



SAFETY-CRITICAL SYSTEMS RECOMMENDED PRACTICE

J3187™

MAY2023

Issued 2022-02
Revised 2023-05

Superseding J3187 FEB2022

(R) System Theoretic Process Analysis (STPA) Recommended Practices
for Evaluations of Safety-Critical Systems in Any Industry

RATIONALE

This document provides recommended practices regarding how System Theoretic Process Analysis (STPA) may be applied to safety-critical systems in any industry.

NOTE: This document is a replacement for the original release of SAE J3187 and has been modified to make it clear that it is applicable for all industries dealing with safety critical systems. This new version is streamlined in nature and is not automotive specific. The original Human-Machine Interactions (HMIs), Model-Based Systems Engineering (MBSE), and Safety of the Intended Functionality (SOTIF) sections have been removed and will be offered as separate independent appendices under the original SAE J3187 number, utilizing different suffix labeling.

TABLE OF CONTENTS

1.	SCOPE.....	4
1.1	Purpose.....	4
2.	REFERENCES.....	4
2.1	Applicable Documents	4
2.1.1	SAE Publications.....	4
2.1.2	Other Publications.....	4
3.	DEFINITIONS	6
4.	ACRONYMS	8
5.	DOCUMENT ORGANIZATION.....	9
6.	BASIC STPA APPROACH AND LESSONS LEARNED.....	9
6.1	Introduction	9
6.2	STPA Method Overview.....	10
6.3	Use of Systems Engineering Perspective.....	11
6.4	STPA Scope Determination	14
6.5	Potential Loss ("Accident") Identification.....	15
6.6	Potential "Hazard" Identification.....	17
6.7	Defining System-Level Constraints (Safety Goals).....	18
6.8	Creating a Control Structure	19
6.9	Define Unsafe/Unwanted/Unexpected Control Actions (UCAs)	21

SAE Executive Standards Committee Rules provide that: "This report is published by SAE to advance the state of technical and engineering sciences. The use of this report is entirely voluntary, and its applicability and suitability for any particular use, including any patent infringement arising therefrom, is the sole responsibility of the user."

SAE reviews each technical report at least every five years at which time it may be revised, reaffirmed, stabilized, or cancelled. SAE invites your written comments and suggestions.

Copyright © 2023 SAE International

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording, or otherwise, without the prior written permission of SAE.

TO PLACE A DOCUMENT ORDER: Tel: 877-606-7323 (inside USA and Canada)
Tel: +1 724-776-4970 (outside USA)
Fax: 724-776-0790
Email: CustomerService@sae.org
http://www.sae.org

SAE WEB ADDRESS:

Full document details: SAE J3187_202305

For more information on this standard, visit
https://www.sae.org/standards/content/J3187_202305/

Documents

Document Sharing Hub

6.10	Define Casual Scenarios.....	24
6.11	Creation of Safety Requirements.....	28
6.12	Lessons Learned.....	29
6.13	Questions to Prepare for STPA.....	30
6.14	Questions While Performing STPA.....	30
6.15	Lessons Learned for Incorporating STPA in Large Organizations.....	31
6.16	Japan Automotive Software Platform and Architecture (JASPAR) Lessons Learned.....	32
6.16.1	Control Structure Legend.....	33
6.16.2	Vehicle Level Control Structure.....	33
6.16.3	System-Level Control Structure.....	34
7.	HIGH-LEVEL USE OF STPA WITHIN SAFETY PROCESSES AND STPA WITH OTHER SAFETY EVALUATION METHODS.....	35
7.1	Introduction.....	35
7.1.1	STPA Assistance in Established System Safety Processes and/or Standards.....	37
7.1.2	STPA Within a System Safety Process.....	45
7.2	Analysis Methodology Perspectives - Inductive, Deductive, and Exploratory.....	46
7.3	Comments on HAZOP and STPA.....	47
7.4	Interactions versus Interfaces.....	47
7.5	STPA in Human-Machine Interaction (HMI) Evaluations.....	48
7.6	STPA Compared to Other System Safety Analysis Methods.....	48
7.6.1	Useful Comparison Studies and Presentations.....	48
7.6.2	Illustrative Analogy for STPA and Other Methodologies.....	49
7.6.3	High-Level Methodology Comparison.....	50
8.	SUMMARY.....	52
8.1	STPA Value.....	52
8.2	STPA Application Perspective.....	52
8.3	STPA Effectiveness for Evaluating Human-Machine Interactions within Operating Scenarios.....	52
8.4	STPA and Emergent Properties and Behaviors.....	53
8.5	STPA within Existing Safety Evaluation Processes.....	53
9.	NOTES.....	54
9.1	Revision Indicator.....	54
Figure 1	Overview of the basic STPA method (Leveson and Thomas, 2018).....	10
Figure 2	Initial allocation to major systems (Vernacchia, 2018).....	12
Figure 3	Interactions between major systems (Vernacchia, 2018).....	13
Figure 4	Rebalance and optimize initial concept (Vernacchia, 2018).....	13
Figure 5	Selected concept/architecture (Vernacchia, 2018).....	13
Figure 6	Multiple level balancing and optimization for continued concept/architecture improvement (Vernacchia, 2018).....	14
Figure 7	STPA Handbook system-level constraints (safety goals) (Leveson and Thomas, 2018).....	19
Figure 8	High-level control structure.....	19
Figure 9	Abstraction progression example.....	20
Figure 10	Iterative cycle using STPA for system control structure updates.....	21
Figure 11	Guideword headings as shown in STPA Handbook (Leveson and Thomas, 2018).....	22
Figure 12	Guideword table with additional helpful phrases (Ressler and Vernacchia, 2019).....	22
Figure 13	Example of appropriate UCA format.....	23
Figure 14	Causal scenario graphic from STPA Handbook (Leveson and Thomas, 2018).....	25
Figure 15	Compressed natural gas system: causal scenario “type” areas (Ressler and Vernacchia, 2018).....	26
Figure 16	JASPAR lessons learned vehicle level control structure.....	34
Figure 17	JASPAR lessons learned system-level control structure.....	35
Figure 18	Systems engineering “V” model and STPA (Leveson and Thomas, 2018).....	36
Figure 19	System engineering “V” model with STPA in a safety process (Vernacchia, 2018).....	45
Figure 20	Different evaluation methodologies (Ressler and Vernacchia, 2019).....	46
Figure 21	Causes and effects matrix (Ressler and Vernacchia, 2019).....	47

Figure 22	Interactions of engine thrust (Vernacchia, 2016)	48
Figure 23	Hidden cube face	50
Table 1	System concept/architecture tasks involving early safety analysis methods (Vernacchia, 2018)	12
Table 2	JASPAR lessons learned control structure legend	33
Table 3	Standards comparison for requirements definition	38
Table 4	STPA high-level main steps	41
Table 5	STPA with ISO 26262, MIL-STD-882E, and ARP4761	42
Table 6	Attribute comparison of different analyses	51

1. SCOPE

This document provides recommended practices regarding how System Theoretic Process Analysis (STPA) may be applied to safety-critical systems in any industry.

1.1 Purpose

This document provides information and lessons learned from experienced STPA practitioners about the use of STPA for the evaluation of control-based safety-critical systems. Appendices explore specific areas of STPA application, such as Human-Machine Interaction (HMI) analysis, medical devices, Model-Based Systems Engineering (MBSE), cybersecurity, automotive, etc., and provide examples for specific industries and/or topics. These appendices are published as separate documents.

2. REFERENCES

2.1 Applicable Documents

The following publications form a part of this specification to the extent specified herein. Unless otherwise indicated, the latest issue of SAE publications shall apply.

2.1.1 SAE Publications

Available from SAE International, 400 Commonwealth Drive, Warrendale, PA 15096-0001, Tel: 877-606-7323 (inside USA and Canada) or +1 724-776-4970 (outside USA), www.sae.org.

SAE J3016 Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles

Khastgir, S., Birrell, S., Dhadyalla, G., and Jennings, P., "The Science of Testing: An Automotive Perspective," SAE Technical Paper 2018-01-1070, 2018, <https://doi.org/10.4271/2018-01-1070>.

Okada, M. and Christensen, A. (2017, October 25). SAE Recommended Task Force - Group 4 Example.

Suo, D., Yako, S., Boesch, M., and Post, K., "Integrating STPA into ISO 26262 Process for Requirement Development," SAE Technical Paper 2017-01-0058, 2017, <https://doi.org/10.4271/2017-01-0058>.

2.1.2 Other Publications

Abrecht, B. (2016). A New Approach to Hazard Analysis for Naval Systems. MIT STAMP Workshop. Cambridge MA: MIT.

Crash Avoidance Metrics Partnership (CAMP) Automated Vehicle Research (AVR). (2016). Automated Vehicle Research for Enhanced Safety. Washington DC: U.S. Department of Transportation.

Fleming, C., Placke, M., and Leveson, N. (2013). Technical report: STPA analysis of NextGen interval management components: Ground interval management (GIM) and flight decn interval management (FIM).

France, M.E. (2017). Engineering for Humans - Human-Automation Interaction in STPA. MIT STAMP/STPA Workshop (p. 18/50). Cambridge: MIT.

France, M.E. (2017). Engineering for Humans: A New Extension to STPA. Cambridge MA: MIT - Master's Thesis.

Georgiou, O.B. (2017). Haptic In-Vehicle Gesture Controls. Proceedings of the 9th International Conference on Automotive User Interfaces and Interactive Vehicular Applications Adjunct (pp. 233-238).

Immersion Corporation. (2017). MAGNA and IPG Media Labs Haptic Ad Study. Immersion Corporation; <https://www.immersion.com/study-unveils-touchsense-ads-lead-to-brand-lift/>.

Kalra, N. and Paddock, S.M. (2016). Driving to safety: How many miles of driving would it take to demonstrate autonomous vehicle reliability? Transportation Research Part A: Policy and Practice.

Khastgir, S., Brewerton, S., Thomas, J., and Jennings, P. (2021). Systems Approach to Creating Test Scenarios for Automated Driving Systems. Reliability Engineering and System Safety.

Leveson, N. (2015). A systems approach to risk management through leading safety indicators. Reliability engineering and system safety.

Leveson, N. and Thomas, J. (2018). STPA Handbook. Cambridge MA: MIT.

Malloy, N.H. (2017). Integrating STAMP-Based Hazard Analysis with MIL-STD-882E Functional Hazard Analysis. MIT STAMP Workshop. Cambridge MA: MIT.

Martinez, P.D. (2017). Agency in Mid-air Interfaces. Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (pp. 2426-2439).

Martinez, R.S. (2015). A comparison of STPA and automotive FMECA. MIT STAMP Workshop. Cambridge MA: MIT.

Object Management Group. (2020). RAAML Examples. Risk Analysis and Assessment Modeling Language (RAAML) Examples (Informative). Milford, MA, USA: Object Management Group.

Object Management Group. (2021). RAAML. Risk Analysis and Assessment Modeling Language (RAAML) Libraries and Profiles. Milford, MA, USA: OMG.

OMG, O.M. (2021). Risk Analysis and Assessment Modeling Language (RAAML) Version 1.0. Milford, MA 01757 USA: OMG.

Pawlicki, T. (2018). Experiences with STPA in Radiation Therapy. MIT STAMP Workshop. Cambridge MA: MIT.

Pope, G. (2019). Risk Management Using Systemic Theoretic Process Analysis (STPA). Livermore, CA (United States): Lawrence Livermore National Lab (LLNL).

Reason, J. (1995). Understanding adverse events: human factors. Quality in Health Care, pp. 4:80-89.

Ressler, G.E. and Vernacchia, M.A. (2018). Designing in Safety Tutorial. International System Safety Society Conference Proceedings. Phoenix: ISSS.

Ressler, G.E. and Vernacchia, M.A. (2019). Helping Beginners Use System Safety Methodologies. International System Safety Conference ISSC37 (pp. 4, 131). Norfolk VA: ISSS - International System Safety Society.

Ribeiro, D.D. (2016). A Systems Approach to the Development of an Aircraft Smoke Control System. MIT STAMP. Cambridge MA: MIT.

Silvis-Cividjian, N., Verbakel, W., and Admiraal, M. (2018). Application of STPA in Radiation Therapy: a Preliminary Study. MIT STAMP Workshop. Cambridge MA: MIT.

Skedung, L.A. (2012). Feeling Small: Exploring the Tactile Perception Limits. Scientific Reports, 1-6.

Thomas, J. (2020). 8.4.1.7 When STPA Results Surprise You - An industry case study employing STPA, Fault Trees, FMEA, and HAZOP. MIT STAMP Virtual Workshop. Cambridge MA: MIT.

Thomas, J. and Gibson, M. (2020). Industry Trials to Evaluate STPA's Effectiveness and Practicality for Digital Control Systems. MIT STAMP Virtual Workshop. Cambridge MA: MIT.

Vernacchia, M.A. (2016). Defining Safety Requirements for Human-Machine Interactions. International System Safety Conference (p. 6). Orlando: ISSS.

Vernacchia, M.A. (2016). Interim General Motors System Safety Training. Milford MI: General Motors Company.

Vernacchia, M.A. (2017). System Concept Analysis Methods. Milford: GM Internal Document.

Vernacchia, M.A. (2018). Delivering Safety Through Early Analysis Methods. ISSC 36 Conference Proceedings (p. 13). Phoenix: ISSS.

Vernacchia, M.A. (2018). Integration of STPA into GM System Safety Process. MIT STAMP Workshop (p. 5). Cambridge MA: MIT.

Vernacchia, M.A. (2019). Integrating STPA into Large Organizations - Lessons Learned at General Motors. ISSC37 - International System Safety Conference. Norfolk VA: International System Safety Society.

Vincoli, J.W. (2005). Basic Guide to System Safety, Second Edition. John Wiley and Sons, Inc., Hoboken, NJ, USA.

Young, W. (2018). STPA in Cybersecurity. MIT STAMP/STPA Workshop. Cambridge: MIT.

3. DEFINITIONS

ACCIDENT: Unexpected and unplanned event or circumstance resulting in a loss. (Leveson, MIT)

ARCHITECTURE: Representation of the structure of the item or functions or systems or elements that allows identification of building blocks, their boundaries, and interfaces, and includes the allocation of functions to hardware and software elements. (ISO 26262)

CONCEPT OF OPERATIONS: A document produced early in the requirements definition process to describe what the system will do (not how it will do it) and why (rationale).

NOTE 1: Usually referred to as “ConOps.”

NOTE 2: It should also define any critical, top-level performance requirements or objectives (stated either qualitatively or quantitatively) and system rationale. (INCOSE Systems Engineering Handbook)

CONDITION: State of something, especially with regard to its appearance, quality, or working order. (ISSS working group)

CONTROL ACTION: A command or feedback item created by, or used by, a system element to perform its function(s).

NOTE: Functions “do things,” whereas, control actions are the commands/activities driven by the functions, and/or are the feedback needed by the element(s) to perform the function(s).

CONTROL STRUCTURE: Hierarchical structures where each level imposes constraints on the activities of the level beneath them and accidents are viewed as the consequence of inadequate control of safety constraints.

NOTE: (<https://systemsthinkinglab.com/stamp/>); provides a method to document and graphically depiction of the functional design of the system (Leveson); systems represented as interrelated components kept in a state of dynamic equilibrium by feedback loops of information and control. (Leveson)

DEDUCTIVE ANALYSIS: Analysis approach starting out with a general statement, or hypothesis, and examines the possibilities to reach a specific, logical conclusion; the general to the specific. (<https://www.livescience.com/21569-deduction-vs-induction.html>)

ELEMENT: System or part of a system including components, hardware, software, hardware parts, and software units. (ISO 26262)

EXPLORATORY ANALYSIS: Analysis approach that considers multiple points of view and tries to anticipate all possible objections to, or flaws in, a particular position, with the goal of seeking the truth; used when cause and effects may not be known or not known in sufficient detail. (https://en.wikipedia.org/wiki/Exploratory_thought)

FAIL SAFE: Characteristic of a system whereby any malfunction affecting the system safety will cause the system to revert to a state that is known to be within acceptable risk parameters. (FAA System Safety Handbook)

FAILURE: Termination of the ability of an element to perform a function as required. (ISO 26262)

FAULT: Abnormal condition that can cause an element or an item to fail. (ISO 26262)

HARM: Physical injury or damage to the health of persons. (ISO 26262)

HAZARD: Any set of system states, events, or conditions together with a particular set of environmental conditions that will lead to a loss.

HAZARD ASSESSMENT: Process involved in identifying and determining the hazard level. (Leveson)

HUMAN FACTORS (ERGONOMICS): Applied science concerned with designing and arranging interfaces people use so that both people and interfaces interact efficiently and safely.

INDUCTIVE ANALYSIS: Analysis approach using Inductive reasoning makes broad generalizations from specific observations, the specific to the general. (<https://www.livescience.com/21569-deduction-vs-induction.html>)

INTERACTION: System state affects or influences one or more vehicle systems, or system elements, even if they may not be connected or communicating with each other.

INTERFACE: Occurrence of direct data/info transfer and/or communication between two systems or system elements within a system.

LOSS: Something no longer possessed or having less of something. May include harm to people; damage or destruction to property or equipment; environmental harm, loss of mission objectives; etc. caused by losing something.

MALFUNCTION: A failure to function in a normal or satisfactory manner. (Oxford Languages)

MISHAP: An accident

REQUIREMENTS: Statements describing essential, necessary, or desired attributes. (FAA System Safety Handbook)

RISK: The combination of the probability of occurrence of harm and the severity of that harm. (ISO 26262)

RISK—ACCEPTABLE: Risk that is understood and agreed to by the program/project, governing authority, mission directorate, and other customer(s) such that no further specific mitigating action is required. (NASA Systems Engineering Handbook)

RISK—RESIDUAL: Risk remaining after the deployment of safety measures. (ISO 26262)

RISK—UNREASONABLE: Risk judged to be unacceptable in a certain context according to valid societal moral concepts. (ISO 26262)

RISK ASSESSMENT: Process of determining the risk level (quantifying risk). (Leveson)

RISK LEVEL: Function of the hazard level combined with (1) the likelihood of the hazard leading to an accident, and (2) hazard exposure or duration. (Leveson)

SAFETY: Absence of losses

SEVERITY: The magnitude or degree of consequences of an accident (mishap/adverse event)

SUB-SYSTEM: An element of a system that, in itself, may constitute a system. (FAA System Safety Handbook)

SYSTEM: Collection of elements that work together to achieve a common goal .

SYSTEM SAFETY ENGINEERING: Subdiscipline within Systems Engineering that deals with safety to identify safety-related risks and eliminate or control them.

SYSTEMS ENGINEERING: Engineering discipline whose responsibility is creating and executing an interdisciplinary process to ensure that the customer and stakeholder's needs are satisfied in a high quality, trustworthy, cost-efficient, and schedule compliant manner throughout a system's entire life cycle. (INCOSE)

TRADE OFF: Decision making actions that select from various requirements and alternative solutions on the basis of net benefits to the stakeholders. (INCOSE)

TRADE STUDY: A means of evaluating system designs by devising alternative means to meet functional requirements, evaluating these alternatives in terms of the measures of effectiveness and system cost, ranking the alternatives according to appropriate selection criteria, dropping less promising alternatives, and proceeding to the next level of resolution, if needed. (NASA Systems Engineering Handbook)

UNSAFE CONTROL ACTION: Control action that, in a particular context and worst-case environment, will lead to a hazard.

VERIFICATION: Confirmation that work products properly reflect the requirements specified for them ("you built it right"). (NASA Systems Engineering Handbook)

VALIDATION: Confirmation that the product fulfills its intended use in all of the environments that the product will be used in ("you built the right thing"). (NASA Systems Engineering Handbook)

4. ACRONYMS

CAST	Causal analysis based on system theory
DFMEA	Design failure mode and effects analysis
FIA	Functional interface analysis
FMEA	Failure mode and effects analysis
FTA	Fault tree analysis
HARA	Hazard analysis and risk assessment
HMI	Human-Machine Interaction
INCOSE	International Council on Systems Engineering
ISO	International Organization for Standardization
JASPAR	Japan Automotive Software Platform and Architecture
MBSE	Model-Based Systems Engineering
NASA	National Aeronautics and Space Administration
ODD	Operational design domain
OMG	Object management group
PHA	Preliminary hazard analysis
RAAML	Risk analysis and assessment modeling language
SLC	System-level constraint
SOTIF	Safety of the Intended Functionality

STAMP Systems theoretic accident modeling and processes

STPA System Theoretic Process Analysis

UCA Unsafe/undesired control action

5. DOCUMENT ORGANIZATION

This document is organized as follows:

- Basic STPA approach, recommended practices, lessons learned.
- High-level use of STPA within safety process and STPA with other safety evaluation methods.
- Summary.

6. BASIC STPA APPROACH AND LESSONS LEARNED

6.1 Introduction

This section describes attributes of a basic System Theoretic Process Analysis (STPA) approach. The content in this section comes from practitioner experience, and at times references, and excerpts from the STPA Handbook (Leveson and Thomas, 2018).

The STPA Handbook defines STPA as a hazard analysis technique based on an extended model of accident causation. In addition to component failures, STPA assumes that accidents can also be caused by unsafe interactions of system components, none of which may have failed. Some of the advantages of STPA over traditional hazard/risk analysis techniques are that:

- Very complex systems can be analyzed. “Unknown unknowns” that were previously only found in operations can be identified early in the development process and either eliminated or mitigated. Both intended and unintended functionality are handled.
- STPA can be started in early concept analysis to assist in identifying safety requirements and constraints. These can then be used to design safety (and security) into the system architecture and design, eliminating the costly rework involved when design flaws are identified late in development or during operations. As the design is refined and more detailed design decisions are made, STPA is refined to help make more and more detailed design decisions. Complete traceability from requirements to all system artifacts can be easily maintained, enhancing system maintainability and evolution.
- STPA includes software and human operators in the analysis, ensuring that the hazard analysis includes all potential causal factors in losses.
- STPA provides documentation of system functionality that is often missing or difficult to find in large, complex systems.
- STPA can easily be integrated into your system engineering process and into model-based system engineering.

STPA provides a system engineering and system thinking approach to the hazards analysis and initial requirements definition steps of system safety assessments. It provides a robust methodology to accomplish these tasks that is useful in any industry containing safety-critical systems. STPA effectively evaluates initial architectures and system concepts so appropriate tradeoffs and system balance discussions may occur or decisions may be made. STPA has been effectively used in many industries, and the approach outlined in this section is valid for any number of industries.

NOTE: This section does not intend to teach the STPA process, as other resources such as the STPA Handbook accomplish that goal. Please refer to these other resources for specific STPA questions. See Section [2](#) for additional resources.

6.2 STPA Method Overview

There are four steps in a basic STPA system safety evaluation. This section outlines these steps and provides recommendations to effectively execute these steps. The steps in a basic STPA are shown in [Figure 1](#) (excerpt from STPA Handbook), along with a graphical representation of these steps.

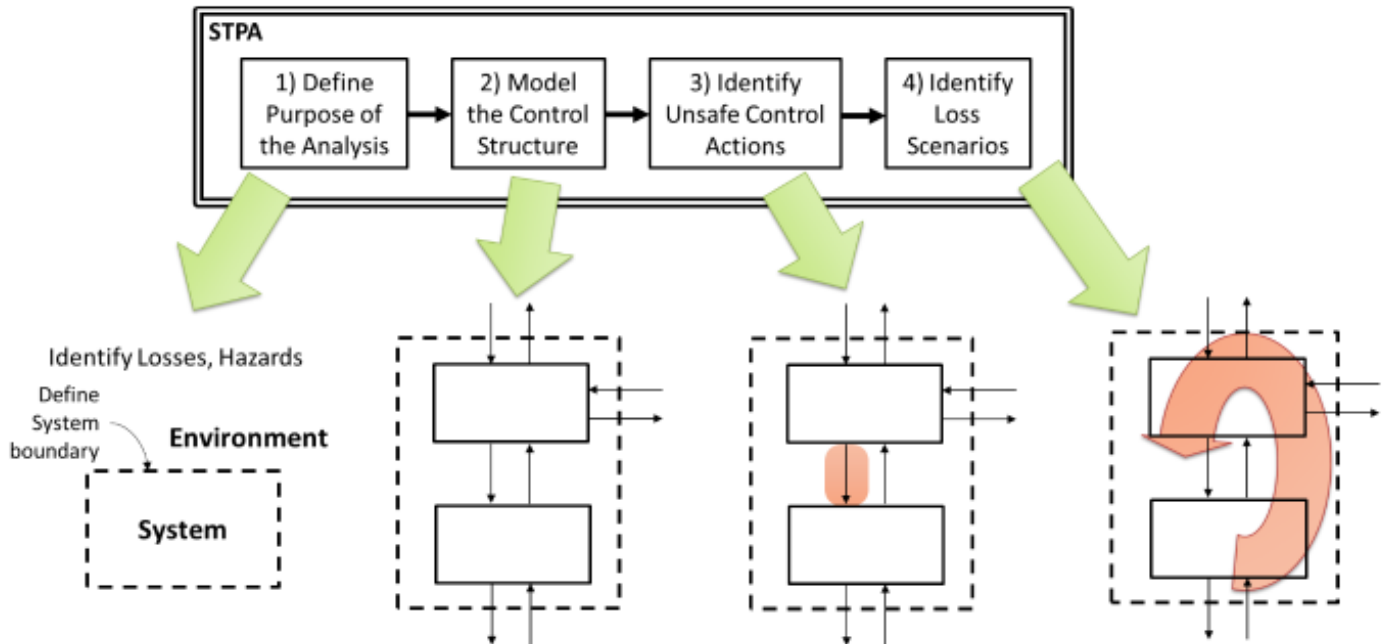


Figure 1 - Overview of the basic STPA method (Leveson and Thomas, 2018)

Used with permission from Leveson and Thomas. All other rights reserved.

The first step is to define the purpose and scope of the STPA evaluation effort. Accidents (losses) and safety goals associated with automotive applications are defined along with potential hazardous states (hazards) that may result from system element misbehaviors.

STPA can be used as a standalone process, or it can be used alongside existing frameworks. When used in an ISO 26262 framework, the STPA accidents and hazards are associated with potential harm to human operators, vehicle occupants, and surrounding pedestrians. (SAE J2980 is a good reference with examples of automotive hazard analysis.) When used in other frameworks, the STPA accidents and hazards can be linked to mishaps, hazards, or similar concepts.

High-level system constraints (safety goals) are developed by reformatting the hazards into constraints that will address the specific hazard. When used in an ISO 26262 framework, the STPA high-level system constraints can be listed as “safety goals,” which provides an approach to link safety requirements to safety goals and then to hazards¹. The STPA high-level system constraints can be linked to safety objectives, high-level safety requirements, or similar concepts when used in other frameworks.

Building a hierarchical “control structure” is the second step. A control structure captures the functional interactions between system elements using a series of command and feedback loops. It is advantageous to model the system using a high-level control structure at an abstract level and then refine it to more detailed levels as the evaluation progresses. The control structure should not include implementation details but rather logical elements representing software controllers, actuators, sensors, and processes that the system controls. A control structure need not be limited to electromechanical elements of the system but may extend beyond the technical system to include elements such as human operators, maintenance personnel, a control room or fleet operations center, and interactions with the external environment.

¹ ISO 26262 defines a “safety goal” as a top-level safety requirement from the hazard analysis and risk assessment (HARA) at the vehicle level.

The third step is to identify unsafe or undesired control actions. This can involve examining functional responsibilities (functions) assigned to the various control structure elements. Undesired or unsafe control actions (UCAs) associated with each action that achieves a function are defined using standard syntax and then linked to the hazards defined in the first step. A UCA may lead to more than one hazard, and hazards may result from multiple UCAs (many-to-many relationships). It is worth noting that developing UCAs may help identify additional hazards not identified in the first step.

STPA is a systematic process to reduce and eliminate open-ended mental processes that result in gaps and oversights. Therefore, this step should avoid using non-systematic methods such as unstructured brainstorming.

The final step identifies the circumstances, conditions, or reasons (causal scenarios/causal faults) why UCAs might occur in the system. These scenarios are created to explain:

1. How unsafe or undesired control action leads to a loss event.
2. How incorrect feedback, inadequate requirements, design errors, component failures, and other factors can independently or collectively contribute to unsafe or undesired control actions ultimately lead to losses.
3. How safe (or desired) control actions, including hazard mitigation actions, might be required but are not commanded, are commanded improperly, or executed improperly.

More than one causal scenario can exist for each unsafe or undesired control action.

Causal scenarios create additional technical requirements, drive the architecture, and highlight conflicting requirements that must be resolved before the design can proceed.

If applied during early development, STPA can provide recommendations and new design decisions. It can also:

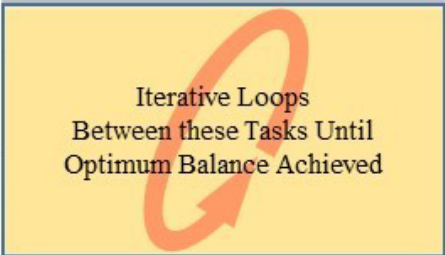
- Evaluate design decisions and highlight gaps when applied to existing concepts.
- Identify functional and behavioral requirements.
- Define human procedures and training requirements.
- Define test cases and create test plans.
- Develop leading indicators of risk.

6.3 Use of Systems Engineering Perspective

Effective safety analysis methods provide value through the entire engineering design, development, testing, and manufacturing phases of a system's implementation process. These analysis methods significantly impact system content in the early phases of a program as major expenditures for critical component procurements and manufacturing tooling have yet to be committed. There is much less cost associated with design changes at this point in the process than there are with changes that occur later in the process (e.g., it is much cheaper to move lines on paper than it is to move/change actual physical tooling lines in a manufacturing environment) (Vernacchia, 2018).

Early in the design process, various design concepts are explored to determine the optimized balance between system imperatives such as cost, timing, mass, performance, and safety "acceptable risk" requirements. This "system architecture balance" phase typically contains tasks such as those shown in [Table 1](#), with some tasks having an iterative nature between them. Of course, if an optimal balance cannot be achieved, the iterative nature would be extended to early tasks to modify stakeholder expectations and initial requirements (Vernacchia, 2018).

Table 1 - System concept/architecture tasks involving early safety analysis methods (Vernacchia, 2018)
Used with permission from Vernacchia. All other rights reserved.

System Concept / Architecture Tasks	Safety Analysis Methods Areas
Stakeholder Expectation Determination	
Initial Requirements Identification	
Architecture Development / Modification	 Iterative Loops Between these Tasks Until Optimum Balance Achieved
Balance, Trade Studies, Optimization Activities Leading to Design Selection within Functional Domain(s)	
Analysis Verification (Functional, Performance, and Safety)	
Program Validation Against Stakeholder Expectations	
Concept / Architecture Selection	

An underlying assumption in [Table 1](#) is that system safety is an integrated part of the system engineering process to be optimally effective. The iterative loop shown in [Table 1](#) may be further illustrated by a systems engineering “cascading conic” depiction shown in [Figures 2, 3, 4, 5, and 6](#). [Figure 2](#) shows the initial allocation of system imperatives to major systems. [Figure 3](#) shows the interactions between these major systems during the balance, tradeoff, and optimization steps. [Figure 4](#) shows the rebalance and optimization at the top system-level leading to an initial concept/architecture, and [Figure 5](#) depicts the concept/architecture selection.

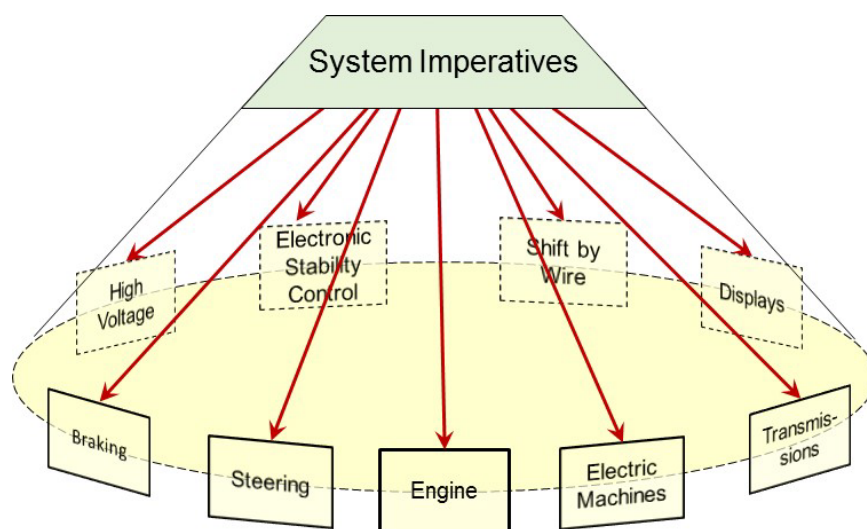


Figure 2 - Initial allocation to major systems (Vernacchia, 2018)
Used with permission from Vernacchia. All other rights reserved.

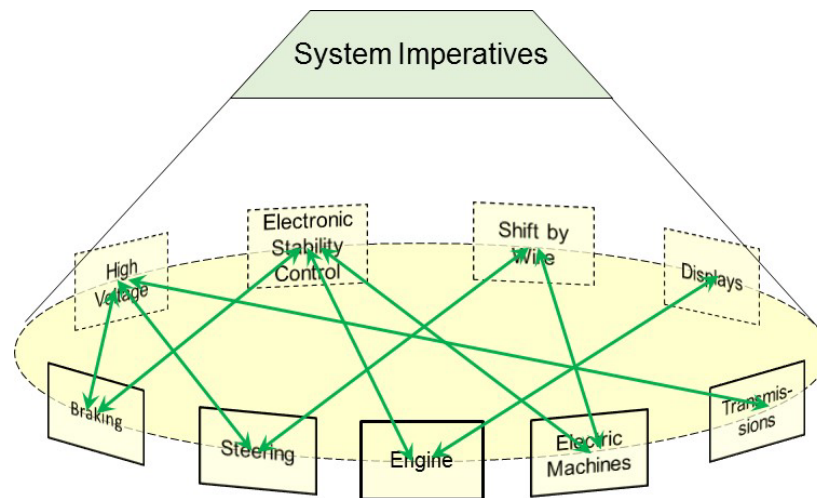


Figure 3 - Interactions between major systems (Vernacchia, 2018)
Used with permission from Vernacchia. All other rights reserved.

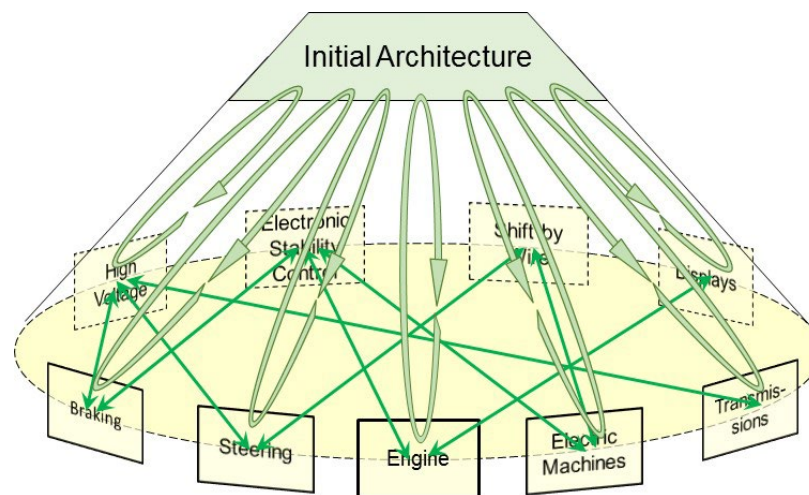


Figure 4 - Rebalance and optimize initial concept (Vernacchia, 2018)
Used with permission from Vernacchia. All other rights reserved.

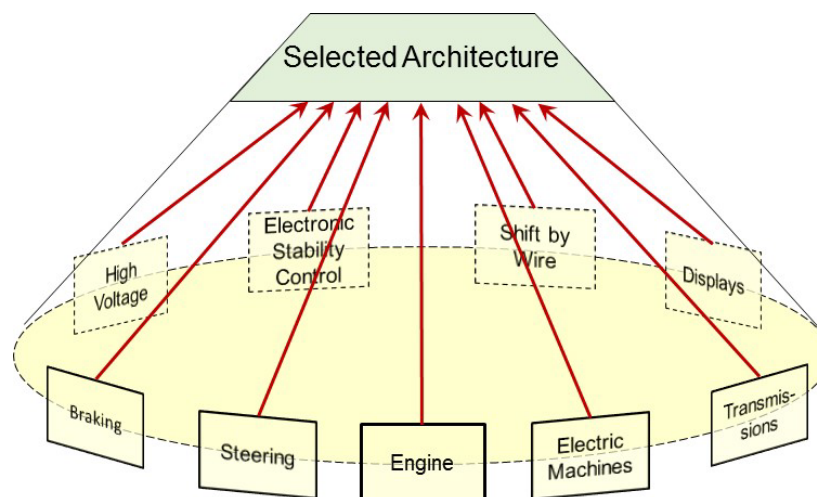


Figure 5 - Selected concept/architecture (Vernacchia, 2018)
Used with permission from Vernacchia. All other rights reserved.

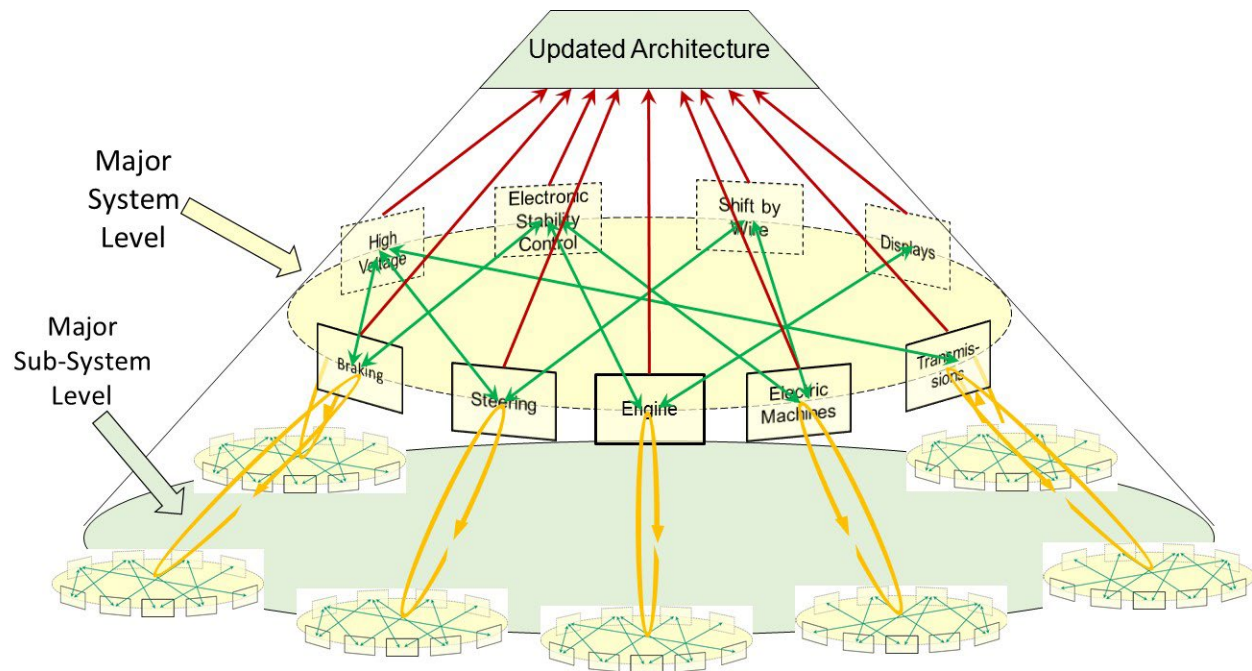


Figure 6 - Multiple level balancing and optimization for continued concept/architecture improvement (Vernacchia, 2018)
 Used with permission from Vernacchia. All other rights reserved.

Figure 6 illustrates the allocation to the major subsystems of each major system where the balance, tradeoff, and optimization process repeats at this lower level, and updates are transposed back up to the top level. Two levels of evaluation tend to be sufficient for initial safety task activities. Deeper levels of the engineering process are very well served by familiar evaluation methodologies such as DFMEAs, FTAs, and FIAs.

Within this system's engineering process framework, safety tasks performed early in the design process support the balance and optimization tasks by identifying what hazardous conditions could arise due to system element misbehavior and assessing the risk associated with such hazards. This effort defines the required system content and the detection and mitigation strategies necessary to eliminate or manage the assigned risk to acceptable levels. To be effective, these detection and mitigation strategies must be part of the balancing and optimization of the system.

A key asset for this top-down, iterative, and interactive process is a "chief system architect" who understands the technical nature of the major system interactions and should have the expertise to broker tradeoffs and compromises between functional areas. If no such person is available, it is useful to engage "system architects" from the different functional domains and engage them in the "roll-down" and "roll-up" activities involving system imperative and high-level constraints.

6.4 STPA Scope Determination

The first step is to determine the system scope under consideration and its hierarchal nature. The scope of the STPA evaluation may change when going from the top hierarchal controller into a deep nested lower level controller, so the previous systems engineering-based section is an effective method to determine system content and element functions. Additional key considerations found effective to consider when defining the system are:

- Define the purpose of the analysis and the strategy/focus of the analysis. (For example, is this a safety analysis, security analysis, systems analysis, or some combination of the three? This purpose should be identified early on as it will lead to the identification and avoidance of losses, as mentioned in the next section.)
- Is safety-relevant security included in the analysis? Security's impact on safety needs to be considered from the beginning. Correlating security-related UCAs and scenarios with the system hazards and losses identifies the causal factors that have priority (Young, 2018).
- List all elements at the highest abstract level.

- The STPA control diagram is revisited throughout the analysis process, serving as a reference for the STPA evaluation and guiding discussions with design and test engineering teams.
- For these reasons, it is important to balance the level of abstraction and detail such that all aspects of the system can be analyzed across diverse teams.
- STPA should be applied initially at the highest abstract levels and then at more defined lower levels. A high level of abstraction is expected at the early definition stage, with details added as the project design progresses.
- Ask these upfront questions:

A system to do {What = Purpose}

By means of. {How = Method}

In order to contribute to {Why = Goals}

While {Constraints, restraints, ODD, etc.}

System actors {Who or what is interacting with the system?}

Delivered by {When = Milestone deliveries}

- The key part is capturing system assumptions. If system assumptions are not detailed early in the system scope and boundary, difficulty will develop when trying to understand causal analysis and process model flaws. So, capture all assumptions of system behavior early on. Examples: “This analysis will only focus on automatic transmissions,” or “This analysis is assumed to have Level 3 autonomy only.”

Additional items to consider while determining scope and boundary are:

- List assumptions as to the basis for focused analysis.
- Think of different ways to frame the problem.
- Avoid considering only the items we control, but intentionally look at all aspects with a specific interest in items that are not controllable (or perceived not to be controllable).
- Must ensure that we define and solve the right (engineering) problem.
- Group list of elements but avoid detailing lower levels and instead focus on top-level interactions of an element. It is okay to show some dependencies, but it is better to focus on top-level understanding.
- Consider not just communication but process interfaces (e.g., high voltage for electric powertrain).
- Identifies the system boundaries as well as defines the components (subsystems) inside the system boundaries targeted by STPA.
- Definitions of the components inside the system boundaries to the granularity of the control structure.

6.5 Potential Loss (“Accident”) Identification

The next step is to identify potential accidents (e.g., “loss,” “harm”) that may occur if certain hazardous states and specific circumstances are present. In STPA, a “loss” can occur due to an accident. ISO 26262 uses “harm” as an equivalent for an accident. This document will use each of the three words—accident, loss, and harm—as equivalent terms.

STPA can be used to target any loss that is unacceptable to stakeholders. If more than one loss is included, they can be ranked and prioritized. Because every STPA result will be traceable to one or more losses, the analysis results can be easily ranked and prioritized based on the losses to which they refer.

Before any analysis begins, the stakeholders must identify the losses they want the analysis to focus on. The losses to be considered may be defined by management, government regulations, or customers. A general approach to identifying losses may involve:

1. Identify the stakeholders, e.g., users, producers, customers, operators, etc.
2. Stakeholders identify their “stake” in the system. What do they value? For example, human life, a fleet of automobiles, electrical power generation, transportation, etc. What are their goals? For example, to maintain a fleet of automobiles, provide transportation, provide medical treatment, provide electrical power generation, etc.
3. Translate each value or goal into a loss, e.g., loss of life, loss of vehicle functions, loss of electrical power generation, loss of transportation, etc.

As an example, automotive application accidents will typically fall into one of the following categories:

- Vehicle collides with another vehicle.
- Vehicle collides with pedestrian(s).
- Vehicle occupant injury.
- Vehicle collides with infrastructure.

In addition to these categories, accidents (i.e., harm) result from other occurrences, such as a fire or an electrical shock that may occur without a vehicle collision.

Depending on the purpose and scope of the STPA effort, the evaluation of accidents (losses) can expand beyond safety-related incidents. Accidents can include various other types of losses, such as performance, financial, quality, risks, and threats. Accidents do not have to be limited to safety-related losses and should include other important design aspects as appropriate. A hazard need not always lead to a loss but may lead to a loss under certain conditions. For example, unintended braking may be a hazard for automated emergency braking (AEB), but it may lead to a loss if there is a following vehicle that collides with the subject vehicle.

Additional key aspects to consider when identifying accidents and losses are:

- Use top-level strategy and priorities to focus on consequences directly related to the strategy, purpose, and definition of the system being analyzed.
- Verbs may be used to help identify system accidents and hazards:
 - Autonomous vehicle (AV) colliding with mobile object.
 - Autonomous vehicle (AV) injuring passengers without a collision.
- If security is included in the analysis, distinguish between safety-related accidents and non-safety-related security losses, such as loss of reputation, intellectual property, privacy or breach of customer personal information. Non-safety-related losses can be related to quality, performance, security, and financial.
- For traceability, use unique labels for each accident and loss (security loss not impacting safety).

Next are potential hazardous situations (system-level hazards) that may lead to accidents/losses in specific circumstances that need to be determined.

6.6 Potential “Hazard” Identification

Hazards should first be identified at the system level, i.e., associated with the high-level control structure representation. (More hazards may be defined as more detailed control structures are evaluated.) High-level control structure hazards for vehicle applications (e.g., road vehicles, aerial vehicles, space vehicles, etc.) might include vehicles exceeding safe operating envelope, vehicles not maintaining a safe distance from nearby objects, or vehicles entering dangerous areas/regions. These hazards may lead to multiple accidents and losses. As such, the STPA process links each hazard to the resulting accidents and losses by typically documenting the accident/loss identifier in brackets after the hazard.²

The following criteria from the STPA Handbook (Leveson and Thomas, 2018) is useful for defining hazards:

- Hazards are system states or conditions (not component-level causes or environmental states).
- Hazards lead to an accident (or loss) in some worst-case environment.
- Hazards must describe states or conditions to be prevented.

The STPA Handbook goes on to outline some common mistakes when identifying hazards:

- Confusing hazards with causes of hazards.
- Too many hazards containing unnecessary detail.
- Ambiguous or recursive wording.
- Confusing hazards with failures.
- Key pitfalls: If you use words like “causing” or “resulting in,” you are likely combining a hazard and a potential cause. See the STPA Handbook for examples.

It may be helpful to list out a few possible mechanisms for developing the initial hazard list before entering the UCAs (e.g., field data). The [link here](#) shows a high-level HAZOP to help find this initial list.

System-level hazards will be a list that will evolve as the STPA process is implemented, especially as the control diagram(s) are refined. Start with the list of elements created for a top-level understanding of the system. Use losses that were identified earlier to identify related hazards. Use verbs to identify hazards and avoid compound sentences. For example:

- AV breaches the minimum safe distance of a forward mobile object.
- AV exposes passengers to unhealthy g-forces.

If security is included in the analysis, distinguish between hazards with the potential cause of an accident and losses that don't impact safety. Listing both safety and non-safety related losses adds clarity to the boundary of the analysis. As the analysis proceeds, it is possible that elements could move to the other list. It is important to look across the analysis boundary to leverage the iterative nature of STPA.

For traceability, use unique labels for each hazard and loss.

² In ISO 26262, “accidents” are not specifically linked to hazards. ISO 26262 investigates when hazardous situations may lead to “harm” under certain operating conditions or modes. An accident is an example of harm. Even with this distinction, ISO 26262 links hazards to an accident (harm) by hazard analysis and risk assessment (HARA). A system item is evaluated without internal safety mechanisms for various operational situations and operating modes. An operational situation or mode may define a hazardous event, which, unmitigated, may lead to a loss.

ISO 26262 provides one way to assess the impact of a potential loss using three parameters: severity, exposure, and controllability. Each of these parameters has multiple levels of degree and then can be combined to define an automotive safety integrity level (ASIL) rating. The five ISO 26262 classifications are (starting with most significant) ASIL D, ASIL C, ASIL B, ASIL A, and QM. QM (quality managed) is the least significant classification and allows a robust engineering process to address system safety concerns.