US and Japanese automakers," *Management science,* vol. 41, no. 9, pp. 1471-1492, 1995.

[42] J. R. Galbraith, "Organization design: An information processing view," *Interfaces,* vol. 4, no. 3, pp. 28-36, 1974.

[43] M. V. Tatikonda and S. R. Rosenthal, "Technology novelty, project complexity, and product development project execution success: a deeper look at task uncertainty in product innovation," *IEEE Transactions on engineering management,* vol. 47, no. 1, pp. 74-87, 2000.

[44] A. J. Shenhar and D. Dvir, *Reinventing project management: the diamond approach to successful growth and innovation*. Harvard Business Review Press, 2007.

[45] P. Svejvig and P. Andersen, "Rethinking project management: A structured literature review with a critical look at the brave new world," *International Journal of Project Management,* vol. 33, no. 2, pp. 278-290, 2015.

[46] F. C. Saunders, A. W. Gale, and A. H. Sherry, "Conceptualising uncertainty in safety-critical projects: A practitioner perspective," *International Journal of Project Management,* vol. 33, no. 2, pp. 467-478, 2015.

[47] D. J. Snowden and M. E. Boone, "A leader's framework for decision making," *Harvard business review,* vol. 85, no. 11, p. 68, 2007.

[48] K. M. Eisenhardt and J. A. Martin, "Dynamic capabilities: what are they?," *Strategic management journal,* pp. 1105-1121, 2000.

[49] L. Argote, *Organizational learning: Creating, retaining and transferring knowledge*. Springer Science & Business Media, 2012.

[50] Y. Petit and B. Hobbs, "Project portfolios in dynamic environments: sources of uncertainty and sensing mechanisms," *Project Management Journal,* vol. 41, no. 4, pp. 46-58, 2010.

[51] F. C. Saunders, A. H. Sherry, and A. W. Gale, "Dualities and dilemmas: contending with uncertainty in large-scale safety-critical projects," *Construction Management and Economics,* vol. 34, no. 9, pp. 657-675, 2016.

[52] M. Famelis, R. Salay, and M. Chechik, "Partial models: Towards modeling and reasoning with uncertainty," in *2012 34th International Conference on Software Engineering (ICSE)*, 2012: IEEE, pp. 573-583.

[53] A. Ferrari, M. Fusani, and S. Gnesi, "Are Standards an Ambiguity-Free Reference for Product Validation?," in *International Conference on Reliability, Safety and Security of Railway Systems*, 2017: Springer, pp. 251-264.

[54] G. Papavassiliou and G. Mentzas, "Knowledge modelling in weakly-structured business processes," *Journal of Knowledge Management,* vol. 7, no. 2, pp. 18-33, 2003.

[55] J. Wang, "A subjective modelling tool applied to formal ship safety assessment," *Ocean Engineering,* vol. 27, no. 10, pp. 1019-1035, 2000.

[56] L. Özdamar and E. Alanya, "Uncertainty modelling in software development projects (with case study)," *Annals of Operations Research,* vol. 102, no. 1-4, pp. 157-178, 2001.

[57] R. W. Mensing, "An Analytic Approach for Treating Uncertainty in Probabilistic Risk Assessments," vol. 2017, ed. Th*irty-First D*o*D Explosives Safety Seminar Proceedings*, 2004.

[58] J. Hammonds, F. Hoffman, and S. Bartell, "An introductory guide to uncertainty analysis in environmental and health risk assessment," *US DOE,* 1994.

[59] M. Goldstein, "Subjective Bayesian analysis: principles and practice," *Bayesian analysis,* vol. 1, no. 3, pp. 403-420, 2006.

[60] A. Kinder, M. Henshaw, and C. Siemieniuch, "A model based approach to system of systems risk management," in *System of Systems Engineering Conference (SoSE), 2015 10th*, 2015: IEEE, pp. 122-127.

[61] R. Flage and T. Aven, "Expressing and communicating uncertainty in relation to quantitative risk analysis," *Reliability: Theory & Applications,* vol. 4, no. 2-1 (13), 2009.

[62] (May 2012). *DoD MIL-STD-882E on System Safety*.

[63] (2015). *Ministry of Defence (2015) Defence Standard 00-56 Issue 6: Safety Management Requirements for Defence Systems*.

[64] (2007). *Ministry of Defence (2007) Defence Standard 00-56 Issue 4: Safety Management Requirements for Defence Systems*.

[65] C. A. Ericson, *Hazard analysis techniques for system safety*. John Wiley & Sons, 2015.

[66] N. Leveson, *Engineering a safer world: Systems thinking applied to safety*. MIT press, 2011.

[67] J. Chen, S. Zhang, Y. Lu, and P. Tang, "STPA-based hazard analysis of a complex UAV system in take-off," in *Transportation Information and Safety (ICTIS), 2015 International Conference on*, 2015: IEEE, pp. 774-779.

[68] J. McDermid, "Risk, Uncertainty, Software and Professional Ethics," *Safety Systems: The Safety-Critical Systems Club Newsletter,* 2008.

[69] J. M. Keynes, *A treatise on probability*. Courier Corporation, 2013.

[70] L. Duan, S. Rayadurgam, M. P. Heimdahl, A. Ayoub, O. Sokolsky, and I. Lee, "Reasoning about confidence and uncertainty in assurance cases: A survey," in *Software Engineering in Health Care*: Springer, 2014, pp. 64-80.

[71] T. P. Kelly, "Arguing safety: a systematic approach to managing safety cases," University of York, 1999.

[72] R. Hawkins, T. Kelly, J. Knight, and P. Graydon, "A new approach to creating clear safety arguments," in *Advances in systems safety*: Springer, 2011, pp. 3-23.

[73] C. B. Weinstock, J. B. Goodenough, and A. Z. Klein, "Measuring assurance case confidence using Baconian probabilities," in *Proceedings of the 1st international workshop on assurance cases for software-intensive systems,* 2013: IEEE Press, pp. 7-11.

[74] L. Cyra and J. Górski, "Supporting expert assessment of argument structures in trust cases," in *Ninth International Probabilistic Safety Assessment and Management Conference PSAM*, 2008, vol. 9, pp. 1-9.

[75] T. K. Ferrell and U. D. Ferrell, "RTCA DO-178C/EUROCAE ED-12C," *Digital Avionics Handbook,* 2017.

[76] J. C. Knight, "Safety critical systems: challenges and directions," in *Software Engineering, 2002. ICSE 2002. Proceedings of the 24rd International Conference on*, 2002: IEEE, pp. 547-550.

[77] R. Kazman, K. Schmid, C. B. Nielsen, and J. Klein, "Understanding patterns for system of systems integration," in *System of Systems Engineering (SoSE), 2013 8th International Conference on,* 2-6 June 2013 2013, pp. 141-146, doi: 10.1109/SYSoSE.2013.6575257.

[78] M. Trapp and D. Schneider, "Safety Assurance of Open Adaptive Systems–A Survey," in *Models@ run. time*: Springer, 2014, pp. 279-318.

[79] I. Habli, "Model-based assurance of safety-critical product lines," University of York, 2009.

[80] M. R. Endsley, "A taxonomy of situation awareness errors," *Human factors in aviation operations,* vol. 3, no. 2, pp. 287-292, 1995.

[81] N. Nwiabu, I. Allison, P. Holt, P. Lowit, and B. Oyeneyin, "Situation awareness in context-aware case-based decision support," in *Cognitive Methods in Situation Awareness and Decision Support (CogSIMA), 2011 IEEE First International Multi-Disciplinary Conference on,* 22-24 Feb. 2011 2011, pp. 9-16, doi: 10.1109/COGSIMA.2011.5753761.

[82] N. Kajtazovic, C. Preschern, A. Höller, and C. Kreiner, "Towards Assured Dynamic Configuration of Safety-Critical Embedded Systems," in *Computer Safety, Reliability, and Security*: Springer, 2014, pp. 167-179.

[83] RTCA, *Integrated Modular Avionics (IMA) Development Guidance and Certification Considerations*. RTCA, 2005.

[84] T. Reiman, C. Rollenhagen, E. Pietikäinen, and J. Heikkilä, "Principles of adaptive management in complex safety–critical organizations," *Safety Science*, vol. 71, pp. 80-92, 2015.

[85] S. Dekker, *Drift into failure: from hunting broken components to understanding complex systems*. Ashgate Publishing, Ltd., 2012.

[86] A. Ghorbani, G. P. Dijkema, and I. Nikolic, "Emergence Engineering: A Review," *Available at SSRN 2138253*, 2012.

[87] M. Ulieru and R. Doursat, "Emergent engineering: a radical paradigm shift," *International Journal of Autonomous and Adaptive Communications Systems*, vol. 4, no. 1, pp. 39-60, 2011.

[88] P. J. Redmond, J. B. Michael, and P. V. Shebalin, "Interface hazard analysis for system of systems," in *System of Systems Engineering, 2008. SoSE'08. IEEE International Conference on*, 2008: IEEE, pp. 1-8.

[89] NATS. "ATC Disruption 7 Dec 13 - Report.pdf." http://www.nats.aero/wp-content/uploads/2014/08/ATC%20Disruption%207%20Dec%2013%20-%20Report.pdf (accessed 10 April, 2015).

[90] I. Osborne. "CAA Standard Letter - Air Traffic Control Disruption on 7th December 2013." http://www.caa.co.uk/docs/5/20140812OsborneRolfeVCSIncident&FollowUp.pdf (accessed 11 April, 2015).

[91] J. Stone. "An unprecedented computer fault cleared London's skies of planes (and the people in charge had never heard of it before)." @Independent. http://www.independent.co.uk/news/uk/home-news/londonwide-air-traffic-control-failure-was-caused-by-an-unprecedented-computer-fault-9922914.html (accessed 3 April, 2015).

[92] NATS. "Service outage at Swanwick 12 Dec - NATS." http://www.nats.aero/news/newsbrief/dec14jan-2015/service-outage-swanwick-12-dec/ (accessed 10 April, 2015).

[93] G. Baxter and I. Sommerville, "Socio-technical systems: From design methods to systems engineering," *Interacting with Computers*, vol. 23, no. 1, pp. 4-17, 2011.

[94] I. Sommerville *et al.*, "Large-scale complex IT systems," *Communications of the ACM*, vol. 55, no. 7, pp. 71-77, 2012.

[95] P. Feiler *et al.*, "Ultra-large-scale systems: The software challenge of the future," *Software Engineering Institute*, vol. 1, 2006.

[96] R. Atkinson, L. Crawford, and S. Ward, "Fundamental uncertainties in projects and the scope of project management," *International journal of project management*, vol. 24, no. 8, pp. 687-698, 2006.

[97] C. W. Johnson and C. M. Holloway, "A Possible Approach for Addressing Neglected Human Factors Issues of Systems Engineering," 2011.

[98] G. Wild, "Risk homeostasis theory and traffic accidents," *Ergonomics*, vol. 31, no. 4, pp. 441-468, 1988.

[99] J. Shi, J. Wan, H. Yan, and H. Suo, "A survey of cyber-physical systems," in *Wireless Communications and Signal Processing (WCSP), 2011 International Conference on*, 2011: IEEE, pp. 1-6.

[100] K. Sampigethaya and R. Poovendran, "Aviation cyber–physical systems: foundations for future aircraft and air transport," *Proceedings of the IEEE,* vol. 101, no. 8, pp. 1834-1855, 2013.

[101] A. Banerjee, K. K. Venkatasubramanian, T. Mukherjee, and S. K. Gupta, "Ensuring safety, security, and sustainability of mission-critical cyber–physical systems," *Proceedings of the IEEE,* vol. 100, no. 1, pp. 283-299, 2012.

[102] L. Pazzi, "Controlling hazards and safety in complex systems: a multi-layered part-whole approach to system safety i," *Business Systems Review,* vol. 1, no. 1, pp. 248-262, 2012.

[103] J. Zhu and A. Mostafavi, "Towards a new paradigm for management of complex engineering projects: A system-of-systems framework," in *Systems Conference (SysCon), 2014 8th Annual IEEE*, 2014: IEEE, pp. 213-219.

[104] M. W. Maier, "Architecting Principles for Systems‐of‐Systems," in *INCOSE International Symposium*, 1996, vol. 6, no. 1: Wiley Online Library, pp. 565-573.

[105] C. Harvey and N. A. Stanton, "Safety in System-of-Systems: ten key challenges," *Safety science,* vol. 70, pp. 358-366, 2014.

[106] Y. Y. Haimes, "Modeling complex systems of systems with phantom system models," *Systems Engineering,* vol. 15, no. 3, pp. 333-346, 2012.

[107] M. Hause, "Model-Based System of Systems Engineering with UPDM," *Omg. org,* 2010.

[108] S. Benade and L. Pretorius, "System architecture and enterprise architecture: a juxta position?," *South African Journal of Industrial Engineering,* vol. 23, no. 2, pp. 29-46, 2012.

[109] J. E. Bartolomei, D. E. Hastings, R. de Neufville, and D. H. Rhodes, "Engineering Systems Multiple‐Domain Matrix: An organizing framework for modeling large‐scale complex systems," *Systems Engineering,* vol. 15, no. 1, pp. 41-61, 2012.

[110] C. Wrigley, "Modelling Systems-of-Systems: Issues and possible solutions," in *SoSE*, 2012, pp. 479-484.

[111] A. Ruiz, I. Habli, and H. Espinoza, "Towards a case-based reasoning approach for safety assurance reuse," in *Computer Safety, Reliability, and Security*: Springer, 2012, pp. 22-35.

[112] C. Alexander, S. Ishikawa, and M. Silverstein, *A pattern language: towns, buildings, construction*. Oxford University Press, 1977.

[113] J. Rauhamäki, T. Vepsäläinen, and S. Kuikka, "Functional safety system patterns," in *Proc. VikingPLoP 2012 Conference*, 2012, pp. 48-68.

[114] C. Wei, B. Xiaohong, and L. Xuefei, "A Study on Airborne Software Safety Requirements Patterns," in *Software Security and Reliability-Companion (SERE-C), 2013 IEEE 7th International Conference on*, 2013: IEEE, pp. 131-136.

[115] G. Jolliffe, "Producing a safety case for IMA blueprints," in *Digital Avionics Systems Conference, 2005. DASC 2005. The 24th*, 2005, vol. 2: IEEE, p. 14 pp. Vol. 2.

[116] J. L. Fenn, R. D. Hawkins, P. Williams, T. P. Kelly, M. G. Banner, and Y. Oakshott, "The who, where, how, why and when of modular and incremental certification," in *System Safety, 2007 2nd Institution of Engineering and Technology International Conference on*, 2007: IET, pp. 135-140.

[117] MoD, "The Acquisition Handbook 2005," 2005.

[118] P. D. Antill, J. C. Smith, and D. M. Moore, "UK AFV and PPV procurement using Urgent Operational Requirements (UOR)," 2012.

[119] M. Chris. (2012) LPPV - Lessons for Defence Procurement. *RUSI DEFENCE SYSTEMS*. 3.

[120] M. Bozzano and A. Villafiorita, *Design and safety assessment of critical systems*. CRC Press, 2010.

[121] C. A. P. Ltd. "Functional Safety Management." (accessed May 2015).

[122] G. Barlow and A. Shanks, "Systems and Safety Engineering–a Combined Approach during Concept Design and Beyond," in *Proc. RINA International Conference on Systems Engineering in Ship and Offshore Design, Bath, 21-22 October 2010*, 2010.

[123] R. Baheti and H. Gill, "Cyber-physical systems," *The impact of control technology*, pp. 161-166, 2011.

[124] E. Institute. "Guidance Article: Top 10 Health Technology Hazards for 2014: Key Safety Threats to Manage in the Coming Year - STF_Top_Ten_Tech_Hazards_2014-06-13.pdf." http://www.healthit.gov/facas/sites/faca/files/STF_Top_Ten_Tech_Hazards_2014-06-13.pdf (accessed 9 April, 2015).

[125] C. Johnson, "Identifying common problems in the acquisition and deployment of large-scale, safety–critical, software projects in the US and UK healthcare systems," *Safety Science*, vol. 49, no. 5, pp. 735-745, 2011.

[126] L. A. Lipsitz, "Understanding health care as a complex system: the foundation for unintended consequences," *JAMA*, vol. 308, no. 3, pp. 243-244, 2012.

[127] G. E. Box, "Science and statistics," *Journal of the American Statistical Association*, vol. 71, no. 356, pp. 791-799, 1976.

[128] E. L. Trist, "On socio-technical systems," *Sociotechnical systems: A sourcebook*, pp. 43-57, 1978.

[129] J. Robinson, *Essays in the theory of economic growth*. Springer, 1965.

[130] M. Kolmar, *Principles of Microeconomics*. Springer, 2017.

[131] B. ISO, "IEC 15026-2: 2011 Systems and software engineering–Systems and software assurance," ed: Part.

[132] O. SACM, "Structured Assurance Case Meta-model," 2013.

[133] T. Kelly and R. Weaver, "The goal structuring notation–a safety argument notation," in *Proceedings of the dependable systems and networks 2004 workshop on assurance cases*, 2004: Citeseer, p. 6.

[134] NATO, "Methods and Models for Life Cycle Costing," *RTO-TR-SAS-054,* 2007.

[135] E. L. S. Yen and O. X. Hui, "SMART DIAGNOSTIC TOOL FOR COMPLEX SYSTEM-OF-SYSTEMS," *DSTA Horizons 2018,* p. 8, 2018.

[136] V. Hamilton, "Criteria for Software Evidence, Goal-based standards require evidence-based approaches," *Safety Systems,* vol. 16, no. 1, 2006.

[137] E. T. Almeida, J. E. Luntz, and D. M. Tilbury, "Modular finite state machines implemented as event-condition-action systems," *IFAC Proceedings Volumes,* vol. 38, no. 1, pp. 373-378, 2005.

[138] J. Bae, H. Bae, S.-H. Kang, and Y. Kim, "Automatic control of workflow processes using ECA rules," *IEEE transactions on knowledge and data engineering,* vol. 16, no. 8, pp. 1010-1023, 2004.

[139] N. Chaudhry, J. Moyne, and E. A. Rundensteiner, "Active Controller: utilizing active databases for implementing multistep control of semiconductor manufacturing," *IEEE Transactions on Components, Packaging, and Manufacturing Technology: Part C,* vol. 21, no. 3, pp. 217-224, 1998.

[140] G. Papamarkos, A. Poulovassilis, and P. T. Wood, "Event-condition-action rule languages for the semantic web," in *Proceedings of the First International Conference on Semantic Web and Databases*, 2003: Citeseer, pp. 294-312.

[141] S. ARP, "Certification considerations for highly-integrated or complex aircraft systems," *SAE, Warrendale, PA,* 1996.

[142] P. Hedström and P. Ylikoski, "Causal mechanisms in the social sciences," *Annual Review of Sociology,* vol. 36, pp. 49-67, 2010, doi: 10.1146/annurev.soc.012809.102632.

[143] C. F. Craver and L. Darden, *In search of mechanisms: Discoveries across the life sciences*. University of Chicago Press, 2013.

[144] R. Ferdous, F. Khan, R. Sadiq, P. Amyotte, and B. Veitch, "Fault and event tree analyses for process systems risk analysis: uncertainty handling formulations," *Risk analysis,* vol. 31, no. 1, pp. 86-107, 2011.

[145] J. H. Saleh, K. B. Marais, and F. M. Favaró, "System safety principles: A multidisciplinary engineering perspective," *Journal of Loss Prevention in the Process Industries,* vol. 29, pp. 283-294, 2014, doi: 10.1016/j.jlp.2014.04.001.

[146] S. A. Shappel and D. A. Wiegmann, "The human factors analysis and classification system--HFACS," US Federal Aviation Administration, Office of Aviation Medicine, 2000.

[147] B. M. O'Halloran, R. B. Stone, and I. Y. Tumer, "A failure modes and mechanisms naming taxonomy," in *Reliability and Maintainability Symposium (RAMS), 2012 Proceedings-Annual*, 2012: IEEE, pp. 1-6, doi: 10.1109/rams.2012.6175455.

[148] Y. Yue and M. Henshaw, "An holistic view of UK military capability development," *Defense & Security Analysis,* vol. 25, no. 1, pp. 53-67, 2009.

[149] F. R. Burton, R. F. Paige, S. Poulding, and S. Smith, "System of systems acquisition trade-offs," *Procedia Computer Science,* vol. 28, pp. 11-18, 2014.

[150] D. J. Hurley, "Defence capability development manual,," *Australian Department of Defence*, 2006.

[151] S. A. Fry, "Joint publication 1-02 - Department of Defence dictionary of military and associated terms," *Department of Defence*, July 2010.

[152] M. Lizotte *et al.*, "Toward a Capability Engineering Process," DEFENCE RESEARCH AND DEVELOPMENT CANADAVALCARTIER (QUEBEC), 2004.

[153] G. Morgan, "G. Images of Organisation," ed: Sage: California, 1986.

[154] C. Johnson, "Failure in safety-critical systems," *A Handbook of Accident and Incident Reporting. Ocotber. University of Glasgow Press, Glasgow, Scotland,* 2003.

[155] W. Li, L. Zhang, and W. Liang, "An Accident Causation Analysis and Taxonomy (ACAT) model of complex industrial system from both system safety and control theory perspectives," *Safety science,* vol. 92, pp. 94-103, 2017, doi: 10.1016/j.ssci.2016.10.001.

[156] Y. K. Dwivedi *et al.*, "IS/IT project failures: a review of the extant literature for deriving a taxonomy of failure factors," in *International Working Conference on Transfer and Diffusion of IT*, 2013: Springer, pp. 73-88, doi: 10.1007/978-3-642-38862-0_5.

[157] E. Hollnagel, *FRAM, the functional resonance analysis method: modelling complex socio-technical systems*. Ashgate Publishing, Ltd., 2012.

[158] H. R. Booher, *Handbook of human systems integration*. John Wiley & Sons, 2003.

[159] M. Van Beuzekom, F. Boer, S. Akerboom, and P. Hudson, "Patient safety: latent risk factors," *British journal of anaesthesia*, vol. 105, no. 1, pp. 52-59, 2010, doi: 10.1093/bja/aeq135.

[160] J. Rasmussen, "Risk management in a dynamic society: a modelling problem," *Safety science*, vol. 27, no. 2, pp. 183-213, 1997, doi: 10.1016/s0925-7535(97)00052-0.

[161] P. Baybutt, "Addressing enablers in layers of protection analysis," *Process Safety Progress,* vol. 33, no. 3, pp. 221-226, 2014, doi: 10.1002/prs.11639.

[162] T. W. Van der Schaaf, J. Moraal, and A. R. Hale, *Near miss reporting in the chemical process industry*. Technische Universiteit Eindhoven, Proefschrift., 1992.

[163] E. K. Cetinkaya and J. P. Sterbenz, "A taxonomy of network challenges," in *Design of Reliable Communication Networks (DRCN), 2013 9th International Conference on the*, 2013: IEEE, pp. 322-330.

[164] E. Hollnagel, *Cognitive reliability and error analysis method (CREAM).* Elsevier, 1998.

[165] P. Carayon, "The balance theory and the work system model… Twenty years later," *Intl. Journal of Human–Computer Interaction,* vol. 25, no. 5, pp. 313-327, 2009, doi: 10.1080/10447310902864928.

[166] J. de Jonge, H. Bosma, R. Peter, and J. Siegrist, "Job strain, effort-reward imbalance and employee well-being: a large-scale cross-sectional study,"

Social science & medicine, vol. 50, no. 9, pp. 1317-1327, 2000, doi: 10.1016/s0277-9536(99)00388-3.

[167] E. Edwards, "Man and machine- Systems for safety(Man machine systems for flight safety, studying accidents, human factors in system design and implementation of personnel)," *Outlook on safety,* pp. 21-36, 1973.

[168] R. P. Bostrom and J. S. Heinen, "MIS problems and failures: a socio-technical perspective, part II: the application of socio-technical theory," *MIS quarterly,* pp. 11-28, 1977, doi: 10.2307/249019.

[169] J. R. Wilson, "Fundamentals of ergonomics in theory and practice," *Applied ergonomics,* vol. 31, no. 6, pp. 557-567, 2000, doi: 10.1016/s0003-6870(00)00034-x.

[170] S. Alter, "Work system theory: overview of core concepts, extensions, and challenges for the future," *Journal of the Association for Information Systems,* vol. 14, no. 2, p. 72, 2013.

[171] B. M. Kleiner, L. J. Hettinger, D. M. DeJoy, Y.-H. Huang, and P. E. Love, "Sociotechnical attributes of safe and unsafe work systems," *Ergonomics,* vol. 58, no. 4, pp. 635-649, 2015, doi: 10.1080/00140139.2015.1009175.

[172] A. B. Bakker and E. Demerouti, "The job demands-resources model: State of the art," *Journal of managerial psychology,* vol. 22, no. 3, pp. 309-328, 2007, doi: 10.1108/02683940710733115.

[173] C. Dejours and J.-P. Deranty, "The centrality of work," *Critical Horizons,* vol. 11, no. 2, pp. 167-180, 2010, doi: 10.1558/crit.v11i2.167.

[174] J. P. Van Gigch, "Modeling, metamodeling, and taxonomy of system failures," *IEEE transactions on reliability,* vol. 35, no. 2, pp. 131-136, 1986, doi: 10.1109/tr.1986.4335383.

[175] J. Reason, *Human error*. Cambridge university press, 1990.

[176] W. A. Wagenaar, J. Groeneweg, P. Hudson, and J. Reason, "Promoting safety in the oil industry. The ergonomics society lecture presented at the ergonomics society annual conference, Edinburgh, 13-16 April 1993," *Ergonomics,* vol. 37, no. 12, pp. 1999-2013, 1994, doi: 10.1080/00140139408964963.

[177] B. M. Kleiner, "Macroergonomics: analysis and design of work systems," *Applied ergonomics,* vol. 37, no. 1, pp. 81-89, 2006, doi: 10.1016/j.apergo.2005.07.006.

[178] S. Khan, P. Phillips, C. Hockley, and I. K. Jennions, "Towards standardisation of no fault found taxonomy," 2012.

[179] W. G. Johnson, "Management Oversight and Risk Tree-MORT," Aerojet Nuclear Co., Scoville, ID (USA), 1973.

[180] R. Lock, "Modelling and analysing standard use within system of systems," in *Engineering of Complex Computer Systems (ICECCS), 2011 16th IEEE International Conference on,* 2011: IEEE, pp. 149-156.

[181] C. H.-C. QC, "The Nimrod Review," *Air Command, House of Commons, London,* 2009.

[182] N. Leveson, "An STPA Primer, Version 1, August 2013," ed, 2013.

[183] E. ISO, "9001: 2015 Quality management systems," *Requirements (ISO 9001: 2015), European Committee for Standardization, Brussels*, 2015.

[184] A. N. Oppenheim, *Questionnaire design, interviewing and attitude measurement*. Bloomsbury Publishing, 2000.

[185] K. F. Punch, *Introduction to social research: Quantitative and qualitative approaches*. sage, 2013.

[186] V. Clarke and V. Braun, "Thematic analysis," in *Encyclopedia of critical psychology*: Springer, 2014, pp. 1947-1952.

[187] R. Likert, "A technique for the measurement of attitudes," *Archives of psychology*, 1932.

[188] D. Bertram, "Likert scales," *Retrieved November*, vol. 2, p. 2013, 2007.

[189] B. P. Subedi, "Using Likert type data in social science research: Confusion, issues and challenges," *International Journal of Conterporary Applied Sciences*, vol. 3, no. 2, pp. 36-49, 2016.

[190] "INCOMPLETENESS | meaning in the Cambridge English Dictionary." https://dictionary.cambridge.org/dictionary/english/incompleteness (accessed June 17, 2019).

[191] "TAXONOMY | meaning in the Cambridge English Dictionary." https://dictionary.cambridge.org/dictionary/english/taxonomy (accessed June 17, 2019).