

Systems Thinking Applied to Automation and Workplace Safety

by
Nathaniel Arthur Peper

B.S. Mechanical Engineering, United States Military Academy, 2007

Submitted to the MIT Sloan School of Management and the Department of Aeronautics
and Astronautics in partial fulfillment of the requirements for the degrees of
Master of Business Administration

and

Master of Science in Aeronautics and Astronautics
in conjunction with the Leaders for Global Operations Program at the
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2017

© Nathaniel Arthur Peper, MMXVII. All rights reserved.

The author hereby grants to MIT permission to reproduce and to distribute publicly paper and electronic
copies of this thesis document in whole or in part in any medium now known or hereafter created.

Signature redacted

Signature of Author

MIT Sloan School of Management and the Department of Aeronautics and Astronautics
May 12, 2017

Signature redacted

Certified by.....

Nancy G. Leveson, Thesis Supervisor
Professor of Aeronautics and Astronautics

Signature redacted

Certified by.....

John S. Carroll, Thesis Supervisor
Gordon Kaufman Professor of Management

Signature redacted

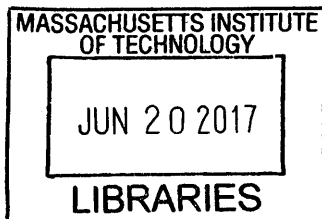
Approved by.....

Youssef Marzouk
Chair of the Graduate Program Committee
Associate Professor, Aeronautics and Astronautics

Signature redacted

Approved by.....

Maura Herson
Director of MIT Sloan MBA Program
MIT Sloan School of Management



ARCHIVES

THIS PAGE INTENTIONALLY LEFT BLANK

Systems Thinking Applied to Automation and Workplace Safety

by

Nathaniel Arthur Peper

Submitted to the MIT Sloan School of Management and the Department of Aeronautics and Astronautics on May 12, 2017, in partial fulfillment of the requirements for the

degrees of

Master of Business Administration

and

Master of Science in Aeronautics and Astronautics

Abstract

This thesis presents the results of a study to compare Systems-Theoretic Process Analysis (STPA), a hazard analysis methodology based on a new model of accident causation called Systems-Theoretic Accident Model and Processes (STAMP), with the traditional assessments recommended by industry standards for analyzing safety risks in modern manufacturing workplaces that are increasingly incorporating automated systems. These increasingly complex, modern socio-technical systems are introducing new problems in the manufacturing environment that traditional methods of analysis were not designed to analyze. While these traditional methods have previously proven effective at analyzing hazards, the increasing levels of complexity and technological advancement in the factories are surpassing the limits of traditional assessment capabilities. Today's continuous search for opportunities to automate manufacturing process makes this a critical time to ensure that the hazard analysis methodologies in use are capable of providing an effective and efficient analysis.

STAMP and STPA were developed specifically to understand and analyze modern, complex socio-technical systems that are introducing new types of accidents with causes beyond traditional component failures. This thesis provides background and discussion of traditional models and methods, of the current industry standard method, and of the proposed method. The current and proposed methods are then used on an actual semi-automated manufacturing process being implemented in an aerospace manufacturing company and analyzed with a set of criteria to determine their effectiveness and efficiency. The results of this analysis determine that STPA is better equipped for the modern manufacturing environment.

Thesis Supervisor: Nancy G. Leveson

Title: Professor of Aeronautics and Astronautics

Thesis Supervisor: John S. Carroll

Title: Gordon Kaufman Professor of Management

*The author wishes to acknowledge the
Leaders for Global Operations Program for its support of this work.*

Acknowledgments

First and foremost, I would like express my deepest gratitude to my wife, Kelly, for putting up with this crazy journey through the Leaders for Global Operations (LGO) program. I know there wasn't anything easy about this program for either of us and without your support, none of this would have been possible. To my daughter, Makenzie, for always being ecstatic to see me walk in the door, no matter how late or distracted I was. To my son, Ian, for giving me the motivation to just stay up and get my work done because you were going to be waking up in the middle of the night anyways. I love you all.

I'd also like to express my thanks to my incredible advisors during this process. To Professor Nancy Leveson for her guidance and advice in class and during this and other work. Your insights and lessons on safety were initially difficult to internalize, but now I feel like I have an entirely new perspective that will hopefully make a difference in the future. To Professor John Carroll, for your insight and support to understand why safety is such a complex issue to change, let alone even address, in organizations.

Thank you to the company that created the opportunity for this work, especially my supervisors, colleagues, mentors, and friends that listened to my countless questions, provided the insight necessary to understand the company, and offered support even when I didn't realize I needed any.

Last but not least, thank you to everyone that has had an influence on my life and helped me get to this moment. I would not be where I am today or have the opportunities that have been presented to me if it weren't for a life filled with incredible family, friends, and mentors from childhood, to school, to life in the Army, and to now. Your support and influence will never be forgotten.

THIS PAGE INTENTIONALLY LEFT BLANK

Contents

1	Introduction	13
1.1	Motivation	13
1.2	Project Objective	15
1.3	Thesis Objective	15
1.4	Overview of Approach	16
1.5	Overview of Thesis	17
2	Background	19
2.1	Industry Overview	19
2.2	Company Overview	20
2.3	Manufacturing Workplace	21
2.4	Workplace Safety	22
3	Literature Review	25
3.1	Risk	25
3.2	Safety	31
3.3	Accidents Models and Assumptions	33
3.3.1	Domino Accident Model	33
3.3.2	Domino Model Extensions	34
3.3.3	Swiss Cheese Accident Model	35
3.3.4	Systems-Theoretic Accident Model and Processes	37
3.4	Analysis Methods	41
3.4.1	Failure Modes and Effect Analysis	42

3.4.2	Fault Tree Analysis	43
3.4.3	Job Hazard Analysis	44
3.4.4	Systems-Theoretic Process Analysis	46
4	Case Study of a Semi-Automated Manufacturing Process	55
4.1	Overview of Manufacturing Process	55
4.2	Current Analysis Process - SRA	56
4.2.1	ANSI and American National Standards	56
4.2.2	ANSI Standard Analysis Process	57
4.2.3	Task-based Risk Assessment Methodology	57
4.2.4	Discussion of Task-based Risk Assessment Methodology	66
4.2.5	Discussion of Task-based Risk Assessment in Practice	70
4.3	Proposed Analysis Process - STPA	80
4.3.1	Defining Workplace Accidents	80
4.3.2	Defining Hazards and System Boundaries	81
4.3.3	Model Development – Safety Control Structure	83
4.3.4	STPA Steps 1 & 2	89
4.3.5	Discussion of STPA Process and Results	95
4.4	Comparison of the Analyses	100
5	Conclusions and Recommendations	105

List of Figures

1-1	Representative Company Comparison Criteria Developed by Company Stakeholders	17
3-1	Risk Matrix by DoD [11]	29
3-2	Components of Risk [18]	31
3-3	Heinrich’s Domino Accident Model	34
3-4	Reason’s Swiss Cheese Accident Model	35
3-5	Emergent Properties	38
3-6	Example Control Loop	39
3-7	General Form of a Model of Socio-Technical Control	40
3-8	Example Fault Tree from the original Bell Laboratory study[33]	43
3-9	Example Job/Activity Hazard Analysis Form	45
3-10	Standard Control Loop	48
3-11	Basic Depiction of a Hierarchical System[22]	50
3-12	Standard Control Structure Diagramming Convention[7]	50
3-13	General Control Loop with Causal Factors	52
4-1	Task-based Risk Assessment Flowchart [6]	58
4-2	Severity - Exposure - Avoidance Ratings and Criteria [6]	62
4-3	Risk Level Decision Matrix [6]	63
4-4	Minimum Risk Reduction Measures [6]	64
4-5	Hierarchy of Control [23]	65
4-6	Example Task-based Risk Assessment Form [6]	66
4-7	Comparison of Task-based Risk Assessment in Theory vs. Practice	72

4-8	Example Results Format from the Task-based Risk Assessment	75
4-9	Example Results Analysis	76
4-10	TRA Risk Reduction Measures Hierarchy of Controls Classifications	79
4-11	AGV Safety Control Structure	85
4-12	Product Transportation Vehicle Safety Control Structure	86
4-13	Combined Product Transportation System Safety Control Structure	87
4-14	Robotic System Safety Control Structure	88
4-15	AGV SCS Segment	90
4-16	General Control Loop with Causal Factors	92
4-17	STPA Risk Reduction Measures Hierarchy of Controls Classifications	99
4-18	Assessment Analysis Criteria Results	101
4-19	Risk Reduction Measures Hierarchy of Controls Classifications Comparison	102

List of Tables

3.1	Qualitative Senses of the Term "Risk" [15]	26
3.2	Quantitative Senses of the Term "Risk" [15]	26
3.3	Example FMECA Worksheet	42
3.4	Control Action Table	51
4.1	Lifecycle Phases IAW RIA TR R15.306-2014	58
4.2	Example Hazards from ANSI/RIA Standard	61
4.3	Example Control Action Table for an AGV	91

THIS PAGE INTENTIONALLY LEFT BLANK

Chapter 1

Introduction

This chapter details the motivation for this work, including the project and thesis objectives. It also provides an overview of the approach of the research and an overview of the thesis.

1.1 Motivation

The modern-day aerospace industry is typically known as being incredibly diverse and technologically advanced, with a strong focus on product safety. This is shown through the increasing number of professional disciplines involved in research and development, of engineering disciplines involved in design, of technologies included in products, of moving parts and pieces required, and of lines of code required in the software to make modern products, all while the number of aircraft incidents has steadily declined since the 1970's.[2] However, this view of safety in the aerospace industry leaves out an important period during each product's lifecycle — manufacturing.

Technological advancements have increased the complexity of not only the aerospace products themselves, but also the production systems used to manufacture those same products. With the increasing number of people, parts, plans, tools, and interactions among them to build a final product, the physical and sociological levels of complexity on the manufacturing floor has drastically increased. Furthermore, the introduction of new technologies, such as the computers and software involved in the automated and semi-automated manufacturing processes, continues to change the nature and types of accidents seen on the manufactur-

ing floor. These advanced and highly complex technologies are leading to accidents that are not the result of component failures, but of unsafe interactions among properly working components of the system.

While a company may know there is an issue and want to correct it, often the solution selected is to refocus on the same traditional tools, processes, and analysis techniques that allowed them to arrive at the current state of worker injury rates. The most common traditional hazard analysis techniques are based on reliability theory and assume that accidents are preventable if the system components are made highly reliable or component failure is planned for in the design.[33] This stems from the fact that an immense amount of effort went into the initial development of reliability analysis to understand the relatively simple, electromechanical systems of the past. Since then, the amount of effort to continuously improve the tools and techniques has not kept pace with the technological advances and complexity of the socio-technical systems built today. However, the accident models and analysis techniques proved to be sufficient in the past and companies usually tend to favor processes that have a historical track record of success.

The aerospace manufacturing industry is currently undergoing a major shift from highly manual processes to increasing levels of automation. This change introduces new factors, such as software and increasingly complex human factors, that can be involved in accidents. While most software-related accidents can be traced back to flawed or incomplete software requirements, traditional analysis methods typically overlook these and look to apply more safeguards before understanding the hazard.[21, 18] Additionally, human responsibility is quickly shifting from executing specific procedures to sharing control of systems with automation as a higher-level decision maker.[25, 31] While humans are commonly still blamed for accidents labeled as operator error, many times the accidents are a result of new types of human error related to factors in the environment in which the humans operate. As a result of these considerations, a new approach to accident models and hazard analysis is needed in order to prevent accidents during the design of and management of workplaces and automation, especially in the manufacturing environment.

1.2 Project Objective

The primary objective of the project was to determine if applying a new systems thinking approach to workplace safety could assist a large, aerospace manufacturing company to reduce workplace injuries. The company has been aggressively working to improve their workplace injury rates but has not fully achieved their desired goals. During the same time period, the company has been increasingly introducing cutting-edge technologies to the manufacturing floors with the intention to improve quality, production schedules and costs, and workplace safety. However, over the course of multiple new technology implementations the company has realized that its traditional methods of assessing risk and safety issues for this equipment are no longer achieving the level of effectiveness they once did.

By engaging with cross-functional stakeholders throughout the organization, we determined that one of the biggest opportunities that may benefit from an entirely new perspective was the safety risk assessment processes for automation and workplace safety. The company had already identified their risk assessment process for new implementations on the factory floor as an area of concern and stakeholders across the company were open to change. A number of traditional tools being used, such as technology readiness levels (TRLs) and manufacturing readiness levels (MRLs), were all being compared across the company to realign and improve their ability to mitigate different forms of risk in the production system. By analyzing the assessment process specifically for safety and hazard analysis, the company would be able to determine if the traditional processes achieved the level of effectiveness and efficiency desired for workplace safety in comparison with newer approaches.

1.3 Thesis Objective

The safety risk assessment process is a longstanding methodology that has been repeatedly used within the company for new technology implementations. While a renewed effort to reinforce the same, historically successful behavior may work in an environment that has not changed, the manufacturing floors of the company are dynamically changing at an accelerating rate. Increasing levels of automation and human-robotic interaction are becoming the

norm and the tools used to understand and analyze them are the same tools developed for a highly manual and much simpler manufacturing environment. In order for the company to achieve the step-change improvement desired in workplace safety, it will need to assess if its efforts are focused in the right areas and if it is even equipped with the right tools for success. The objective of this thesis is to determine if a systems thinking approach to workplace safety, named Systems-Theoretic Accident Model and Process (STAMP)[20], is more effective and efficient than the recommended industry standard hazard analysis process [5] for managing safety in modern manufacturing workplaces.

1.4 Overview of Approach

In order to determine if STAMP is more effective and efficient than the industry standard risk assessment process for managing safety in modern manufacturing workplaces, a list of criteria was established with the company stakeholders in addition to the widely-accepted Hierarchy of Control [23] used to analyze the results of each assessment. The list of criteria was developed to assess the feasibility and quality of the analysis process when considering automation and workplace safety within the company. Feasibility criteria were developed to analyze which levels of authority may have implementation issues for either methodology due to the high amount of regulation and oversight in the aerospace industry. Quality criteria were developed to analyze the robustness of the assessment results for contributing to the company's changing strategy toward and goals for workplace safety. A representation of the list of criteria developed is shown in Figure 1-1.

In order to reasonably compare the results of the two entirely different assessment methodologies, a case study of a new semi-automated manufacturing process being implemented into the factory was used. The teams designing and integrating the equipment were highly qualified and relatively small. Due to the sizes of these teams and the benefit of having a cross-functional assessment team, the two separate assessments had teams of similar composition but tried to minimize the number of similar personnel. Additionally, the execution was separated by multiple months to mitigate the effects of learning from the traditional assessment to contribute to the results of the proposed assessment. Lastly, the

Criteria	TRA	STPA
Feasibility		
Compliant to Government Safety Regulations and Standards		
Compliant to Industry Safety Regulations and Standards		
Compliant to Company Safety Regulations and Standards		
Documented Analysis Process		
Quality		
Includes hardware failure accidents		
Includes technological factors in accidents beyond hardware failures, such as system design and requirements flaws (software and component interactions)		
Includes the role of management, operations, and procedures in accidents		
Includes the role of the environment in accidents		
Goes beyond specifying what humans did wrong to explain why they did what they did (includes sophisticated human factors in analysis)		
Creates thorough understanding of the problem before implementing controls IAW the hierarchy of controls		

Figure 1-1: Representative Company Comparison Criteria Developed by Company Stakeholders

traditional assessment process was led by a longstanding company expert in regards to the process, while the proposed STAMP analysis process was led by a newly trained employee.

During the execution of these methodologies, I observed the participants and their execution of each process to collect data on each participant’s personal and professional behaviors during the assessment, the participants’ interactions with each other, the team’s adherence to or deviations from the prescribed methodology, the assessment environment, and the results of each assessment. During and after each assessment, members of the groups were interviewed to obtain a more in-depth understanding of the processes, perspectives, and experiences of individuals within the company.

1.5 Overview of Thesis

In order to present the analysis and results of this thesis, a large amount of context and foundational safety and hazard analysis concepts will be introduced and analyzed. Chapter 2 provides a broad overview of the aerospace industry and the participating company on various dimensions. Chapter 3 reviews the foundational building blocks of risk and safety, accident

model frameworks and their underlying assumptions, and risk assessment and hazard analysis methodologies in the context of complex human and software intensive systems. Chapter 4 will then describe a case study of a semi-automated manufacturing process, review the current company hazard analysis process and discuss its results, discuss the results of a STAMP-based hazard analysis tool called Systems-Theoretic Process Analysis (STPA) [20], and finally compare the results.

Chapter 2

Background

This chapter provides background information on the aerospace industry and the company in which the research was conducted. It further details the dynamics of the company's manufacturing workplace and safety management.

2.1 Industry Overview

The aerospace industry is primarily focused on the research, development, design, manufacturing, and certification of flight vehicles, including their associated subsystems and support systems. These vehicles can be divided into a number of categories based on certain characteristics, such as fixed-wing aircraft, rotary-wing aircraft, rocket and missile systems, space launch vehicles, and spacecraft. Within these flight vehicles there are also industry specific subsystems that include control surfaces, propulsion, and avionics. Additionally, there are numerous support systems that help to build, test, operate, and maintain the flight vehicle and associated subsystems.[2]

Due to the complex nature of flight, the industry is highly focused on technological advancements and relies upon a diverse, cross-disciplinary workforce to bring these complex systems from concept to reality. However, it is also highly regulated by the government due to the characteristics and use of the products once they are delivered to the customers. Because of the scale and complexity of the products in this industry, it is also among the largest manufacturing industries in the world. The organizational industry structure has

relatively few large companies for the actual vehicles and major sub-components, but there are many tier two and three suppliers that include international partnerships.[2]

Although aircraft have progressed technologically from manually flown spruce and fir frames with sewn-on fabric to modern designs that involve metal and composite fibers for structures and flight surfaces, the manufacturing process is still described by some to be a craft process with a mass production mentality. More recently, the continued progress and development in other fields, such as robotics, is providing the ability to automate more and more of these craft processes. Jobs and tasks that could typically only be completed by a human for complexity and quality reasons, such as riveting, drilling and filling aircraft structures with fasteners, are now being designed for manufacturing and automation and given to robotic systems where humans are now supervising and managing the process at a higher level of decision-making.

2.2 Company Overview

For the purpose of this study, the scope of this project will focus on a large aerospace manufacturing company that builds a variety of commercial, fixed-wing airplanes for transportation companies. The company has a long and successful history that includes many firsts in the aerospace industry and has always led the competition in regards to product safety. However, this history of success has not come without numerous internal issues, struggles, and failures. Historically, aircraft have been solely designed for performance, leaving producibility and the production system as afterthoughts. Most recently the company has developed another new airplane that includes many new technologies unmatched across the competitive landscape. In addition to the new technologies being used in the aircraft, the company has started to consider production earlier in the product lifecycle. There are also a number of technologies that have been developed in parallel to revolutionize the manufacturing process. The company is actively searching for opportunities to implement many of these and other new technologies and materials to increase quality, capability, and safety of its production system. However, these studies and implementations must occur concurrently to the current production operations due to various resource and time constraints, such as limited manu-

facturing space, established number of company personnel, fully utilized production system, and committed production schedule.

2.3 Manufacturing Workplace

Due to the massive size and complexity of the products being manufactured and the changing dynamics of competition in the global economy, the company's production system involves a large number of tightly-coupled interactions among people, parts, plans, and tools across the globe and is becoming more complex over time. Manufacturing workplaces that were once solely responsible for fabricating and assembling basic aircraft are now primarily assembling complex components and subsystems to deliver advanced aircraft. The increasing number of tightly-coupled interactions within the factory, reliance upon suppliers, rate of aircraft deliveries, consequences of failure, and focus on cost competitiveness are exposing the production system to new levels of risk in many categories. Due to these dynamics, the company is actively looking for new ways to reduce the overall risk to the production system.

One solution the company is looking toward to mitigate multiple risks is automation within the manufacturing workplace. Aerospace manufacturing has historically been a highly manual process, but technological advances in automation and robotics have reached the level that makes automated or semi-automated manufacturing a possibility for the aerospace industry. In theory, automated manufacturing provides three primary benefits that will help mitigate risks in the production system for quality, production capability, and workplace safety.[35] One example of this potential is the drill and fill process commonly used all over the aircraft during production. Previously, workers had to manually drill precision holes in different metals of various thickness at different angles. These holes were also in locations that caused the workers to contort their bodies into various positions, adding additional stress on their bodies from the process. Ideally, by automating the process, the holes should be drilled with higher quality, the drilling rate should increase, and the safety of the workplace and workers should increase.

However, the implementation of new technology and automation into the manufacturing workplace has introduced new hazards that the company has little to no experience with.

project. Additionally, the delineation of the responsibilities, accountability, and authority within the safety organization and with the other major organizations in the company were still developing as the groups were figuring out how to work together optimally.

THIS PAGE INTENTIONALLY LEFT BLANK

Chapter 3

Literature Review

This chapter provides an overview and discussion of the foundational concepts critical to this thesis, such as risk and safety, accident model frameworks and their underlying assumptions, and risk assessment and hazard analysis methodologies in the context of complex human and software intensive systems.

3.1 Risk

One foundational concept that is common to all industries is risk. Company and functional experts dedicate large amounts of time and resources to mitigate different types of risk that directly impact their sector of industry: manufacturing with component variation, finance with returns on equity and cash, project managers with schedule and performance, etc. For the purpose of this thesis we are concerned with risks involved in the safety of modern socio-technical systems. While it may not be immediately clear that different financial, schedule, and performance risks are a major concern for automation and workplace safety, these risks also have a significant impact on system safety.

In order to see how all of these different risks have an impact on safety, it is important to understand some of the issues within this foundational concept of risk. First, due to the ubiquitous nature of risk, it is not as clearly defined and agreed upon as one would think. There are two primary senses in which the word is used across disciplines – qualitative and quantitative. The qualitative use appears simple and easy to understand when we are only

focused on the general concept of quality of the risk, but its use to discuss a specific event can still lead to confusion. Differing examples of qualitative uses for risk that change the focus of discussion are shown in table 3.1.

Definition of Risk	Example Usage
An <i>unwanted event</i> which may or may not occur	Lung cancer is one of the major risks that affect smokers.
The <i>cause</i> of an unwanted event which may or may not occur	Smoking is by far the most important health risk in industrialized countries.

Table 3.1: Qualitative Senses of the Term "Risk" [15]

In the qualitative interpretations shown here, it is quite easy to see how there could be an issue determining the risks associated with the issue at hand. Is the risk the event or the cause of the event being discussed? Although this may seem like an insignificant difference, two people assessing risk qualitatively can come to entirely different conclusions by focusing on different areas of an issue. In order to fix this issue of vagueness with the use of qualitative terms, many people turn to more structured ideas that possibly include some form of calculation with numbers that makes the concept more tangible. Concepts of risk in the quantitative sense are shown in Table 3.2.

Definition of Risk	Example Usage
The <i>probability</i> of an unwanted event which may or may not occur	The risk that a smoker's life is shortened by a smoking-related disease is about 50%.
The statistical <i>expectation value</i> of an unwanted event which may or may not occur	The expectation value of a possible negative event is the product of its probability and some measure of its severity.

Table 3.2: Quantitative Senses of the Term "Risk" [15]

However, in the quantitative interpretations, there is a significant increase in complexity for the term "risk" when trying to quantify not only the probability of an unwanted event occurring, but also some measure of the severity of the unwanted event. Expectation values

have been used since the 17th century, but they were only recently introduced by Rasmussen to the concept of risk in the influential Reactor Safety Study, WASH-1400.[26] While this report drew much criticism for how it determined different probabilities and severities, the expected value formulation of the term “risk” is regarded by some analysts as the only correct usage of the term today.

$$RiskE(Loss) = Probability(Loss) \times Consequence(Loss)$$

In order for the expected value formulation to be useful for comparing and making decisions, the probability of loss must be perfectly known and the consequence of the loss can be calculated. This may prove true in the simplest cases, but it is impossible to determine an exact probability and consequence of a real-world event. Additionally, expected values are averages, which introduce even greater possibility for error in an attempt to simplify the problem.

Focusing on probabilities, it is common practice to assume that behavior is random and that each behavior is independent and identically distributed. However, it has repeatedly been shown that humans and software do not behave in accordance with these assumptions. Humans are very adaptable and can change their behavior over time. Typically, this adaptation is not random and different humans behave differently. In the case of software, the lines of code are either programmed correctly or incorrectly but the modules will behave the same way, every time.

To further complicate the issue of determining exact probabilities, how does someone calculate a probability for a new technology or methodology that has never been used in a specific application before? Historical data may be irrelevant and while there may be a number of tests leading up to the final product, the pace of change in today’s world and focus on schedule and budget make it infeasible to completely test a new technology in a realistic environment.

After detailing how easily a seemingly simple task, such as calculating the probability for an event, can quickly become not so straightforward, it is important to discuss the difficulties of determining the exact value for the consequence of a loss. In some industries this is a very

straightforward endeavor. The financial industry works in different measures of currency that can all be directly converted to provide a comparable measure of units. On the other hand, a safety-focused industry like aerospace may think in terms of human injuries or death. How do you quantify the wide range of severity from no effect to death? Is losing the use of a right index finger in an accident considered the same level of severity for a right- and left-handed person? Someone from the finance department may argue that quantifying on an injury/death scale is too subjective and that we could quantify the injuries in terms of dollars. Interestingly enough, this type of scenario may actually reverse the severity levels of two accidents such that a fatality may have a single, lump-sum life insurance payout with a lawsuit settlement that is less expensive for a company than a significant lifetime disability claim for someone in their twenties with a lawsuit settlement.

Let us now suppose that someone was actually able to work through all of the difficulties to getting exact values for probability and consequence that were just mentioned and now just needed to make a decision between two different options. One of the options creates a one out of four chance of losing forty dollars and the other option creates a one out of one million chance to lose ten million dollars.

$$\begin{aligned}
 RiskE(Option1) &= Probability(Option1) \times Consequence(Option1) \\
 &= \frac{1}{4} \times 40 \\
 &= 10
 \end{aligned}$$

$$\begin{aligned}
 RiskE(Option2) &= Probability(Option2) \times Consequence(Option2) \\
 &= \frac{1}{1,000,000} \times 10,000,000 \\
 &= 10
 \end{aligned}$$

$$RiskE(Option1) = RiskE(Option2)$$

While a perfectly rational person would understand that the expected values are the same in either scenario, a person's risk perception will add another variable called subjective risk

to the decision-making.[17] The weighting of the subjective risk varies from person to person based on a number of factors, to include previous experience or cultural background. Even with perfect and objective information, people will likely make the final decision.

One popular tool used to organize risk data and help people make an objective decision is called the Risk Matrix. It is commonly known to help simplify and visualize the determined values of probability and severity for managerial decision-making. In addition to all of the potential issues previously discussed with the inputs to a risk matrix, there are also other characteristics of the matrix that create qualitative and mathematical flaws. One example is how we scale the axes of probability and severity for the risk matrix.

Using the risk matrix in Figure 3-1, consider an assessment with three different causes that need to be addressed, assuming the probabilities and severities are known with certainty.

RISK ASSESSMENT MATRIX				
SEVERITY PROBABILITY	Catastrophic (1)	Critical (2)	Marginal (3)	Negligible (4)
Frequent (A)	High	High	Serious Cause 1	Medium
Probable (B)	High	High	Serious	Medium
Occasional (C)	High	Cause 2 Serious	Medium	Low
Remote (D)	Serious Cause 3	Medium	Medium	Low
Improbable (E)	Medium	Medium	Medium	Low
Eliminated (F)	Eliminated			

Figure 3-1: Risk Matrix by DoD [11]

When the axes are scaled linearly on the interval $[0, 1]$, the three separate risks are calculated as follows,

$$Risk_i = Probability_i \times Severity_i$$

$$Risk_1 = 1 \times 0.25 = 0.25$$

$$Risk_2 = 0.5 \times 0.7 = 0.35$$

$$Risk_3 = 0.25 \times 1 = 0.25$$

However, with logarithmic axes on the intervals $[10^0, 10^8]$ and $[10^{-10}, 10^0]$ the risks are

$$Risk_i = Probability_i \times Severity_i$$

$$Risk_1 = 10^0 \times 10^3 = 10^3$$

$$Risk_2 = 10^{-5} \times 10^5 = 1$$

$$Risk_3 = 10^{-7} \times 10^8 = 10$$

All three causes receive the same risk ranking of “Serious” but are all qualitatively and quantitatively different. The first cause has a high probability of a relatively small loss, while the third cause has a low probability of a catastrophic loss. The second cause is quantitatively greater than the others when using linear axes, but is quantitatively less than the others when using logarithmic axes. Still, all three causes are considered to be within the same risk category. Is it always true that a 50% chance of a critical loss is actually worse than a certainty of a marginal loss? Or even a 25% chance of a catastrophic loss? Advocates defend the use of the traditional approaches by stating that the results are only meant to guide the decision making. But as discussed earlier, is this guide for decision makers even remotely accurate with the amount of variability and potential errors in the inputs? Due to the inherent amount of variability and uncertainty in the risk assessment matrix, even the largest amount of effort put into creating one can still provide inconsistent and misleading qualitative information, whether the users know it or not. [10]

Other experts[18] in the field of safety, who acknowledge the issues around determining a specific number for severity and probability and multiplying them together, view risk in

a more complex manner that encompasses hazard levels as a combination of severity and probability individually measured by the methods previously discussed. This group defines risk as:

Risk - Risk is the hazard level combined with the likelihood of the hazard leading to an accident and hazard exposure or duration.

This expanded definition has multiple components of risk, as depicted in Figure 3-2, and does not try to combine them with basic mathematical functions such as multiplication. While risk does involve the hazard level that most people traditionally refer to as the risk level, it also includes the relationship between the hazard and the accident and the duration that the hazard is present.[18]

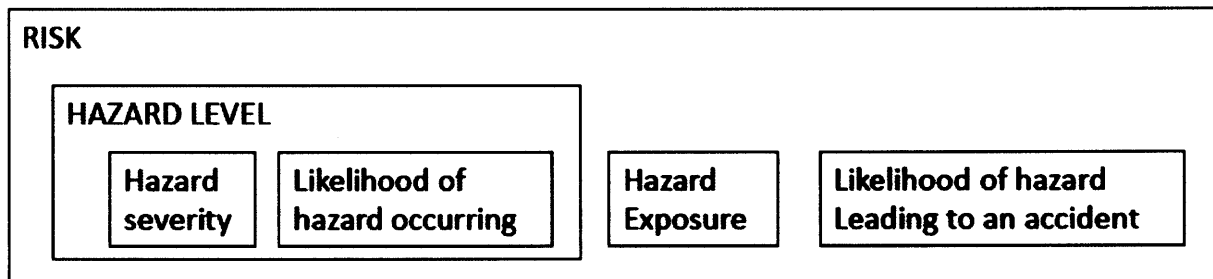


Figure 3-2: Components of Risk [18]

3.2 Safety

Traditionally, people relate risk and safety only in industries that are concerned with occupational health. While risk is a common term across industries, safety tends to only be associated with risk to human lives. Additionally, many people live under the assumption that safety is improved by increasing system or component reliability and that if components or systems do not fail, then accidents will not occur. However, safety and reliability are different properties and while this lack of distinction may have prevented accidents in electromechanical systems of the past, it is proving to be ineffective for modern socio-technical systems. While this may seem like an obvious statement, comparing the traditional defini-

tions of reliability and safety shows how reliability could have easily been used in place of safety depending on the context and complexity of the system.

Reliability - *the ability of a system or component to perform its required functions under stated conditions for a specified time*

Safety – *the condition of being protected against harmful conditions or events, or the control of hazards to reduce risk*

Consider walking under a grand piano that a moving company trying to move into a 5th story apartment by using a pulley system to lift the piano over the sidewalk and up to the apartment balcony. The system is made of entirely mechanical parts, such as rope that is attached to the piano with knots and run through a pulley that is attached to the top of the building or another object from a similar vantage point. The physics are pretty straightforward, movers pull down on the rope with more force than the 1,200 pound piano creates. As long as the force generated on any of the components does not exceed the maximum force allowed, the piano will make it up to the apartment and we will safely make it under the piano. In this scenario, if we wanted to increase the safety of the piano lift we might decide to get a stronger, more reliable rope. Now think about replacing the pulley that the rope is run through with a helicopter on a dark and stormy night. Could we just increase the reliability of the rope to increase the safety of the piano lift?

When systems require humans, software, and complex and tightly coupled sub-systems to operate in a dynamic environment, it becomes more apparent that humans may need to adapt, that software must have complete requirements, that component interactions are increasingly important to the safe operation of modern systems, and that conditions change.[20]

Additionally, traditional views of safety focus primarily on human injuries or deaths. Because of this view, the safety models and analysis techniques developed for a traditional view of safety do not take an extended view to include human losses, mission losses, equipment/material losses, and environmental losses. Due to the limited scope of the initial problem these techniques were developed to support, the potential effectiveness of traditional methodologies was limited from the beginning.

3.3 Accidents Models and Assumptions

The goal of all safety-related activities in an organization is ultimately to prevent accidents from occurring. In order to prevent accidents, we have to understand these accidents through a model that represents how the world works, which is our theory of accident causation. These accident models are critical as they provide the foundation for investigating and analyzing the cause of accidents, designing to prevent future losses, and assessing the risk associated with using the systems and products we create. However, models are simply representations of reality based on a set of underlying assumptions, which is why we must understand the assumptions to identify the model's strengths and weaknesses.

Traditional accident models' most common assumption is that accidents are caused by chains of directly related events and that we can understand accidents and assess risk by looking at the chain of failure-events leading to the loss.[20] Additionally, these models typically assume that there is an initial cause that sets the event chain in motion. Two explanations of accident causality based on this model are Heinrich's Domino Accident Model and Reason's Swiss Cheese Accident Model.

3.3.1 Domino Accident Model

Herbert Heinrich first introduced his Domino Accident Model in 1931 after reviewing numerous industrial accident reports.[16] This model depicts accident causality through a linear propagation of events, which he organized into the five stages depicted in Figure 3-3.

If one event in the sequence triggers, then the following event would trigger unless a domino could be removed. The first two dominos focus on the person involved in the accident, the third is the unsafe act or condition, the fourth is the resultant accident, which then results in the last domino triggering, which is the injury. Heinrich believed that nearly all accidents were caused by unsafe acts or unsafe conditions caused by human error.[16] Because humans are complex and difficult to control, the majority of his work was focused on removing the middle domino through things like standardized work processes. While the intent of this model was to understand injuries in an industrial environment, it received great popularity and was spread to nearly every industry.

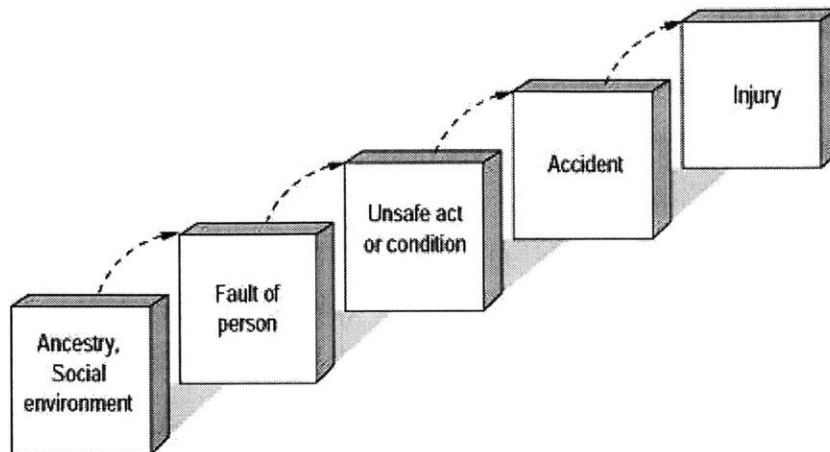


Figure 3-3: Heinrich's Domino Accident Model

3.3.2 Domino Model Extensions

Heinrich's Domino model was eventually revised twice in 1976, attempting to extend the causal factors to include the more systemic factors in an organization and the managerial decisions. The first was by Bird and Loftus, who settled on the following naming conventions for the five stages: [8]

1. Lack of control by management
2. Basic causes (personal and job factors)
3. Immediate causes (substandard practices/conditions/errors)
4. Accident or incident
5. Loss

Later Adams decided upon a slightly different representation of management by renaming the stages again: [1]

1. Management structure (objectives, organization, and operations)
2. Operational errors (management or supervisor behavior)
3. Tactical errors (caused by employee behavior and work conditions)

4. Accident or incident
5. Injury or damage to persons or property

Although these proposals both expand the original focus of the Domino Accident Model from ancestry and social factors to organizations and the role of management in accidents, they continue to assume a root cause and explain accidents in terms of multiple events sequenced as a forward chain over time.

3.3.3 Swiss Cheese Accident Model

In 1990, James Reason introduced the Swiss Cheese Accident Model as a new model of accident causation.[28] However, while this model uses the analogy of Swiss cheese rather than dominos, it continues to be based upon a chain of failure events model. The Swiss Cheese Accident Model depicts accidents as active and latent failures that result in failed or absent defenses, varying randomly in size and location, within the accident protection barriers. These protection barriers are organized into four groups, which include organizational influences, unsafe supervision, preconditions for unsafe acts, and unsafe acts, illustrated in Figure 3-4.

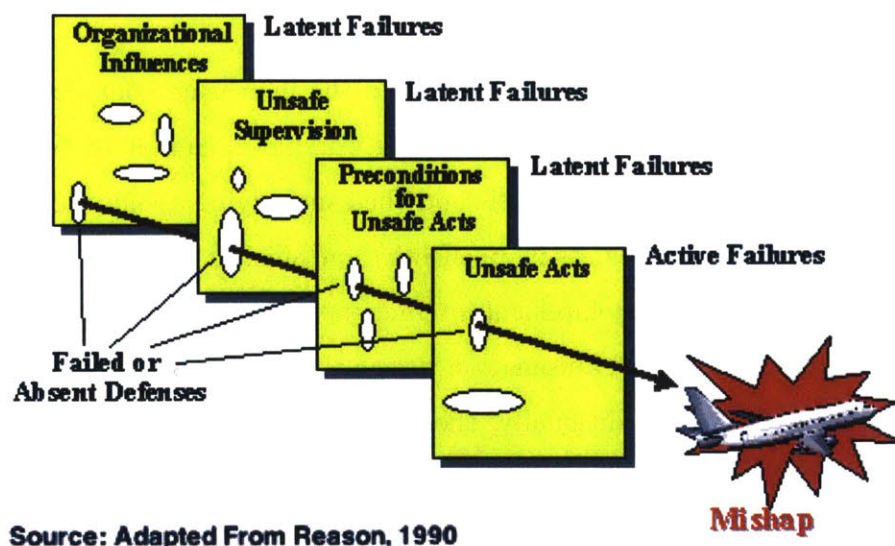


Figure 3-4: Reason's Swiss Cheese Accident Model

While the protective layers are in place to prevent an accident or mishap, eventually the holes may reach an alignment that allows a trajectory through all the slices of cheese. This situation represents a sequence of failures through several layers of defense that ultimately cause an accident. The first three layers are considered to have latent failures, which are typically designer, high-level decision maker, manager, and maintenance staff decisions or actions lying dormant in the system well before the accident. The last layer represents where the active failures occur, leading immediately to the accident.[28] These failures are typically found to be the real time errors or failures of the operational controllers, leading the model to focus again on human error. The model also contains two new assumptions that include random behavior of components and independence between failures in each layer. However, some processes may create holes in every layer and many important systemic factors can simultaneously affect the behavior of the components and layers of this model.

The Swiss Cheese Accident Model was introduced as a new and improved accident model, as it expands the sphere of influence for the accident and includes new assumptions. However, the Swiss Cheese and Domino Models are actually the same model using different analogies to explain the classic chain of events model of accident causation. Underlying assumptions result in looking for multiple events sequenced as a forward chain over time, discounting systemic factors that could impact randomness and independence, searching for a root cause, and blaming accidents on some type of failure or human error.

The advancing rate of technological change and the increasing complexity of modern, socio-technical systems, such as the manufacturing workplace, cannot be fully understood through these previously acceptable models and their underlying assumptions. In order to understand hazards and manage safety in an effective and efficient manner, a new approach is needed that considers hazardous component interactions, software requirement flaws, human complacency or confusion due to automation design, inappropriate human behavior, and traditional hardware faults. Additionally, the approach must be applicable from system conceptualization so that safety-related requirements can be identified and used to mitigate or eliminate hazards during design.

3.3.4 Systems-Theoretic Accident Model and Processes

Nancy Leveson introduced Systems-Theoretic Accident Model and Processes (STAMP) in 2002.[20, 19] It was developed to capture more types of causal factors in accidents, including social and organizational structures, new kinds of human error, design and requirements flaws, and dysfunctional interactions among non-failed components. In order to achieve this goal, STAMP defines safety in an absolute sense:

Safety is freedom from accidents or losses. [18]

Additionally, rather than treating safety as a failure problem or simplifying the cause of accidents as a linear chain of events, STAMP treats safety as an emergent property of the system. This seemingly small but powerful difference turns safety into a control problem in which accidents arise from complex, dynamic processes that may operate concurrently and interact to create unsafe situations.

Under this view of safety, accidents can then be prevented by identifying and enforcing constraints on component interactions. This model not only captures accidents due to component failures, but also explains increasingly common component interaction accidents that occur in complex systems without any component failures. For example, software can create unsafe situations by behaving exactly as instructed or operators and automated controllers can individually perform as intended but together may create unexpected or dangerous conditions.

The foundations of the STAMP model of accident causation are based on three basic concepts — hierarchical safety control structures, safety constraints, and process models — along with basic concepts from systems and control theory. Systems theory is based upon two pairs of primary concepts that include emergence and hierarchy and communication and control. The first concept is that although the focus is on the system as a whole and not on the parts taken separately, a model of a complex system can be built as a hierarchy of levels of organization where each level is characterized by having emergent properties. Each level's emergent properties are those system properties that arise from the interactions among components and do not exist at lower levels. Safety is a type of emergent property for systems, as it can only be determined in the context of the whole, as illustrated in Figure

3-5.

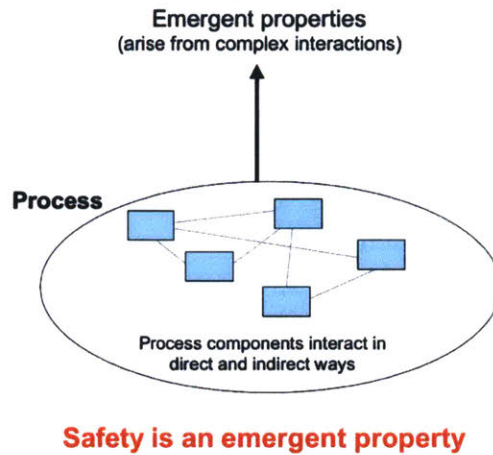


Figure 3-5: Emergent Properties

The second concept comes from control theory, incorporating the concepts of communication and control into systems theory. Within the hierarchical system, control is imposed on lower levels in the form of constraints. Additionally, because STAMP includes the environment in which the systems operate, they are considered to be open systems and require communication. In control theory, open systems are viewed as interrelated components that are kept in a state of dynamic equilibrium by feedback loops of information and control. Therefore, while higher levels of the hierarchy issue control actions to lower levels, lower levels must provide feedback to the higher levels as depicted in Figure 3-6.

Through these ideas, system safety can be viewed as a system control problem rather than a component reliability problem. Accidents occur due to violating required safety constraints on component behavior, such as when component failures, external disturbances, and/or potentially unsafe interactions among system components are not handled adequately or controlled. System controls may be managerial, organizational, physical, or operational, etc., which allows for more flexible and effective solutions that can be tailored to the specific application. An example of a general safety control structure for a socio-technical system is depicted in Figure 3-7.

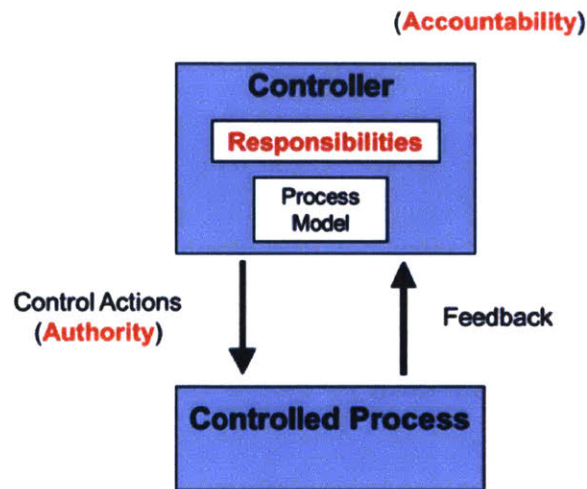


Figure 3-6: Example Control Loop

In order to develop effective safety constraints, STAMP defines four types of unsafe control actions that must be eliminated or controlled to prevent accidents:

1. A control action required for safety is not provided or is not followed
2. An unsafe control action is provided that leads to a hazard
3. A potentially safe control action is provided too late, too early, or out of sequence and leads to a hazard
4. A safe control action is stopped too soon or applied too long, leading to a hazard

One common potential cause of a hazardous control action in STAMP is an inadequate process model used by human or automated controllers. A process model contains the controller's understanding of the current state of the controlled process, the desired state of the controlled process, and the ways the process can change state. This model is used by the controller to determine what control actions are needed. In software, the process model is usually implemented in variables and embedded in the program algorithms. For humans, the process model is often called the "mental model." [19] Software and human errors frequently result from incorrect process models. Accidents like this can occur when an incorrect or incomplete process model causes a controller to provide control actions that are hazardous. While process model flaws are not the only cause of accidents in STAMP, it is a major contributor.

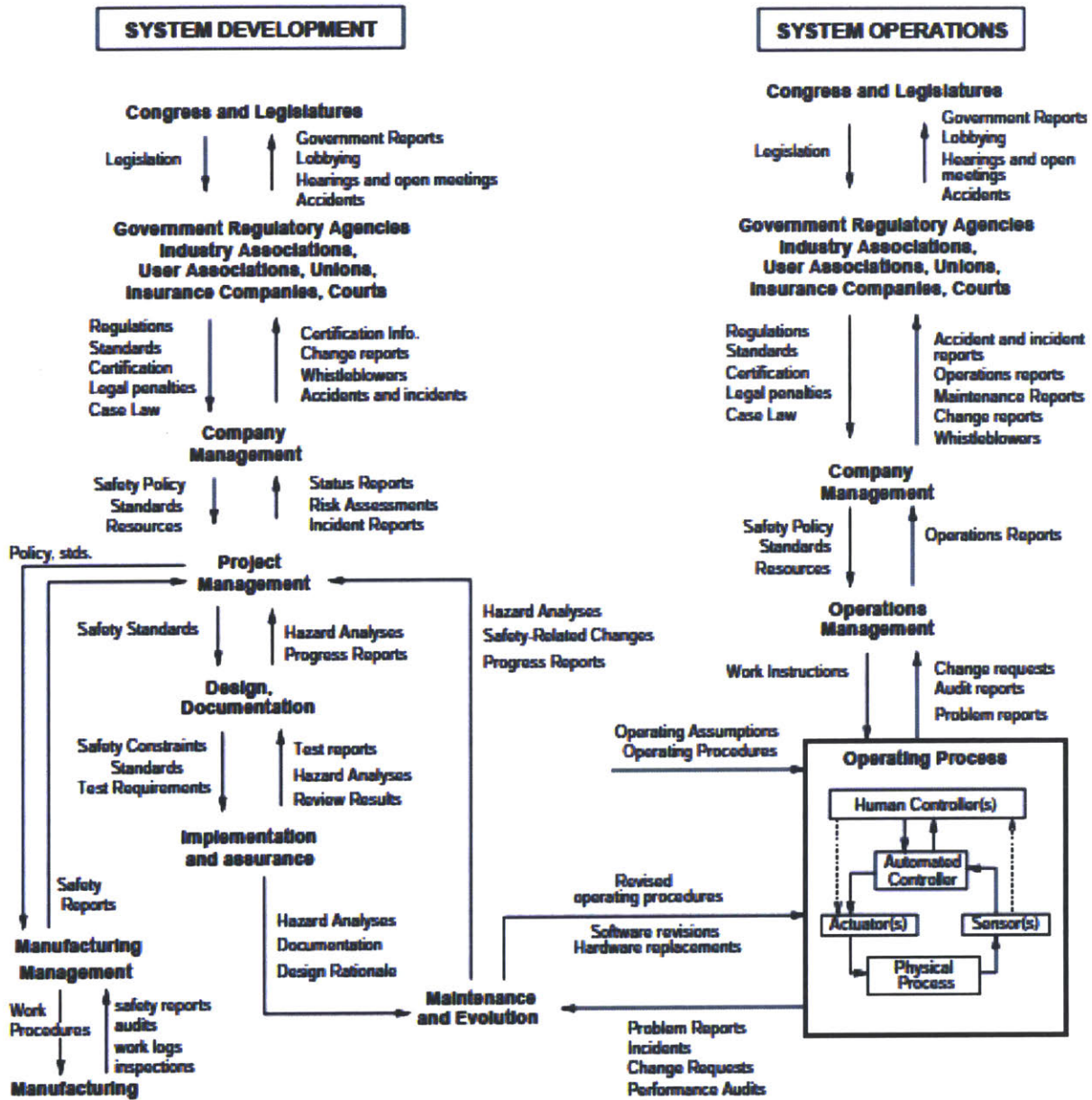


Figure 3-7: General Form of a Model of Socio-Technical Control

Using STAMP, accidents can be understood by identifying the safety constraints that were violated and determining why the controls were inadequate in enforcing them. Additionally, STAMP not only allows for the consideration of more accident causes than simple component failures, but also allows for more sophisticated analysis of failures and component failure accidents. Instead of simply determining that the reliability of a component needs to be increased, the failure may have been a result of inadequate constraints on the manufacturing process, inadequate engineering design such as missing or incorrectly implemented fault tolerance, lack of correspondence between individual component capacity and task requirements, unhandled environmental disturbances, inadequate maintenance, physical degradation, and/or many more possible considerations.

By developing a thorough understanding of the entire socio-technical system, STAMP enables the identification of more causes of accidents, to include component failure accidents, unsafe interactions among components, complex human and software behavior, design errors, and flawed requirements. Once these are identified through initial analysis or monitoring of the dynamic system, a number of appropriate controls can be developed in relation to the other constraints and circumstances for the system. This flexibility ensures that the levels of effectiveness and efficiency required for the specific application can be achieved.

3.4 Analysis Methods

Hazard analysis, or causal analysis, is the process of investigating an accident before it happens. Although hazard analysis does not ensure safety alone, the goal is to identify the causes of accidents so that we can eliminate or control them during system design and operation. While simply identifying hazards in a system may allow safeguards to be applied, the more information known about the hazard and its causes allows for more effective and less expensive specific controls to be put in place. In order to accomplish the analysis, the people involved must have an accident model and system design model to work from. Due to this necessity, the hazard analysis methods are subject to the same capabilities and limitations of their underlying accident models.

3.4.1 Failure Modes and Effect Analysis

Failure Modes and Effect Analysis is an analysis method that systematically evaluates how individual component or subsystem failures will affect the overall system. It was developed and introduced around 1950 by reliability engineers in defense industries.[27] Due to its effectiveness for the complex systems of the time, it was adapted by the overall aerospace industry and eventually spread to automotive, oil and gas, food, drug, and cosmetic industries.[12]

FMEA generally follows an inductive, bottom-up methodology and starts with identifying the components of a system. Next, the mechanisms by which a component may fail to achieve its designed functions, or failure modes, are determined. Then the potential causes and effects on the system are investigated. A closely related methodology called Failure Modes and Effect Criticality Analysis (FMECA) uses the same process but includes determining a criticality of each failure mode by examining the severity of each effect. An example of a FMECA worksheet is shown below that can be used to summarize the analysis.

Component	Failure Mode	Cause	Effect	Severity	Probability of Occurrence	Criticality

Table 3.3: Example FMECA Worksheet

Assuming the FMECA is completed with full and accurate information, the results of the analysis may be useful to make decisions about the physical system. Failures with high criticality are identified to be studied further and possibly eliminated from the system. This method is capable of identifying single point failures in a system and characterizing their effects. The method is simple and can be applied to a broad spectrum of systems. However, its effectiveness is limited in today's complex systems due to its underlying assumption of a linear progression of events. This results in an inability to assess scenarios with nonlinear and feedback relationships, scenarios with multiple failures, and scenarios of component interaction with no failures, which typically include humans, software, the environment, and

more. FMECA is also a very time-consuming process and results in wasted effort assessing component failures that do not lead to an accident.

3.4.2 Fault Tree Analysis

Fault Tree Analysis (FTA) is an analysis method that identifies the causes of undesired events. This method was developed at Bell Laboratories in 1961 to evaluate the Minuteman missile system for the U.S. Air Force.[34] Instead of starting with the components, like in the bottom-up approach of FMEA, FTA is considered a deductive, top-down methodology that begins with selecting the undesired event. After selecting the undesired event, the method continues in a top-down manner to identify the causal events that could lead to the top event. These events are placed as branches underneath the undesired event. FTA can be either qualitative or quantitative in that the results of the analysis could be a list of different combinations or a probability of the undesired event's occurrence.[33] As example of a fault tree is shown in Figure 3-8 to illustrate the methodology.

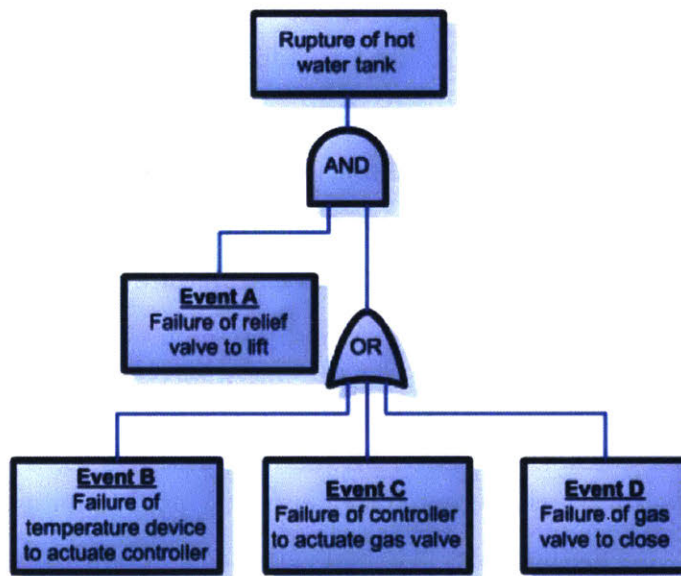


Figure 3-8: Example Fault Tree from the original Bell Laboratory study[33]

By using a top-down approach, Boolean Logic, and a standardized set of symbols to build a tree of faults and failures that can be combined in any combination, FTA has proven to be

an effective tool for simpler systems. However, little guidance is provided on how to find the faults and failures in a fault tree. Additionally, like FMEA/FMECA, this method assumes a failure chain of events and is subject to the same limitations. This method is not as capable of analyzing systems with significant software, such as robotics. As previously mentioned, software does not fail like a physical component and FTA does not address software requirements established earlier in the system lifecycle. Although FTA addresses some of the issues presented with FMEA/FMECA, it still has many limitations in effectiveness and efficiency.

3.4.3 Job Hazard Analysis

Job Hazard Analysis (JHA) is an analysis method that focuses on job tasks as a way to identify hazards before they result in an accident. It is also commonly referred to as Job Safety Analysis (JSA), Activity Hazard Analysis (AHA), and Task Hazard Analysis (THA). Its origins are traced back to the 1920s, during the beginning of the scientific management practice of job analysis.[14] In its initial development, job analysis was being used to analyze streetcar operations.[9] The belief was that job analysis should identify hazards of the operations so that standard procedures could be established in order to avoid accidents.[32]

JHA focuses on the relationship between the worker, the task, the tools, and the work environment. This is accomplished by breaking down a job into its component steps and then evaluating each step for potential or known hazards.[30] Controls are then determined for each hazard. Additionally, users can also assess each hazard identified for a risk level in which the initial score is the inherent risk level and the score with the determined controls in place is the residual risk level. A sample Activity Hazard Analysis worksheet is shown in Figure 3-9.

Job Hazard analysis is a widely-accepted tool currently used in industry to assess workplace safety. However, this methodology has little guidance on how to properly perform an analysis and has a very informal structure and lack of standardized process for continuous learning and improvement. Additionally, JHA appears to approach the analysis from both a top-down and bottom-up methodology, depending on the step in the process, with no systematic structure to ensure a thorough analysis. It also recommends repeatedly asking “what if”[30] throughout the analysis, constantly expanding the bounds of the analysis that were

Activity/Work Task:		Overall Risk Assessment Code (RAC) (Use highest code)					
Project Location:		Risk Assessment Code (RAC) Matrix					
Contract Number:		Severity		Probability			
Date Prepared:				Frequent	Likely	Occasional	
Prepared by (Name/Title):		Catastrophic	E	E	H	Seldom	
Reviewed by (Name/Title):		Critical	C	H	H	M	
Notes: (Add Notes, Review comments, etc.)		Marginal	M	M	L	L	
		Negligible	N	L	L	L	
		Step 1: Review each "Hazard" with identified safety "Controls" and determine RAC (See above) "Probability" is the likelihood to cause an incident, near miss, or accident and identified as: Frequent, Likely, Occasional, Seldom or Unlikely. "Severity" is the outcome/degree if an incident, near miss, or accident did occur and identified as: Catastrophic, Critical, Marginal, or Negligible. Step 2: Identify the RAC (Probability/Severity) as E, H, M, or L for each "Hazard" on AHA. Annotate the overall highest RAC at the top of AHA.				RAC Chart E = Extremely High Risk H = High Risk M = Moderate Risk L = Low Risk	
Job Steps		Hazards		Controls		RAC	
Equipment to be Used		Training Requirements/Competent or Qualified Personnel name(s)		Inspection Requirements			

Figure 3-9: Example Job/Activity Hazard Analysis Form

never established in the first place. Most information on achieving a quality analysis through the methodology relies upon having a great amount of previous experience. Although this methodology has the potential to provide a more systems-like approach to hazard analysis, it lacks the structure to promote learning and improvement and the technical depth to understand complex engineering systems that are being integrated into the workplaces of today. Finally, the assignment of risk is totally arbitrary and often incorrect.

3.4.4 Systems-Theoretic Process Analysis

Systems-Theoretic Process Analysis (STPA) is an analysis method based on the STAMP accident causation model. STPA provides a systematic approach to building a model of the system's safety control structure and to analyzing the model. The analysis identifies present and potential issues with the system's ability to provide adequate control or enforce constraints on safety-related behavior at each level of the system development and system operations control structures. While it is a detailed and systematic process, its basis in systems theory enables it to be applied to any application, limited only by the user's discretion on scope and granularity.

Defining Accidents (Losses)

The analysis begins by defining the accidents and unacceptable losses to be considered for the system. Because the analysis is rooted in systems theory, the term accident takes on a more general definition:

***Accident:** An undesired or unplanned event that results in a loss, including loss of human life or human injury, property damage, environmental pollution, mission loss, etc.*

This definition of an accident allows for the inclusion of any loss undesired by the system's stakeholders, which may include the operators, surrounding workers, management, company investors, local population, etc. Some examples of losses may include death or injury to an operator or another person, major equipment damage, environmental pollution, or non-accomplishment of the mission.

Defining System Hazards and System Boundaries

The next two steps involve identifying the system-level hazards and defining the system boundaries for use to develop the system's safety constraints. In system safety, a hazard is considered to be within the system being designed and not just in the environment. More specifically, STAMP defines the term hazard as follows:

Hazard: *A system state or set of conditions that, together with a particular set of worst-case environmental condition, will lead to an accident (loss).*

To illustrate this definition, consider a self-driving vehicle. In the traditional thinking, a hazard would be considered something in the environment of the system, such as a guardrail along the freeway, another vehicle on the road, or a person crossing the street. However, the above definition would define the hazard as the vehicle not maintaining minimum separation from those surrounding objects.

$$\text{Hazard} + \text{Environmental Condition} \Rightarrow \text{Accident(Loss)}$$

The ability to define the system boundaries is very helpful in that they can be defined to include only conditions related to the accident over which the system designer has control. An accident may involve aspects of the environment that the engineers have no control over. By defining boundaries around the system hazards, engineers can design systems to eliminate or control hazards by enforcing safety constraints and preventing accidents.

Modeling a Safety Control Structure

Modeling how the system is enforcing safety constraints provides numerous benefits, such as promoting a consistent view of the system, using a repeatable process to update the model, and supporting the generation, design, analysis, and verification of requirements. [29] Continuing with the three basic concepts from STAMP — hierarchical safety control structures, safety constraints, and process models — the system can be modeled and analyzed for its ability to enforce the determined safety constraints.

In order to develop a control structure, it is important to review the four conditions required for process control: [20, 4]

1. *Goal* condition: the controller must have a goal or goals;
2. *Action* condition: the controller must be able to affect the state of the system, typically by means of an actuator or actuators;
3. *Model* condition: the controller must contain a model of the system; and

4. *Observability* condition: the controller must be able to ascertain the state of the system, typically by feedback from a sensor.

Additionally, a typical control loop is shown in Figure 3-9 to discuss the basic building block of the control structure. Within each control loop, the controller is responsible to maintain set points (or the goal condition) by providing control actions to manipulate controlled variables for the controlled process. The controller also receives information about (observes) the state of the controlled process through measured variables (feedback) that update the controller process model that is used by the control algorithm to generate the next control action. [20]

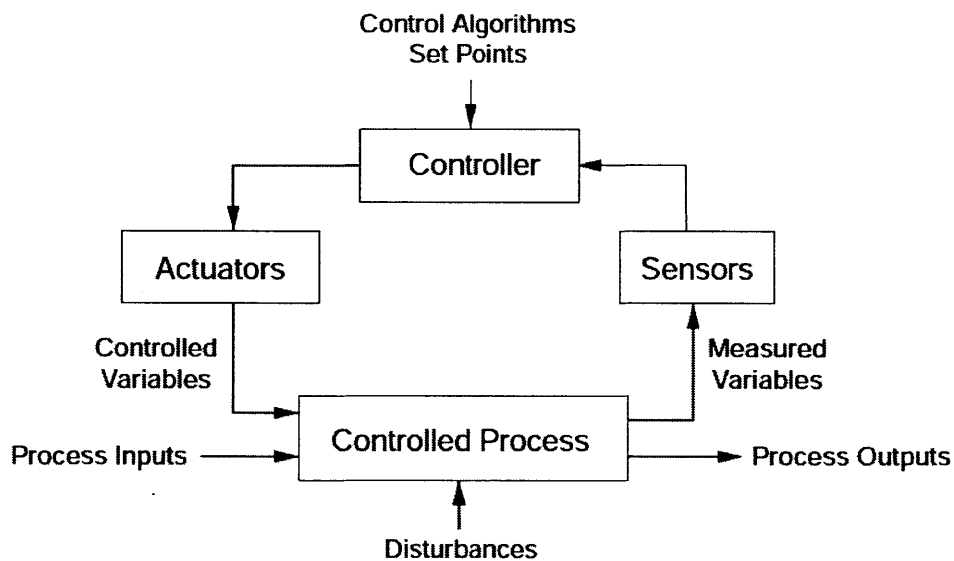


Figure 3-10: Standard Control Loop

It is relatively easy to discuss a single control loop in isolation, however the process of modeling a complex, socio-technical system requires a more detailed explanation of how to properly identify and relate the different components of a functional control structure. The following descriptions provide a more detailed explanation of the characteristics for each component of a control loop to consider when analyzing a system: [13]

The Controller:

- creates, generates, or modifies control actions based on algorithm or procedure and perceived model of system

- processes inputs from sensors to form and update process model
- processes inputs from external sources to form and update process model
- transmits instructions or status to other controllers or entities in the system

The Actuator:

- translates controller-generated action into process-specific instruction, force, heat, torque, or other mechanism

The Controlled Process:

- interacts with environment via forces, heat transfer, chemical reactions, or other input
- translates higher level control actions into control actions directed at lower level processes (if it is not at the bottom of a control hierarchy)

The Sensor:

- transmits continuous dynamic state measurements to controller (i.e., measures the behavior of controlled process via continuous or semi-continuous, digital data)
- transmits binary or discretized state data to controller (i.e., measures behavior of process relative to thresholds; e.g., sensor has algorithm built-in to determine a threshold but has no control authority)
- synthesizes and integrates measurement data (i.e., takes location data from different types of sensors to create an estimate, like a Kalman filter)

These building blocks of the control structure are organized in a hierarchical system, where each level imposes constraints on the level below and provides feedback to the level above while the components are interacting with their environment. A basic depiction of a hierarchical system is provided in Figure 3-11, however as systems increase in complexity the model is rarely this linear vertically. Multiple controllers of different processes may need to pass information back and forth about the state of different process variables to update the other controller's process model in order to control each controller's respective process. Additionally, the model and analysis can include the socio-technical factors present in modern complex systems, as previously discussed and shown in Figure 3-7.

In order to depict and organize the diagram of the functional control structure, a standard control structure diagramming convention can be used like the one shown in Figure 3-12.

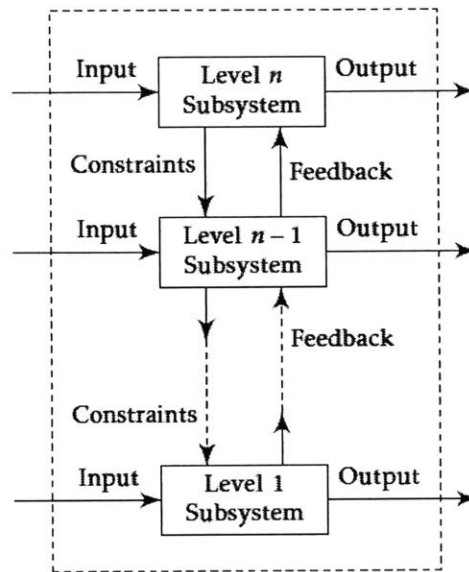


Figure 3-11: Basic Depiction of a Hierarchical System[22]

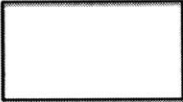



Symbol	Description
	Functional Element, such as a Controller, Actuator, Controlled Process, or Actuator
	Control Action
	Control Feedback
	Physical and/or Informational Interaction, other than Control Action and Control Feedback

Figure 3-12: Standard Control Structure Diagramming Convention[7]

STPA Step 1

Once the system's accidents and related system hazards are identified, along with the associated safety constraints to control, and a safety control structure is determined, STPA Step 1 can begin. This step identifies the potential for inadequate control that leads to a hazard by assessing the safety controls in the system design. In order to do this, the analysis analyzes each control action identified in the development of the safety control structure using four general scenarios:

1. A control action required for safety is not provided or is not followed
2. An unsafe control action is provided that leads to a hazard
3. A potentially safe control action is provided too late, too early, or out of sequence
4. A safe control action is stopped too soon or applied too long

As depicted in Table 3.4, a simple table can be used to capture the results of this portion of the analysis. These control actions determined to be potentially unsafe are then used to create new system and component safety requirements and constraints.

Control Action	Not Providing Causes Hazard	Providing Causes Hazard	Incorrect Timing/Order	Stopped Too Soon / Applied Too Long

Table 3.4: Control Action Table

STPA Step 2

The next step in STPA is designed to determine how each unsafe control action can occur by examining each control loop in the safety control structure. Figure 3-13 shows a generic control loop with many of the causal factors to be considered during this step of the analysis.

By viewing the world of safety through concepts from systems- and control-theory, STPA is able to identify accident causes well beyond traditional component failures. Taking broader views of accidents, defining hazards and system boundaries in ways that keep system designers in control, and systematically analyzing the defined safety control structure, gives

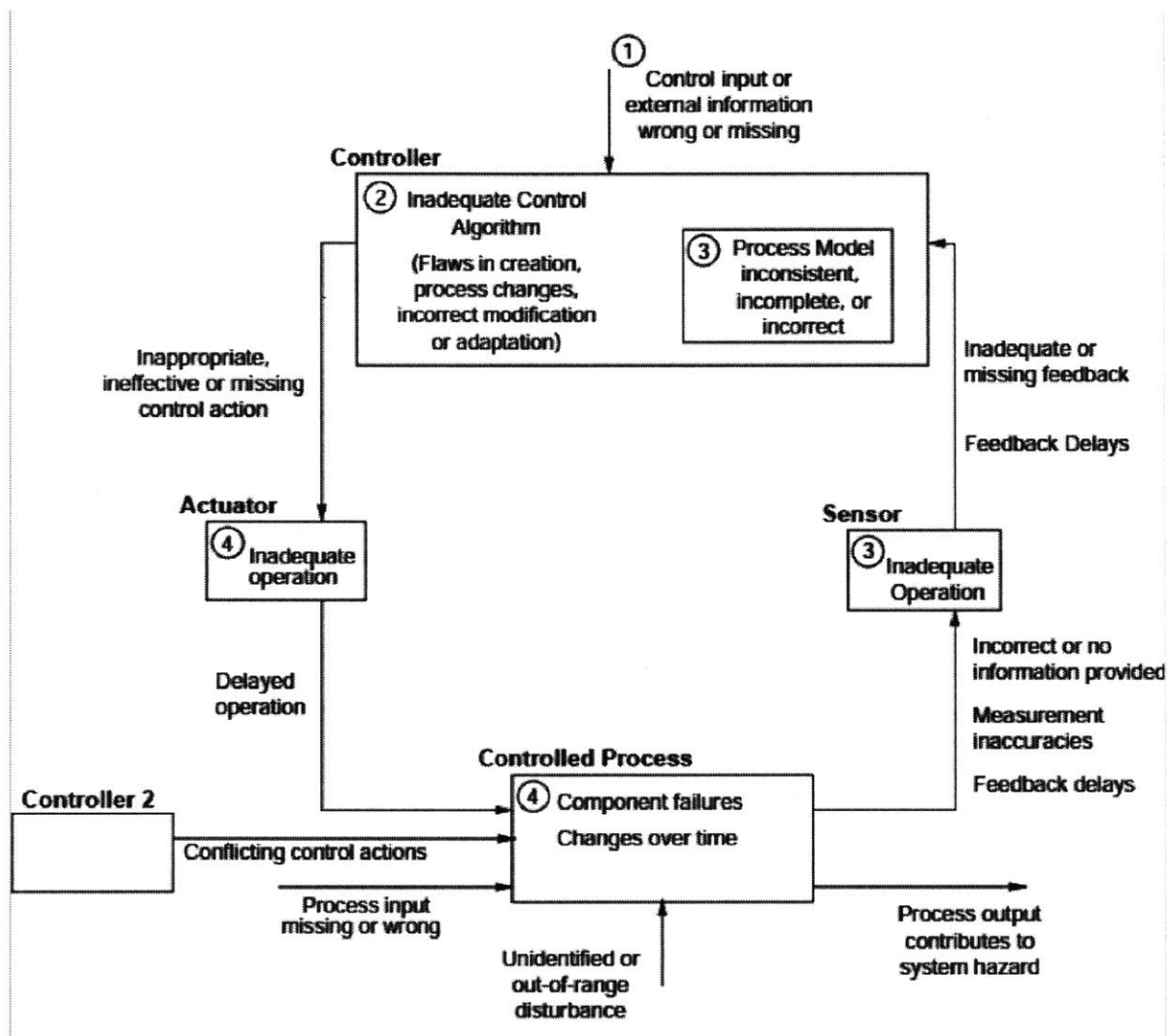


Figure 3-13: General Control Loop with Causal Factors

this analysis technique the ability to identify requirement flaws, design errors, organizational factors, and complex human behavior in addition to component failures. It also enforces a deep understanding of the system prior to determining solutions to any identified issues in the system and provides the flexibility to choose the best solution for the specific application. While it is a relatively new method in comparison to the previously discussed traditional techniques, it is based on an entirely new and coherent foundation of assumptions to handle the issues facing today's increasingly complex, socio-technical systems.

THIS PAGE INTENTIONALLY LEFT BLANK

Chapter 4

Case Study of a Semi-Automated Manufacturing Process

This chapter begins by providing an overview of the case study used for the assessment process analysis. It then provides background on the company's current safety risk assessment process, details the process in use in the case study, and discusses the results of the analysis. The chapter continues by detailing the implementation of the proposed analysis process on the case study, discussing the results of the analysis, and comparing the proposed analysis results to the traditional analysis results.

4.1 Overview of Manufacturing Process

At a very high and abstract level, the process begins with sub-components for the aircraft entering the factory on the back of a vehicle operated by a worker. Within the factory, two more workers are each responsible for driving an automated ground vehicle (AGV) by hand-held control from an AGV docking and charging location within the factory to connection points on a Product Transportation Vehicle (PTV) for the incoming sub-component. After connecting to the PTV, the two drivers must collaboratively drive the combined Product Transportation System (PTS) around the factory to the location of the sub-component on the first mentioned vehicle. At this point, another group of workers will then work together to move the sub-component from the vehicle to the PTS. The AGV operators will then drive

the PTS through the factory to a location at which the sub-component can be prepared for entry into an automated manufacturing cell. Inside the cell are multiple robots and associated support equipment that will perform the automated process of drilling holes in the sub-component and filling the holes with additional components. When the sub-component, the PTS, the robotic cell, and the operators are all ready, the cell operator and controller will assume control of the PTS and move it through the automated process to perform its work statement. During this time, there will also be concurrent work occurring to portions of the sub-component that are not inside the robotic cell perimeter. Once the entire automated process is complete, workers will re-assume control of the PTS and move it down the factory line to the next position for more manual work processes.

4.2 Current Analysis Process - SRA

The company conducts a Safety Risk Assessment (SRA) during the implementation of new, complex automated equipment. It is intended to identify the safety risks being introduced into the factory workplace and determine ways to mitigate the risks. The company currently conducts its Safety Risk Assessment using the methodology recommended by the American National Standards Institute (ANSI).

4.2.1 ANSI and American National Standards

ANSI serves as the administrator and coordinator of the United States private sector voluntary standardization system. It was originally founded in 1918 by five engineering societies and three government agencies and remains a private, nonprofit organization. ANSI facilitates the development of American National Standards (ANS) by accrediting the procedures of standards developing organizations (SDOs). These U.S. groups work cooperatively to develop voluntary national consensus standards that enhance global competitiveness and American quality of life. ANSI also represents the United States interests in the international community when working with other global organizations such as the International Organization for Standardization (ISO) and International Electromechanical Commission (IEC). In this capacity it ensures that American standards are current with the interna-

tional community and promotes the use of U.S. standards internationally. [3]

4.2.2 ANSI Standard Analysis Process

In the case of the specified semi-automated manufacturing process, the company primarily uses the standards developed by the Robotic Industries Association in collaboration with a number of partner organizations. In the document titled “American National Standard for Industrial Robots and Robot Systems – Safety Requirements” (ANSI/RIA R15.06-2012) a common list of voluntary safety requirements is detailed for both the specific industrial robot and the industrial robot systems and integration. The document clearly specifies “the integrator shall perform a risk assessment to determine the risk reduction measures required to adequately reduce the risks presented by the integrated application” and seems to take a stance on using a task-based risk assessment. However, the introduction section of ANSI/RIA R15.06-2012 names a recommended methodology to meeting this requirement, called the Task-based Risk Assessment Methodology. [5]

4.2.3 Task-based Risk Assessment Methodology

The risk assessment methodology outlined by the Robotic Industry Association technical report was developed to identify “hazards associated with tasks that operators would foreseeably have to conduct around industrial robots, as well as related machinery and equipment.” [5] The method looks to combine aspects of bottom-up and top-down approaches, such as FMEA and FTA respectively, for an engineering system with a task based approach, similar to job hazard analysis. The Task-based Risk Assessment procedural document also emphasizes the inclusion of risk considerations beyond its operational, maintenance, and troubleshooting tasks to its entire lifecycle as prescribed in Table 4.1.

Pre-operations phase	Operation phase	Post-operation phase
Initial concept	Set-up	Decommissioning
Design	Start-up	Disposal
Fabrication	Use and production	
Installation	Troubleshooting	
Commissioning	Fault recovery	
	Maintenance	
	Repair	

Table 4.1: Lifecycle Phases IAW RIA TR R15.306-2014

The assessment is intended to be conducted with a team of qualified personnel that is representative of the system integrator and the user. This team should involve personnel from system integration, operations, maintenance, and engineering and “be led by an individual who is familiar with industrial robot and other machine tool applications and who has experience in the risk assessment process.” Once the team is assembled, the team leader begins the eight-step process, illustrated in Figure 4-1.

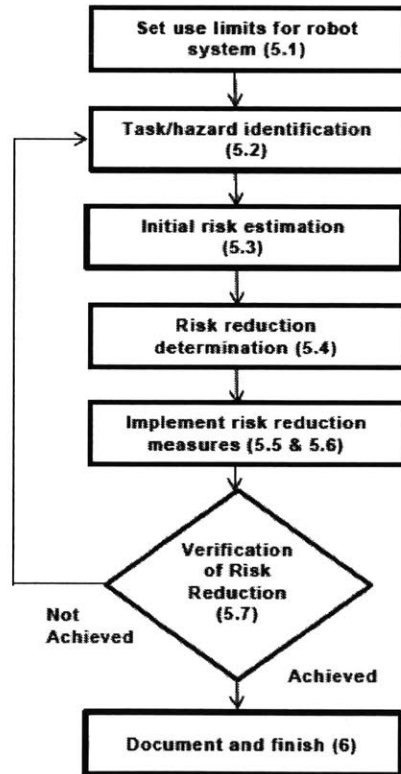


Figure 4-1: Task-based Risk Assessment Flowchart [6]

The first step involves considering the system as a whole, specifying its intended use and determining the limits of the system that will be involved in the risk assessment process. Within the process document, the limits of the robot system are listed as four broad categories with a few examples of each, while the robotics document expands the four categories even further.

From the Task-based Risk Assessment procedural document, RIA TR R15.306-2014, the four limit categories are [6]:

1. Use limits, including the anticipated functionality of the system
2. Space limits, including machine movement range, space needed for personnel access, etc.
3. Time limits, including life limit of the machinery or components, process flows, and recommended service intervals, and
4. Other limits, such as environmental, cleanliness requirements, hazardous conditions, etc.

An example of the more detailed description of one limit category as provided in the Robotics Safety Requirements document, ANSI/RIA 15.05-2012, for Space limits: [5]

1. required machine movement range
2. required space for installation and maintenance
3. required space for operator tasks and other human intervention
4. reconfiguration capabilities
5. required access
6. foundations
7. required space for supply and disposal devices or equipment

Establishing the overall system limits in the beginning of the assessment process helps to ensure the team understands the operations of the system and the different constraints on the design when identifying hazards and selecting mitigations.

During the next step of the assessment, the team begins determining all of the possible tasks to assess. This is done by brainstorming all of the foreseeable tasks associated with

the system, within each applicable stage of development (Table 4.1). The standard provides the following list of tasks to consider as a possible starting point: [5]

1. process control and monitoring;
2. workpiece loading;
3. programming and verification;
4. brief operator intervention not requiring disassembly;
5. set-up (e.g., fixture changes, tool change);
6. troubleshooting;
7. correction of malfunction(s) (e.g., equipment jams, dropped parts, event recovery and abnormal conditions);
8. control of hazardous energy (including fixtures, clamps, turntables, and other equipment);
9. maintenance and repair;
10. equipment cleaning.

Once the team agrees that the task list for the system is complete, the team looks at each task and brainstorms all of the possible hazards involved. Additionally, the team is supposed to identify any hazards that are not related to any particular task. According to the ANSI and the Robotic Industry Association standard, a hazard is defined as:

Hazard: *potential source of harm, or physical injury or damage to health*

The procedural document also provides a list of hazard types and examples for robotics and automation as a reference for the assessors to use while considering each task. Examples of the hazards that are to be considered are shown in Table 4.2.

Type	Example Hazard
Mechanical	Crushing, Pinching, Impact, Entrapment, Stabbing
Electrical	Electrocution, Short, Arcing, Sparking, Shock, Burn
Thermal	Burn (hot or cold), Radiation Injury
Noise	Loss of hearing, Loss of balance, Loss of awareness/disorientation
Vibration	Fatigue, Neurological damage, Vascular disorder, Impact
Radiation	Burn, Damage to eyes and skin, Related illnesses
Material/Substance	Sensitization, Fire, Chemical burn, Inhalation illness
Ergonomic	Unhealthy posture, Excessive effort, Repetitive strain, Fatigue
Environment	Burn, Disease or illness, Slipping, Falling, Respiratory damage, Impact
Combination	Any other consequence of combinations of hazards and hazardous situations

Table 4.2: Example Hazards from ANSI/RIA Standard

After all of the task and hazard combinations that the team can brainstorm have been identified, the team begins the process of assessing the initial risk for each. The risk levels are determined by the three factors of severity, exposure, and avoidance. In the first iteration of the assessment, the team determines the initial risk by viewing the task and hazard without any of the system's risk mitigations or safeguards in place. The team discusses the hazard in terms of the three factors and their respective rating criteria and once the rating for each factor is agreed upon, the team can continue to the next hazard. The risk assessment procedural guide describes severity, exposure, and avoidance as: [6]

Injury severity: *Injury severity is a function of the degree of estimated injury the operator would incur if the hazard were to come into physical contact with the operator while the operator is performing the associated task.*

Exposure: *Exposure to a hazard is a function of the estimated incidence of an operator being in the presence of the hazard, and takes into consideration the following:*

- *How frequently the operator would need to enter the hazard zone*
- *The duration in which the operator would need to stay in the hazard zone after entering*
- *Whether the task is routine or non-routine or other task frequency considerations.*

Avoidance: *Avoidance of a hazard is an assessment of the operator's ability to sense and elude a hazardous situation.*

Along with the definitions for each category, the team selects the appropriate rating for each of these factors using the rating criteria provided in the procedural document, shown in Figure 4-2.

Factor	Rating	Criteria (Examples) – choose most likely <i>Read criteria from the top for each factor</i>
Injury Severity	Serious S3	Normally non-reversible; likely will not return to the same job after recovery from incident: <ul style="list-style-type: none"> - fatality - limb amputation - long term disability - chronic illness If any of the above are applicable, the rating is SERIOUS
	Moderate S2	Normally reversible; likely will return to the same job after recovery from incident: <ul style="list-style-type: none"> - broken bones - severe laceration - short hospitalization - short term disability - loss time (multi-day) - fingertip amputation (not thumb) If any of the above are applicable, the rating is MODERATE
	Minor S1	First aid; no recovery required before returning to job: <ul style="list-style-type: none"> - bruising - small cuts - no loss time (multi-day) - does not require attention by a medical doctor If any of the above are applicable, the rating is MINOR
Exposure	High E2	<ul style="list-style-type: none"> - Typically more than once per hour - Frequent or multiple short duration - Durations/situations which could lead to task creep and does not include teach If any of the above are applicable, the rating is HIGH
	Low E1	<ul style="list-style-type: none"> - Typically less than or once per day or shift - Occasional short durations If either of the above are applicable, the rating is LOW
Avoidance	Not possible A3	<ul style="list-style-type: none"> - Insufficient clearance to move out of the way and safety-rated reduced speed control is not used - The robot system layout causes the operator to be trapped, with the escape route toward the hazard - Safeguarding is not expected to offer protection from the process hazard (e.g. explosion or eruption hazard) If any of the above are applicable, the rating is NOT POSSIBLE
	Not likely A2	<ul style="list-style-type: none"> - insufficient clearance to move out of the way and safety-rated reduced speed control is used - obstructed path to move to safe area - hazard is moving faster than reduced speed (250 mm/sec) - inadequate warning/reaction time - might not perceive the hazard exists If any of the above are applicable, the rating is NOT LIKELY
	Likely A1	<ul style="list-style-type: none"> - sufficient clearance to move out of the way - hazard is incapable of moving greater than reduced speed (250 mm/sec) - adequate warning/reaction time - positioned in a safe location away from the hazard If any of the above are applicable, the rating is LIKELY

Figure 4-2: Severity - Exposure - Avoidance Ratings and Criteria [6]

While the criteria for each category do not require specific numbers, the methodology still takes an expected value approach to determine risk levels. While severity is still considered

as before, exposure and avoidance take the place of probability to be combined and result in one of the determined risk levels — Very High, High, Medium, Low, or Negligible. The risk level decision matrix is shown in Figure 4-3.

Severity of Injury	Exposure to the Hazard	Avoidance of the Hazard	Risk Level	
S1 - Minor	E1 - Low	A1 - Likely	NEGLIGIBLE	
		A2 - Not Likely		
		A3 - Not Possible	LOW	
	E2 - High	A1 - Likely		MEDIUM
		A2 - Not Likely		
		A3 - Not Possible	HIGH	
S2 - Moderate	E1 - Low	A1 - Likely		MEDIUM
		A2 - Not Likely		
		A3 - Not Possible	HIGH	
	E2 - High	A1 - Likely		HIGH
		A2 - Not Likely		
		A3 - Not Possible	VERY HIGH	
S3 - Serious	E1 - Low	A1 - Likely		HIGH
		A2 - Not Likely		
		A3 - Not Possible	VERY HIGH	
	E2 - High	A1 - Likely		VERY HIGH
		A2 - Not Likely		
		A3 - Not Possible		

Figure 4-3: Risk Level Decision Matrix [6]

These risk ratings then have an impact on the minimum risk reduction measure requirements and on the functional safety requirements associated with the safety-related parts of the control system (SRP/CS). The overall trend of the risk reduction measures is that Medium to Very High Risk Levels require at least one risk reduction measure to involve elimination of, substitution of, or safeguarding from the hazard while Negligible to Low Risk Levels can be mitigated by any means. This sets the grounds for discussion during the next

step of the process as the team determines how they will mitigate the identified hazards. The minimum risk reduction measures as a function of the risk level chart are shown in Figure 4-4.


	Risk Reduction Measure	Risk Level				
		VERY HIGH	HIGH	MEDIUM	LOW	NEGLIGIBLE
<p style="text-align: center;"> Most Preferred  Least Preferred </p>	Elimination	One or a combination of Elimination, Substitution, and Safeguarding or SRP/CS is REQUIRED to reduce risks to an acceptable level.			One or any combination of the Risk Reduction Measures that will acceptably reduce the Risk Level may be used.	
	Substitution					
	Safeguarding/ SRP/CS					
	Warnings and Awareness Means	Complementary Protective Measures may be used in conjunction with the above risk reduction measures but shall not be used as the primary risk reduction factor				
	Administrative Controls					
	PPE					

Figure 4-4: Minimum Risk Reduction Measures [6]

Once all tasks, hazards, and initial risk levels have been determined, the team begins determining risk reduction measures to mitigate the hazards. The process uses the hierarchy of control, shown in Figure 4-5, favoring mitigations on one end of the spectrum that eliminate the hazards from the tasks over mitigations that simply provide protective equipment to the people potentially affected by the hazard.

Once all of the risk reduction measures for all of the task and hazard combinations are determined, the severity, exposure, and avoidance ratings are reanalyzed in order to determine if the measure was sufficient in lowering the risk to an acceptable level. If the risk is not lowered to an acceptable level, the hazard mitigations are redetermined and reassessed until an acceptable level is reached.

Finally, the last two steps of the process are to document the assessment in a format that captures the discussion and data points discussed throughout and to update the assessment periodically. An example of a completed risk assessment form showing this step is given in Figure 4-6.

The time interval of the periodic update is left at the discretion of the team, however, the process also recommends updating the assessment when the system is changed, moved, or is part of an accident.

Most Preferred



Least Preferred

1) Elimination	<ul style="list-style-type: none"> Process design, redesign or modification including changing layout to eliminate hazards such as falls, hazardous materials, noise, confined spaces, pinch points (increase clearances) & manual handling Eliminate or reduce human interaction in the process Automate tasks, material handling systems (e.g. lift tables, conveyors, balancers), automated ventilation systems 	Elimination or Substitution
2) Substitution	<ul style="list-style-type: none"> Substitute for less hazardous material Intrinsically safe (energy containment) Reduce energy (e.g. lower speed, force, amperage, pressure, temperature, and noise) 	
3) Safeguarding and Safety Related Parts of the Control System (SRP/CS)	<ul style="list-style-type: none"> Guarding Acoustic enclosures (typically interlocked) Circuit breakers Platforms and guard railing Interlocks / interlocking Sensitive protective equipment Two-hand controls Enabling devices Safety controls / safety logic / safety systems (and integration of interlocking, safety parameters, and safeguarding) Safety-rated speed, position, location limits 	Safeguarding
4) Warnings and Awareness Means	<ul style="list-style-type: none"> Emergency stop devices and systems¹ Lights, beacons and strobes Backup alarms, beepers, horns, sirens, other visual and audible means Computer warnings Signs, labels 	Complementary Protective Measures
5) Administrative Controls	<ul style="list-style-type: none"> Training and safe job procedures Safety equipment inspections Rotation of workers Changing work schedule Lockout Hazard communications Confined space entry 	
6) Personal Protective Equipment	<ul style="list-style-type: none"> Safety glasses, face shields Hearing protection Safety harness and lanyards Gloves Hard hats Respirators Clothing & footwear for specific safety purposes (e.g. Kevlar sleeves, metatarsal shoes) 	

Figure 4-5: Hierarchy of Control [23]

Line No	Task Description	Hazards	Initial Risk				Risk Reduction Measures	Final Risk				
			Severity	Exposure	Avoidance	Risk Level		Severity	Exposure	Avoidance	SRPCS of mitigation (if applicable)	Risk Level
			S1 - S3	E1 - E2	A1 - A3	Tbl 2		S1 - S3	E1 - E2	A1 - A3	Tbl 5	Tbl 2
1	Task 1	mechanical : unexpected start	S2	E2	A2	R1	Use of pendant & enabling devices, safety scanners, Safety PLC, Systems programming requiring operator confirmations, Training & Procedures	E1	A1	S2		R3B
2	Task 1	mechanical : Crushing, pinching, impact movement system	S2	E1	A1	R2B	Training and procedures; situational awareness	E1	A1	S2		R3B
3	Task 1	slips / trips / falls : trip	S1	E1	A1	R4	Housekeeping, caution and situational awareness	E1	A1	S1		R4
4	Task 2	ergonomics / human factors : lifting / bending / twisting	S1	E1	A2	R3B	Proper techniques; adherence to shop practices	E1	A1	S1		R4
5	Task 3	lasers : eye exposure	S1	E2	A1	R3A	Position the laser to a safe location or shut down when not in use. Class 2 laser. Limit access to the build area.	E1	A1	S1		R4

Figure 4-6: Example Task-based Risk Assessment Form [6]

4.2.4 Discussion of Task-based Risk Assessment Methodology

The ANSI/RIA Task-based Risk Assessment methodology attempts to use a Job Hazard Analysis methodology combined with bottom-up and top-down processes similar to FMEA and FTA. However, in an attempt to combine the best practices of all the traditional methodologies, the Task-based Risk Assessment suffers from two significant issues when used in a complex, semi-automated manufacturing environment. The first is that the assessment simply identifies, assesses, and determines mitigations for hazards with no causal analysis, discounting the importance of understanding the hazards or assuming that the assessors already know the causes. The second is that throughout the entire process the assessment retains the unchanged foundation of flawed assumptions for modern socio-technical systems.

Initially, the assessors view the system in its entirety and discuss its functionality in order to break it into its tasks or functional “component parts” for a bottom-up, functional assessment process similar to a FMEA. Once all of the tasks are determined, the process takes a top-down approach similar to a FTA by identifying the hazards to avoid for the task and determining how the hazards can be prevented or mitigated. This process becomes a repeated exercise of a simplistic Job Hazard Analysis methodology to identify and assess hazards, assuming the causes are already known. Then the process attempts to aggregate and prioritize all of the data to determine a risk mitigation strategy for the system. While

this process is an attempt to combine the best practices of all the traditional methodologies, it limits its ability to be effective, lacking a formal causal analysis and being rooted in traditional methods, and limits its ability to be efficient, attempting to be everything at once.

As seen in the process steps of the assessment, this methodology is focused on simply identifying hazards and does not have a detailed means to understand them. While this may be sufficient for a job hazard analysis of simple manual processes or mechanical systems, the automated systems being developed today for humans to supervise, operate, maintain, and repair are more complex than ever before and are introducing new hazards with causes that were never considered in the past. Complex hazards that include factors such as hazardous component interactions, software requirement flaws, human complacency and/or confusion due to automation design, or inappropriate human behavior are simply labeled and accepted as unexpected, unintended, and/or inappropriate. This lack of causal analysis to understand the hazards has a detrimental impact on the remainder of the assessment process.

The first impact is when the assessment methodology turns to the subjective criteria of severity, exposure, and avoidance to determine expected values for assessing the risk level. Consider the criterion of exposure, which is generally a function of frequency, duration, and possibility that a person is in the vicinity of the hazard when it occurs. It is not possible to determine these without understanding the potential causes of the hazard. It may be possible to determine if someone will be in the vicinity of the hazard being assessed, but without fully understanding the causal factors of the hazard it could be arguably impossible to determine the frequency and duration. If the cause is a software error, the frequency will be every time that the process with the software error is executed, not some arbitrary number assigned to unexpected movement. The same types of considerations apply to the avoidance and severity criteria.

Compounding the impact of not understanding the hazard, this methodology is subject to the same fundamental flaws previously discussed for using just probability and severity to determine expected values of risk. Will the same hazard injure every person in the same manner? Is everyone in the manufacturing environment equally able to avoid a moving object? What if they can't see or hear the robot moving towards them? How does someone

determine an accurate assessment of exposure in different scenarios such as operating in an ideal state versus when an issue arises and the task requires multiple people to be in vicinity of the hazard while the system is operating to fix the issue? The wording used in the rating criteria for each factor alone shows the potential for error when the directions state to choose the most likely rating category and the criteria for those rating categories begin with normally, likely, and typically in the examples. Under this process, the risk level assessment will only be as useful as the flawed information that went into the methodology allows.

In addition to the lack of causal analysis, underlying this entire process are the assumptions of looking for multiple events sequenced as a forward chain over time, discounting systemic factors that could impact randomness and independence, searching for a root cause, and blaming accidents on some type of failure of human error. This impacts the subjective discussion to determine each hazards' risk level and the determination of appropriate risk reduction measures. Furthermore, the error-prone expected value process determines the prioritization and determination of minimum risk reduction measures that the company must implement. As long as the determined values for the subjective criteria result in a risk level that is low, they can be mitigated with PPE or administrative controls when no actual understanding of the hazard has been reached. On the other hand, hazards rated as very high must be mitigated with potentially expensive and unnecessary solutions that do not benefit the company.

Not having a formalized and thorough causal analysis and the traditional underlying assumptions have a significant negative impact on the effectiveness for assessing modern, complex systems with the Task-based Risk Assessment.

In terms of efficiency, the assessment lacks a detailed and organized process for the underlying steps, which allows the analysis to be conducted to the extent the practitioner sees fit. For example, there is minimal guidance given in the two procedural documents for identifying the tasks for the system. Consider one of the listed example tasks, such as troubleshooting. Does simply listing "troubleshooting" as a task provide the appropriate level of granularity to identify all of the possible hazards? Would the assessment team be able to trace all of the potential hazards back to the task in an effective manner to then

determine an appropriate risk reduction measure in a later step? There are many different levels and types of troubleshooting that could occur on numerous parts of the system, all potentially with multiple hazards associated with the specific tasks.

To illustrate the difficulties and potential variability with determining the appropriate granularity for tasks, consider the average passenger vehicle and the task of troubleshooting. In the most simple and least granular view, we could encapsulate any and all troubleshooting under one task. With a view from the other end of the spectrum, levels of troubleshooting may include basic tasks the operator can perform, tasks requiring additional tools and expertise that the common garage mechanic can perform, and then in-depth tasks requiring specific specialized equipment and potentially proprietary knowledge only a mechanic at the originating company dealership can perform. Additionally, the average car has approximately 30,000 physical parts that could all potentially require some type of intervention. Considering three different levels of maintenance for 30,000 physical parts on the car, we could potentially have as many as 90,000 troubleshooting tasks to consider. While this may be an extreme view of what could be considered a task, a standardized process that provides very little guidance on how to determine the appropriate level of granularity can easily cause an assessment team to become completely inefficient as it potentially moves from assessing one task to 90,000 tasks for the associated hazards. This methodology can quickly begin to consume an inappropriate amount of time for the amount of value the results provide.

Although this analysis provides an argument that the Task-based Risk Assessment does not achieve the level of effectiveness and efficiency desired for a safety assessment, numerous companies continue to use this methodology endorsed by the established standards organization. The source document for the methodology states that “any risk assessment methodology that identifies tasks and prescribes risk reductions equivalent to or more stringent than the requirements of this methodology is acceptable.” One could argue that this is an implicit endorsement of the methodology that anchors readers to the idea that the suggested methodology is the only approved and acceptable one. Dependent upon the company culture and receptiveness to change, people in a company may be concerned that they would be to blame if someone was to use another methodology and potentially miss a hazard that led to an accident. Nevertheless, some companies decide to adopt the recommended methodology

and integrate it into their own policies and procedures for safety risk assessments.

Using a combination of concepts from multiple analysis processes and neglecting the importance of a causal analysis, the Task-based Risk Assessment inherits a number of foundational assumptions that reduce the effectiveness of the assessment results. The underlying goal of proving that the system is safe combined with the subjective assessments of factor rating criteria to determine the overall risk level for each hazard undermines the process and provides optimistic results. By solely identifying and not fully understanding the hazards and the causes, the engineers and system designers are not able to determine mitigations that eliminate or substitute hazards in the system to improve workplace safety. Additionally, the lack of a systematic methodology within the process sub-steps allows a great deal of variability in the assessment breadth and depth, depending on the assessor.

4.2.5 Discussion of Task-based Risk Assessment in Practice

When applied to the semi-automated manufacturing process case study, the Task-based Risk Assessment included all of the issues stemming from the theoretical application of the process as designed and a number of other issues that could be common to other organizations.

The Assessment Team

The assessment team composition met and exceeded all of the specified recommendations from the procedure. The cross-functional team averaged 12 personnel from integration, engineering, operations, and maintenance. The majority of these people were considered the leads or subject matter experts within their scope of responsibility for the system. Additionally, this analysis was for a system that was being procured through a supplier and they were represented in the assessment as well. The facilitator was a well-respected and influential figure for the organization in terms of industrial robot applications and the risk assessment process. He had also led many of these discussions and brainstorming sessions in the past.

Adherence to the Process

The analysis started the same way as prescribed in the process, discussing the analysis process, discussing the system and its limits, determining the portions of the system lifecycle to include in the analysis, and developing the entire list of tasks to be considered in the assessment. However, at this step the analysis deviated from the standard procedure and began identifying hazards for a single task. Once the list of hazards was developed for a task, the team focused on one hazard at a time to assess severity, exposure, and avoidance, assuming no safeguards installed, to determine the initial risk level. Immediately the team would begin developing the risk mitigation measures and then reassess the hazard with the mitigations in place for the final risk level determination. This iterative process was then followed for the remaining hazards for the task and then repeated for the remaining tasks. An illustration of the differences in the theoretical and actual processes is depicted in Figure 4-7 for further clarification.

Deviating from the standard process allowed the team to assess, mitigate, and reassess each hazard in a seemingly more efficient manner. However, this methodology viewed each hazard individually as one might in a reliability assessment, viewing the individual hazards (components) of the task (system) and mitigating them in isolation. It prevented the group from gaining a thorough understanding of the system in its entirety, to include all of the identified hazards, before determining what the appropriate risk mitigations are for the system in terms of effectiveness and cost.

Assessment Environment

The analysis was completed over the course of five days. The entire team gathered in a conference room for eight hours a day during the first three days in an attempt to complete the analysis. Due to numerous competing tasks and concurrent requirements for each of the individuals involved and the limited amount of time in each day, most of the people in the room were “multitasking” on their laptops and phones. Additionally, there was some turnover for people in the group that could participate in the assessment on one day, but not another due to other meetings and requirements. This led to a number of unanswered

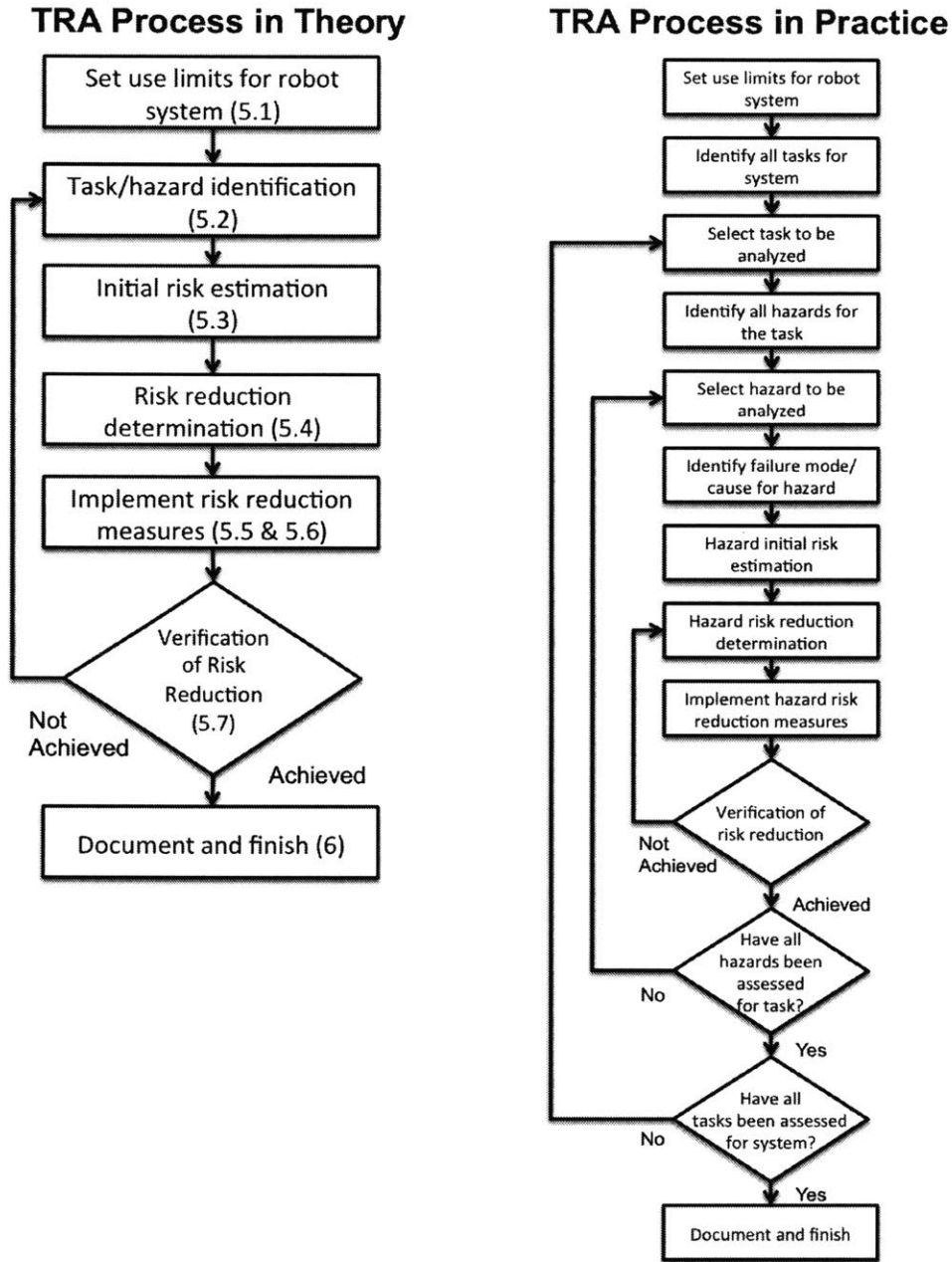


Figure 4-7: Comparison of Task-based Risk Assessment in Theory vs. Practice

questions being written down and saved for later or best guesses and assumptions made by the group to continue the assessment.

The assessment was originally scheduled for only three days, but was not completed by the end of the third day. Due to the assessment's group execution method and the number of different competing conflicts for the individuals in the group, the assessment could not be completed in the same format during the following days. The final consensus among the group was to have the assessment leader and another knowledgeable engineer from the integration group finish the remaining half of the assessment, send it out to the rest of the group for feedback, and then aggregate the feedback and finalize the assessment. The draft assessment was completed in the following two days and then sent out to the group. There were no changes between the draft assessment and the final version. Without considering the time that may have been involved in reviewing the draft assessment for the final version, the total personnel hours involved to complete the assessment was approximately 1,000 hours.

With the number of competing work tasks and other requirements for everyone involved in the assessment, the execution process was a significant strain on the assessment team. The only person that was solely dedicated to the assessment over the course of the five days was the facilitator. The majority of the remaining group rotated into and out of the process as their other priority requirements dictated and even on the days they were involved they had numerous distractions that arose over the course of the eight-hour day. Even though some may suggest that they should have made this a priority and paid more attention, due to the structure of the process and the topics of discussion some members of the team sat through many hours of discussion that did not involve their opinions or expertise at all. Additionally, due to the subjective factors that were the primary topics of discussion, there was not much need for the different subject matter experts. This was demonstrated at the halfway point of the analysis when only two people were able to complete the assessment for the rest of the group. Conducting an assessment in a fast-paced, dynamic, and matrix-organized company that requires a large amount of time from numerous people all at once requires an even larger commitment from the people and organizations involved.

Task-based Risk Assessment Results

The final version of the report contained 53 individual tasks for the system across its lifecycle with an associated 500 hazards. From an assessment completion point of view, the team supposedly determined risk reduction measures that mitigated the 500 hazards, argued that the system is safe, and moved one step closer to getting the equipment out onto the factory floor. An isolated example section from the report is presented in Figure 4-8.

Task	Hazard	Failure Mode/Cause	Severity	Exposure	Avoidance	Risk Category	Risk Reduction Methods	Exposure	Avoidance	Severity	Risk Category
Attach AGV to PTS	mechanical : unexpected start	Unintentional or unexpected movement	S2	E2	A2	R1	Use of pendant & enabling devices, safety scanners, Safety PLC, Systems programming requiring operator confirmations, Training & Procedures	E1	A1	S2	R3B
Attach AGV to PTS	mechanical : Crushing, pinching, impact movement system	Person standing between AGV and PME; person reaching into harms way	S2	E1	A1	R2B	Training and procedures; situational awareness	E1	A1	S2	R3B
Attach AGV to PTS	slips / trips / falls : trip	Tripping over object on floor	S1	E1	A1	R4	Housekeeping, caution and situational awareness	E1	A1	S1	R4
Attach AGV to PTS	ergonomics / human factors : lifting / bending / twisting	Exposure during process	S1	E1	A2	R3B	Proper techniques; adherence to shop practices	E1	A1	S1	R4
Attach AGV to PTS	lasers : eye exposure	Person looks into the laser beam. Laser beam not aimed in a safe location (not in use)	S1	E2	A1	R3A	Position the laser to a safe location or shut down when not in use. Class 2 laser. Limit access to the build area.	E1	A1	S1	R4
Manual drive PTS to the cell	mechanical : unexpected start	Unintentional or unexpected movement	S2	E2	A2	R1	Use of pendant & enabling devices, safety scanners, Safety PLC, Systems programming requiring operator confirmations, Training & Procedures	E1	A1	S2	R3B
Manual drive PTS to the cell	mechanical : Crushing, pinching, impact movement system	Failure of navigation & guidance system; inattention of operator; operator error	S2	E1	A1	R2B	AGV Scanners, compliance with ANSI 56.5; training and procedures	E1	A1	S2	R3B
Manual drive PTS to the cell	slips / trips / falls : trip	Tripping over object on floor	S1	E1	A1	R4	Housekeeping, caution and situational awareness	E1	A1	S1	R4
Manual drive PTS to the cell	ergonomics / human factors : excessive force / exertion	Long term use of AGV pendant	S1	E2	A2	R2C	Investigate personnel rotation	E1	A2	S1	R3A
Manual drive PTS to the cell	ergonomics / human factors : duration	Long term use of AGV pendant	S1	E2	A2	R2C	Investigate personnel rotation	E1	A2	S1	R3A
Manual drive PTS to the cell	lasers : eye exposure	Person looks into the laser beam. Laser beam not aimed in a safe location (not in use)	S1	E2	A1	R3A	Position the laser to a safe location or shut down when not in use. Class 2 laser. Limit access to the build area.	E1	A1	S1	R4

Figure 4-8: Example Results Format from the Task-based Risk Assessment

However, a closer look at the results of the report show how the theoretical issues discussed previously manifest into ineffective risk reduction measures being accepted and believed to reduce risk levels. The most frequent cause of a hazard, such as mechanical crushing, pinching, and impact, was unexpected or unintentional movement. A few of these instances from the assessment process are shown in Figure 4-9.

Hazard	Failure Mode/Cause	Risk Reduction Methods
mechanical : Crushing, pinching, impact movement system	Unexpected movement of PTS or stations	Standard Shop practices, procedures, training, etc. Situational awareness. Use of safety scanners during movement. Standing clear of movement
mechanical : Crushing, pinching, impact by robo/end effector	Unexpected movement. Safety components/systems not functioning properly.	Enabling and Robot Pendant. "Operator Lock-out", interlock gate. Operator training. Safety PLC
mechanical : unexpected start	Unintentional or unexpected movement	Perimeter fencing w/ interlocked gates; use of pendant & enabling devices, Safety PLC, Systems programming requiring operator confirmations, Training & Procedures
mechanical : Crushing, pinching, impact movement system	Safety components/systems not functioning properly	When possible test safety systems without exposure to risks or with limited exposure
ergonomics / human factors : duration	Long term use of pendant	Investigate personnel rotation
ergonomics / human factors : lifting / bending / twisting	Exposure during process	Proper techniques; adherence to shop practices

Figure 4-9: Example Results Analysis

Overall, unexpected movement was determined as the cause of an identified hazard more than 70 times. The first listed hazard was determined to be mitigated to an acceptable level using risk reduction measures such as standard shop practices, situational awareness, and standing clear of movement. How does someone in the workplace stand clear of something that unexpectedly moves? The next hazard expands beyond unexpected movement and also includes safety components/systems not functioning properly, but then lists a safety system as the risk reduction method. The third hazard listed reduces the risk of unintentional or unexpected movement by safeguarding the space with a fence and using pendants and enabling devices if the users need to enter the space. However, if the company does not know how or when the system will move, how does a controller in someone’s hand mitigate the risk? If anything it makes the situation more hazardous by providing a false sense of security for the people nearest to the system and its associated hazards.

One of the common themes of observing the risk assessment process was the large amount of variability between individual’s ideas of severity, exposure, and avoidance due to the subjective nature of the categories. In order to achieve a group consensus on the appropriate

level, the team would have to have a conversation and make numerous assumptions to develop the context. By settling on a specific context for the answer, the resulting risk level was typically one of the more optimistic possibilities and the potential effectiveness of the process was limited.

When analyzed against the assessment criteria, the traditional process fulfills all of the feasibility criteria as the mandatory government and voluntary industry safety regulations and standards require that a safety assessment be completed and only the industry standards provide a recommended methodology. Additionally, the assessment meets the internal company requirements to complete a safety risk assessment and the process is standardized and published in the procedural document from the Robotics Industries Association, called RIA TR R15.306-2014.

The traditional process has a number of issues when analyzed using the quality criteria. Currently the assessment process is one of the final steps to implementing equipment on the manufacturing floor and the majority of the system development work is already complete. While the process may state that it can be used earlier in the development lifecycle, the results are not easily update-able or useful for later stages due to the composition of the results and the overall subjective nature of the assessment. A large amount of context and detail is required to assess the severity, exposure, and avoidance of an identified hazard, which is not available earlier in development. Additionally, because the goal and structure of the traditional process is simply to identify all of the possible hazards involved with a task and mitigate them, a wide range of hazards can be easily identified and assessed, as shown in the 500 hazards across the 53 tasks. While these identified hazards may encompass the hardware failures and software and component interactions, the traditional process does not promote developing an understanding of how the hazards may occur, as shown in all of the examples of unexpected movement. The majority of the mitigations for unexpected or unintentional movement were all focused on controlling the operator or surrounding workers and not on controlling or understanding the system behavior.

In addition to the technical system, the assessment process stops with the human action and does not attempt to understand why the operators or surrounding workers did what they did, as shown in the examples of unintentional movement. Because the assessment only

requires the human action involved in identifying a possible hazard, it does not foster a safety related discussion of human factors, the roles of management, operations, and processes, or the environment. All of these items were either not discussed due to various reasons or just accepted as presented due to how late the assessment was occurring in the development process when considering budget and schedule, such as human factors considerations for or management control of the AGV controllers. Overall, the traditional process is limited to the technical system and basic human actions for the assessment to identify hazards but does not create a thorough understanding before developing hazard controls.

The last method used to assess the outcome of the Task-based Risk Assessment was by reviewing all selected risk reduction measures and classifying them in accordance with the hierarchy of hazard control. All of the individual mitigations were classified, even if there were multiple mitigations for a single identified hazard. If a mitigation seemed to be a blend of two categories, the more preferred category was selected. After successfully mitigating 500 hazards for the system over the course of five days with a team of highly trained personnel, the results of the assessment were surprising. The most preferred method of designing hazards out of the system only accounted for 4% of the total risk reduction measures, the next preferred method of implementing engineering controls accounted for 34%, and the least preferred method of implementing administrative controls accounted for a strong majority of 62%. These results are presented with the hierarchy of hazard control chart, showing potential mitigation examples, on Figure 4-10.

As applied in the case study, the execution and the results of the Task-based Risk Assessment demonstrated the theoretical flaws of the assessment process. Due to the assessment goals and format, it directs the team's mindset toward proving that a system is safe, resulting in many best-case scenario assumptions for context and subjective decisions such as determining severity, exposure, and avoidance. The results are further diminished by the unstructured sub-steps that create variability and inefficiency in the process execution, which is also compounded by the competing requirements for the assessment team's time and attention. The lack of technical rigor prevents the assessment team from using the process to better understand the hazards present in the system. It promotes acceptance of not understanding the system, as shown in the number of unexpected and unintentional

TRA		Risk Reduction Measures	Examples	Influence on Risk Factors	Classification
4%	<p style="text-align: center;">Most Preferred</p> <p style="text-align: center;">Least Preferred</p>	Elimination or Substitution	<ul style="list-style-type: none"> Eliminate pinch points (increase clearance) Intrinsically safe (energy containment) Automated material handling (robots, conveyors, etc.) Redesign the process to eliminate or reduce human interaction Reduced energy Substitute less hazardous chemicals 	<ul style="list-style-type: none"> Impact on overall risk (elimination) by affecting severity and probability of harm May affect severity of harm, frequency of exposure to the hazard under consideration, and/or the possibility of avoiding or limiting harm depending on which method of substitution is applied 	Design Out
34%		Guards, Safeguarding Devices, and Complementary Measures	<ul style="list-style-type: none"> Barriers Interlocks Presence sensing devices (light curtains, safety mats, area scanners, etc.) Two hand control and two-hand trip devices 	<ul style="list-style-type: none"> Greatest impact on the probability of harm (Occurrence of hazardous events under certain circumstances) Minimal if any impact on severity of harm 	Engineering Controls
62%		Awareness Devices	<ul style="list-style-type: none"> Lights, beacons, and strobes Computer warnings Signs and labels Beeper, horns, and sirens 	<ul style="list-style-type: none"> Potential impact on the probability of harm (avoidance) No impact on severity of harm 	Administrative Controls
		Training and Procedures	<ul style="list-style-type: none"> Safe work procedures Safety equipment inspections Training Lockout / Tagout / Verify 	<ul style="list-style-type: none"> Potential impact on the probability of harm (avoidance and/or exposure) No impact on severity of harm 	
		Personal Protective Equipment (PPE)	<ul style="list-style-type: none"> Safety glasses and face shields Ear plugs Gloves Protective footwear Respirators 	<ul style="list-style-type: none"> Potential impact on the probability of harm (avoidance) No impact on severity of harm 	

Figure 4-10: TRA Risk Reduction Measures Hierarchy of Controls Classifications

movement causes, but focuses on simple and easily understood tasks, such as wearing a protective mask for dust particles while drilling into concrete. The isolated and iterative nature of the process reduced the effectiveness of the risk reduction measures for the system and most likely increased the necessary costs. While the Task-based Risk Assessment may fulfill a mandated requirement to conduct a risk assessment for an automated system in the manufacturing workplace, its limited effectiveness and inefficient process does not support a company's goal of a safe workplace.

4.3 Proposed Analysis Process - STPA

As companies identify opportunities to introduce automated systems to complete traditionally manual manufacturing processes, they also introduce entirely new levels of complexity and hazards to the process, the surrounding workspace, and the overall organization. In order to safely implement and continuously operate an automated system, the company must thoroughly understand the increasingly complex socio-technical system in which the automation operates. The proposed analysis process from STAMP, called Systems-Theoretic Process Analysis (STPA), is a systems thinking approach to safety that was designed specifically to assess modern, complex socio-technical systems. STPA is intended to identify missing or improper hazard controls that allow a system to migrate to higher levels of risk in a dynamic environment. The semi-automated manufacturing process being implemented by the company is a perfect example of the type of socio-technical system for which STAMP and STPA were developed.

The Assessment Team

The STPA assessment team was comprised of a cross-functional team from integration, engineering, operations, and maintenance that averaged eight personnel. All of the people involved were considered to be the leads or subject matter experts within their scope of responsibility for the system. The facilitator was a new employee to the company, recently trained in the assessment methodology with little familiarity with the new manufacturing process.

4.3.1 Defining Workplace Accidents

Although system safety and STAMP take a broad view of accidents that include more than just injury and death to humans, the company stakeholders involved in this case study agreed that it would be most beneficial to confine our analysis in this assessment to those losses in order to compare the results of both analysis processes. In review, System Safety and STAMP define an accident as:

Accident - *An undesired or unplanned event that results in a loss, including loss of human life or human injury, property damage, environmental pollution, mission loss, etc.*

For the purpose of the team's analysis of a complex and dynamically changing manufacturing environment, the team wanted to ensure we were consciously considering hazards beyond typical physical impact. A number of chemicals and hazardous materials are present in most manufacturing workplaces due to the processes involved in producing a product. To accomplish this, the team defined the system accident or loss as:

A: Death, Injury, or Illness to Humans

4.3.2 Defining Hazards and System Boundaries

The team then began working with multiple stakeholders from engineering and operations to determine what the company could consider to be inside the design space for drawing the system boundaries or outside as part of the environment when considering workplace safety. This was a very important step in the analysis because it established the boundary in a way that allows the designers to control the hazards as they are defined for the system. In review, system safety and STAMP define a hazard as:

Hazard – *A system state or set of conditions that, together with a particular set of worst-case environmental condition, will lead to an accident (loss).*

By empowering the team and allowing them to think beyond the basic system functions of completing a specified manufacturing task, they were challenged to take a more robust view of the system functionality in order to improve workplace safety. Additionally, the team discussed and reviewed accident and injury data for what are traditionally considered workplace hazards to ensure we could map all of them to the defined system-level hazards. Under this approach, the team defined the following list as the workplace system-level hazards to avoid:

- H1:** Exposure to uncontrolled energy (or energy at a level that could lead to a loss)
- H2:** Potentially injurious movement of the human body (or stress on the body) that could lead to injury
- H3:** Exposure to toxic materials above a safe level
- H4:** Exposure to noise levels that could affect hearing
- H5:** Extended exposure to an environment not providing basic human health requirements

From this overarching list of workplace system-level hazards, the team refined the hazards for the specific application. As seen in the traditional assessment, the majority of the hazards that the company identified but did not understand were related to the motion related functions of this and many other systems being implemented across the company. The refined list of hazards to avoid were defined as:

- H1:** Exposure to uncontrolled energy (or energy at a level that could lead to a loss)
 - H1.1:** Violation of minimum separation between AGVs and external objects
 - H1.2:** Violation of minimum separation between PTV and external objects
 - H1.3:** Violation of minimum separation between AGVs and PTV combination and external objects
 - H1.4:** Violation of minimum separation between Robotics and external objects

Other energy sources present in the current design that must be controlled for H1 include electrical, thermal, pneumatic, hydraulic, gravitational, mechanical, and non-ionizing radiation (lasers).

- H2:** Potentially injurious movement of the human body (or stress on the body) that could lead to injury
 - H2.1:** Potentially injurious movement of the human body (or stress on the body) that could lead to injury during routine operation
 - H2.2:** Potentially injurious movement of the human body (or stress on the body) that could lead to injury during maintenance, servicing, or troubleshooting

H2.3: Potentially injurious movement of the human body (or stress on the body) that could lead to injury during installation, repair, or overhaul

H3: Exposure to toxic materials above a safe level

H3.1: Workers are exposed to toxic chemicals above a safe level while operating the system

H3.2: Workers are exposed to toxic chemicals above a safe level while servicing the equipment

H3.3: Workers are exposed to toxic materials above a safe level from the manufacturing process

H4: Exposure to noise levels that could affect hearing

H5: Extended exposure to an environment not providing basic human health requirements

4.3.3 Model Development – Safety Control Structure

The team then worked together to develop the Safety Control Structure for the system. Due to the different functions and possible configurations of the system, the control structures were developed in each possible scenario. The team developed the model structure and then detailed the responsibilities, accountability, and authority of each controller. Within the responsibilities, the team detailed the process model and process model variables. Within the authority of each controller, the team detailed the possible control actions the controller could impose, how they were sent, when they were sent, and where the control action is sent. Within the accountabilities for each controller, the team detailed what feedback was provided, when it was provided, where the feedback went in the control structure, and how it was given. For this application, the team began by developing the control structure as a group to ensure that all of the major details were captured. However, the process continued with the individuals and small groups who had the technical expertise for different portions of the system to refine the models and ensure we captured the appropriate details.

As discussed in the manufacturing process overview, in the first scenario the operators must navigate an AGV from a charging station located in the factory to the location of

the product transportation vehicle and establish a connection. In the second scenario, the product transportation vehicle has functionality to connect to, hold, and disconnect from a product and to allow or not allow movement so it is modeled as an individual system as well. For the third scenario, the primary method for the product transportation vehicle to move around the factory is through system interactions with two AGVs. The fourth scenario is the robotic system included in the manufacturing process due to its functionality as a standalone system and its control responsibility for the product transportation vehicle and AGV control during the manufacturing process. These four scenario-based models are shown in Figures 4-11, 4-12, 4-13, and 4-14, respectively.

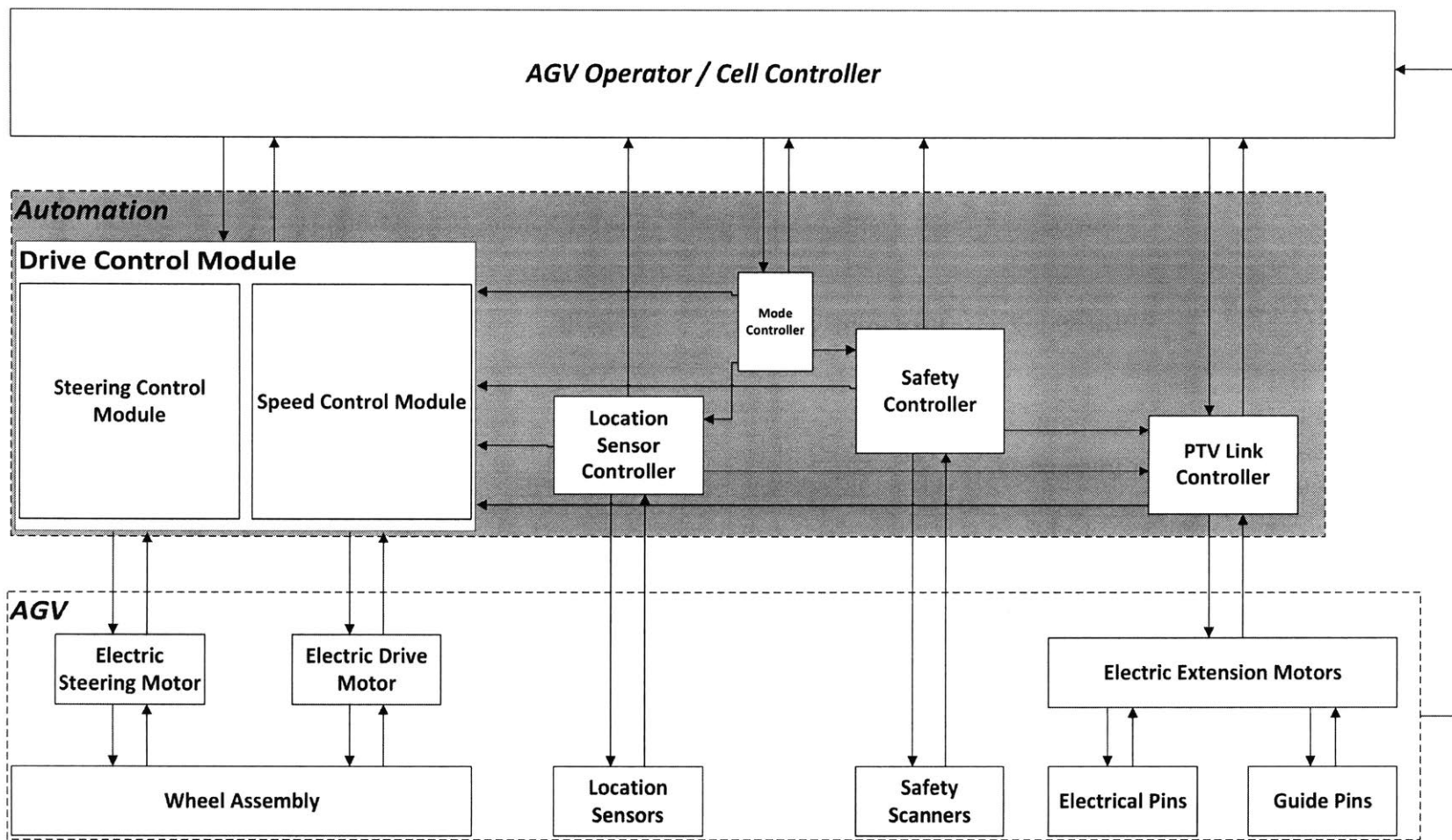


Figure 4-11: AGV Safety Control Structure

85

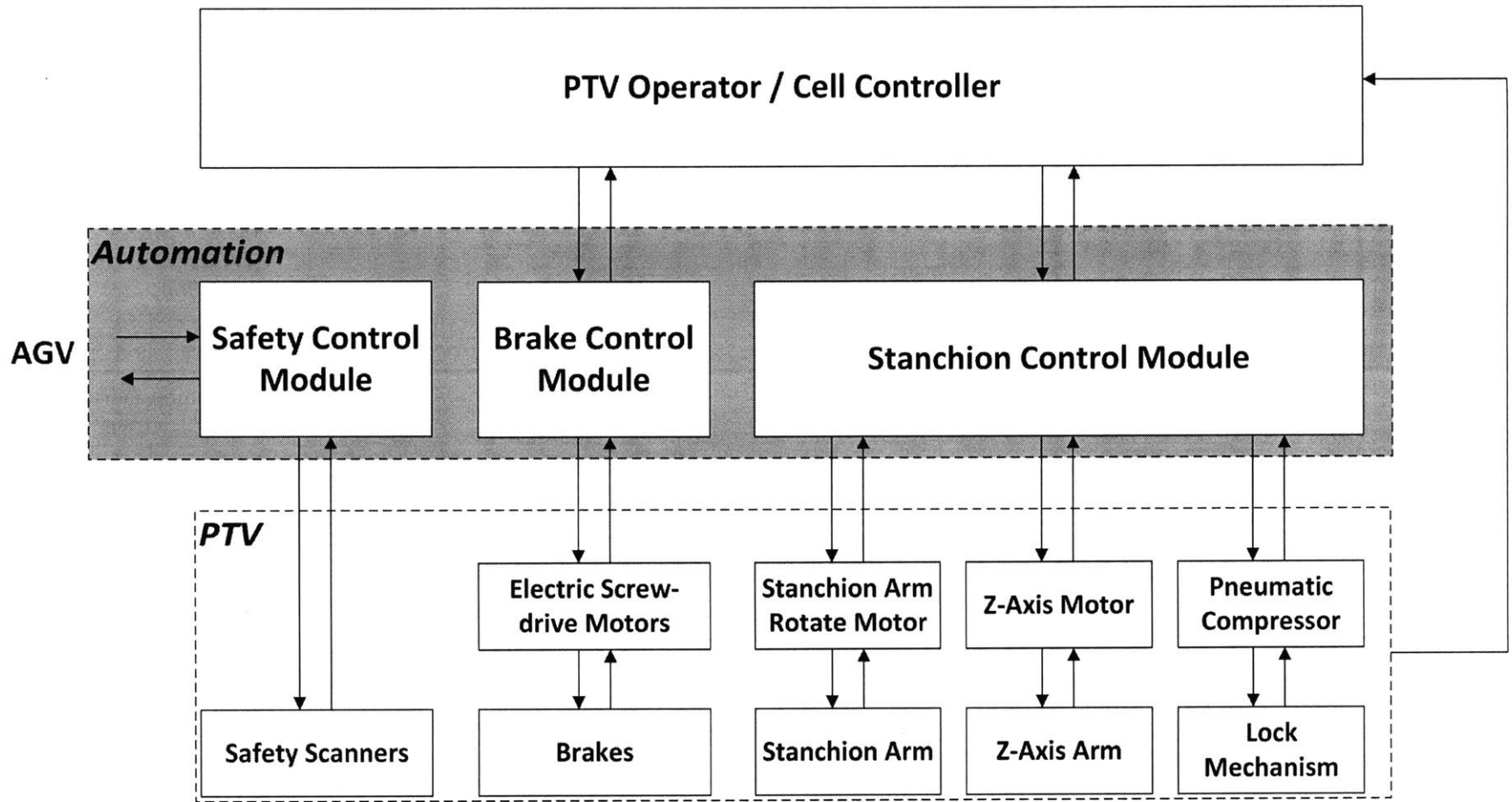


Figure 4-12: Product Transportation Vehicle Safety Control Structure

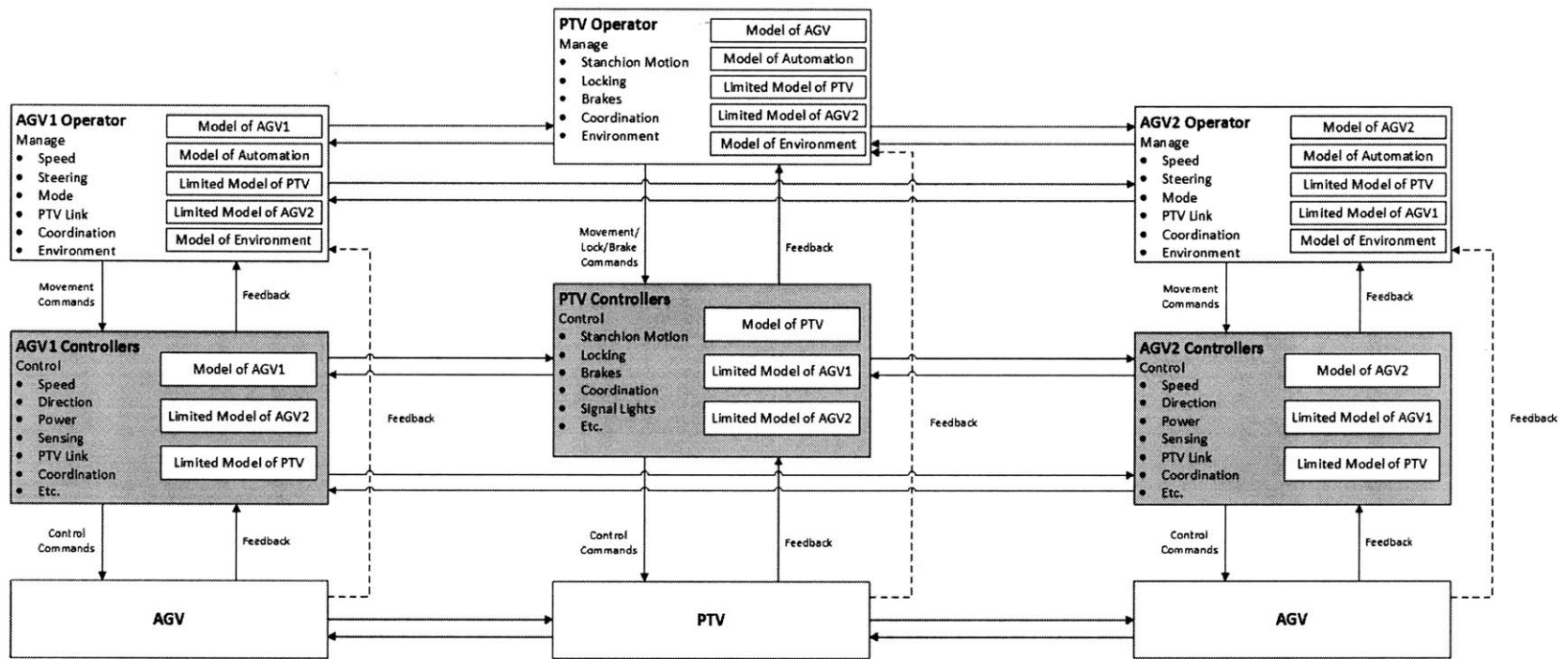


Figure 4-13: Combined Product Transportation System Safety Control Structure

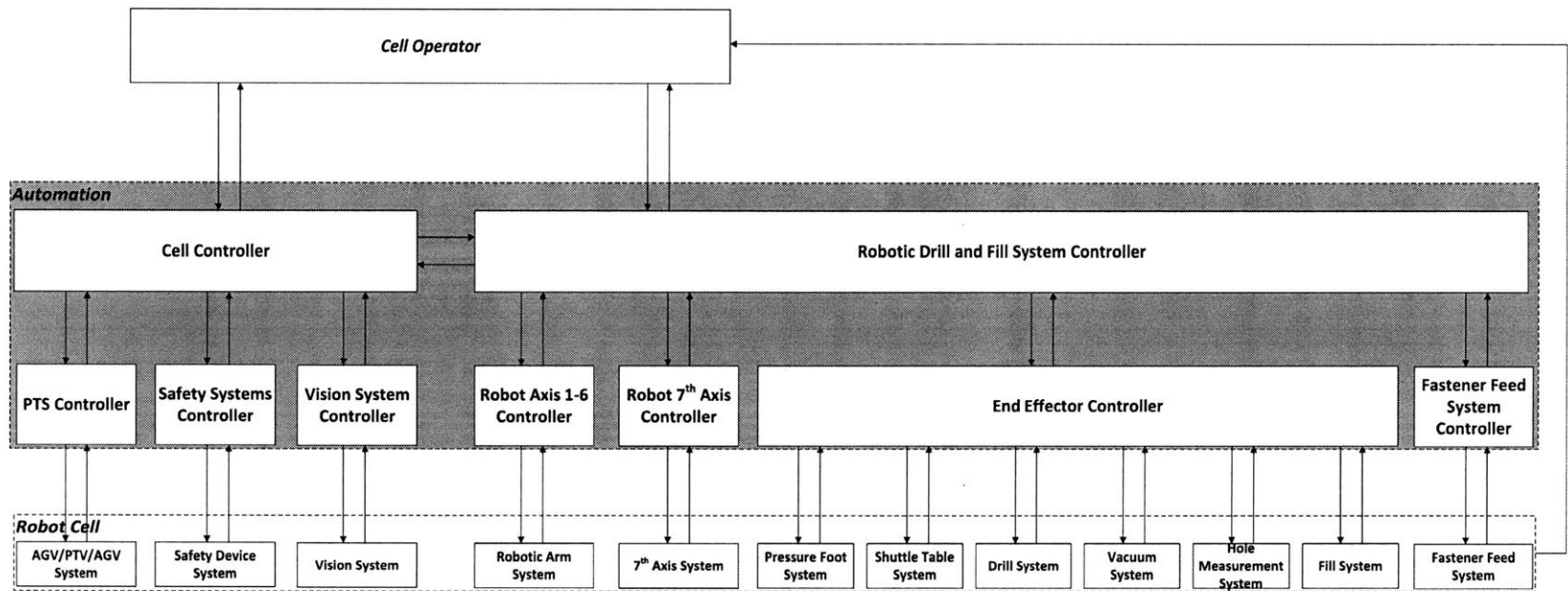


Figure 4-14: Robotic System Safety Control Structure

4.3.4 STPA Steps 1 & 2

While the control structures were primarily detailed as a group with some individual development, STPA Steps 1 and 2 were primarily performed as individuals and small groups. This had numerous benefits, such as allowing for more detailed and engaging discussions with the people involved in the different portions of the safety control structures' control loops and providing flexibility to work around individual competing work requirements. Additionally, while STPA Steps 1 and 2 can either be completed as two distinct steps or merged together while discussing each control action, the team used both methods. This allowed the team to first gain a basic understanding of the mechanics and importance of each step, but then work more efficiently through the potential causal factors in the control loop for each control action.

As previously noted, to identify potentially hazardous control actions in STPA Step 1, the team assessed each control action using the fact that control actions can be hazardous in four ways:

1. A control action required for safety is not provided or is not followed
2. An unsafe control action is provided that leads to a hazard
3. A potentially safe control action is provided too late, too early, or out of sequence
4. A safe control action is stopped too soon or applied too long (for continuous or nondiscrete control actions)

As an example of the work performed by the team, a more detailed segment of the AGV safety control structure developed is depicted in Figure 4-15 and the results from the analysis of one of the control actions is shown in Table 4.3.

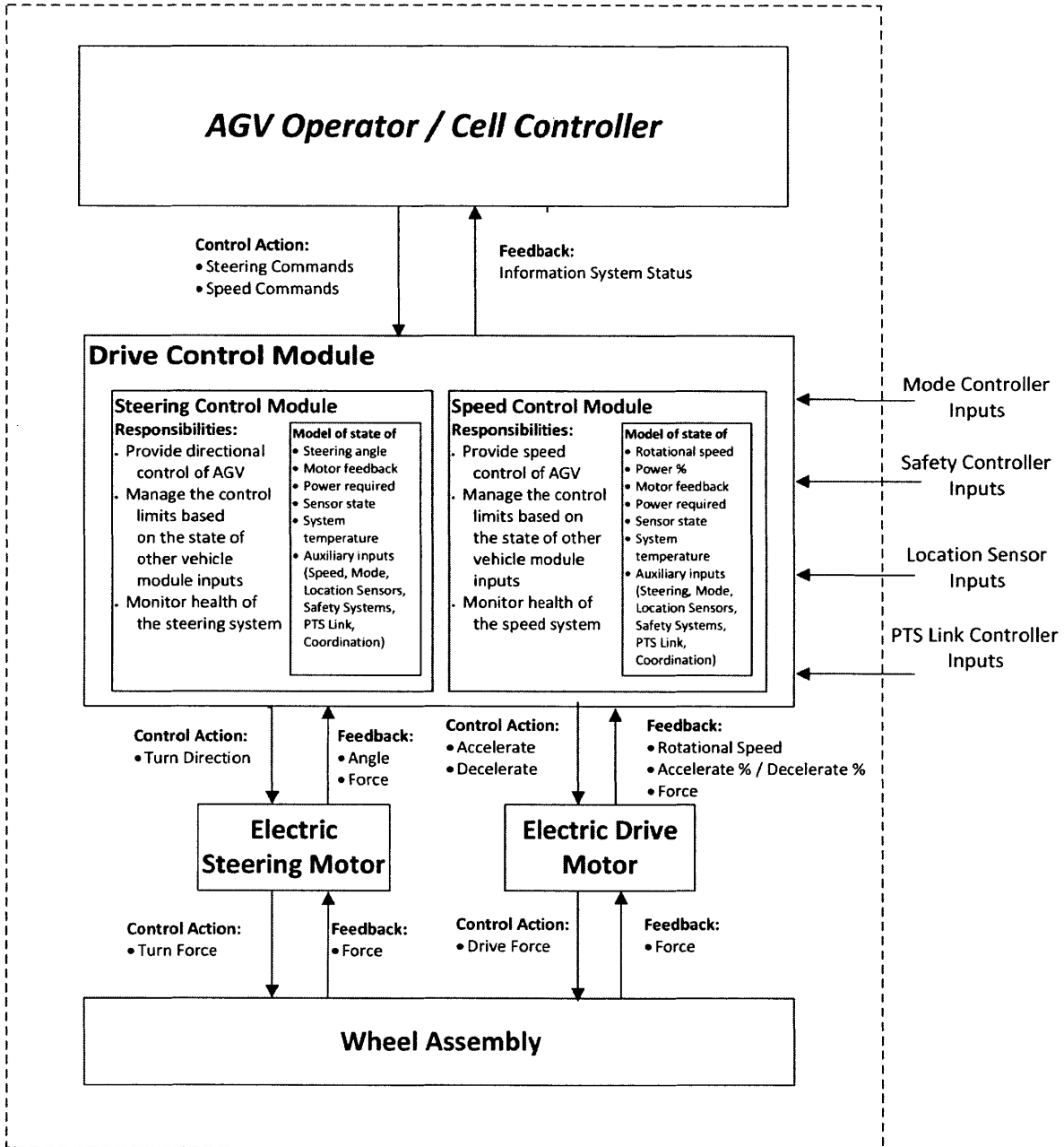


Figure 4-15: AGV SCS Segment

Control Action	Providing Causes Hazard	Not Providing Causes Hazard	Incorrect Timing/Order	Stopped Too Soon / Applied Too Long
Operator provides drive commands to drive control module	UCA1: Drive control module commanded to drive when the movement will violate minimum separation with an object [H1.1]	UCA3: Drive control module not commanded to drive when the movement will prevent a violation of minimum separation with an object [H1.1]	UCA4: Drive control module commanded to drive before or after a safe path direction [H1.1]	UCA7: Drive control module commanded to drive too long when the movement violates minimum separation with an object [H1.1]
	UCA2: Drive control module commanded to drive when a human is handling components that will move [H1]		UCA5: Drive control module commanded to drive before a human stops handling components that will move [H1]	
			UCA6: Drive control module commanded to drive after a human starts handling components that will move [H1]	

Table 4.3: Example Control Action Table for an AGV

Each of these items were evaluated and mapped to a system-level hazard if they were considered hazardous, as shown in the table. Each control action determined to be unsafe could then be written into a component-level safety constraint for the system.

During STPA Step 2 the team would determine how the control actions could potentially become hazardous. As previously noted, the potential causal factors annotated on the generic control loop in Figure 4-16 provided a framework for the team to consider.

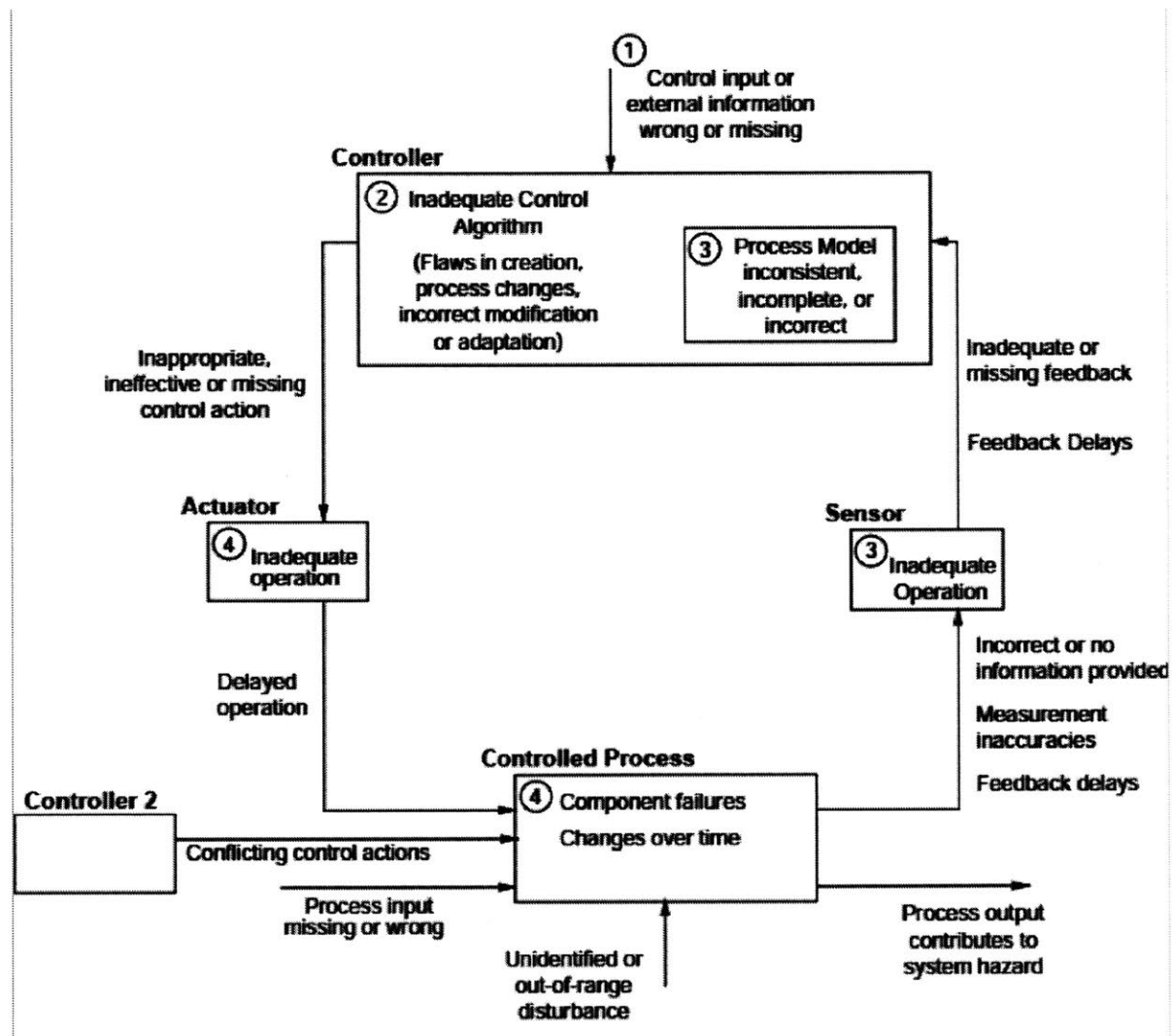


Figure 4-16: General Control Loop with Causal Factors

An example of the causal factors developed is provided below for the control action "Operator provides drive commands to the drive control module" in Table 4.3.

UCA1: Drive control module commanded to drive when the movement will violate minimum separation with an object. [H1]

Causal Scenario 1: The AGV is driven inappropriately because the operator is not familiar with its operation.

- New operator that is not or inadequately trained

Causal Scenario 2: The AGV is driven toward an external object because the operator does not see the object or misjudges the safe path.

- Operator inattention due to task overload, changes to the environment, or other external factors.
- Operating in cluttered/restrictive areas
- Objects are blocked from view by other workers, the vehicle itself, or the spar load on the AGV/PTV/AGV combination

Causal Scenario 3: The AGV is driven into an external object because the AGV speed/steering settings have changed.

- Modifications to the controller that the operator does not know about or is not familiar with and operating under their previous understanding.
- Degradation to the controller that the operator does not know about or is not familiar with and operating under their previous understanding.
- Modifications to the AGV system that the operator does not know about or is not familiar with and operating under their previous understanding.

Causal Scenario 4: The AGV is driven into an external object because the controller has a hardware failure.

- Input is frozen and continues to be communicated even though the operator changes the command.

Causal Scenario 5: The AGV is driven into an external object because the operator holds the drive command for too long due to a delay in the command starting movement in the vehicle.

- Processing delays and computing overloads
- No timeout of control inputs to AGV

Causal Scenario 6: The AGV is driven into an obstacle because the operator is provided incorrect or no information on critical system statuses relied upon for navigation, such as safety scanners and location sensors.

- Operator misses any feedback because it is not easily accessible, small HMI on side of the AGV, next to the ground.
- No status of scanners being on or off
- No status of location sensors being on or off

Causal Scenario 7: The operator drives the AGV into an obstacle that is not detected by the scanners. Possible causes include:

- The object is outside of the safety scanners field of view.
 - Obstacles in the factory at the PTV or spar level.
 - PTV guide rails above the AGV
 - AGV being used for unintended use, such as carrying objects that extend past the scanner FOV.
- The object is in the safety scanners field of view but below the detection threshold.
- The object enters the field of view and is impacted by the AGV at a rate faster than the scan and reaction rate of the vehicle.
- The scanner capabilities have degraded over time.
- The scanner fails into an unsafe state

Causal Scenario 8: The operator drives the AGV into an obstacle that is detected by the scanners but does not stop the AGV. Possible causes include:

- Software coding errors do not command the vehicle to stop as intended
- There is a delay in the signal processing due to CPU overload

- The interface with the PTV scanners is inadequate and the PTV scanners cannot send the signal to the AGV Safety Control Modules

Causal Scenario 9: The AGV operator drives the AGV into an external object when the scanners are deliberately turned off.

- The scanners were disabled and the operator did not know it, relying on the scanners to stop the AGV.
- The scanners on the PTV were disabled with no feedback to the AGV operator
- The operator did not see the object.
- Operator is relying on the location sensing module to avoid possible obstacles
- The operator believes the vehicle is in a different mode and expects the vehicle to respond differently
- The AGV has been modified and the operator does not know or understand the updates.
- The commands are delayed and the operator provides the command for too long

Once the team completed STPA Steps 1 and 2, the team was able to analyze the results with a thorough understanding of how the AGV or entire system could enter hazardous states and begin developing new controls for the system.

4.3.5 Discussion of STPA Process and Results

Assessment Environment

The analysis was completed over the course of a three-week period. A group setting for the assessment was only used twice during the process for less than three hours each, while the rest of the assessment was completed in one-on-one meetings between the facilitator and the respective expert for the portion of the system being discussed.

Because STAMP and STPA are relatively new for the company and entirely new for the team, the first group meeting was used as an instruction period to provide an overview of the methodology being used and then to define the accidents (or losses) the team wanted to avoid, the hazards and system boundaries, and an initial draft of the safety control structure

for this application. After developing the initial draft of the safety control structure, follow-on meetings were scheduled with each respective expert to add detail to the specific portions of the control structure. Because the methodology uses systems- and control-theory to build a model of the control structure, the assessment can be focused on specific loops within the system after the system architecture is built. This provided for easy scheduling of meeting times and numerous meetings within the available time frames of everyone within the group. During these meetings, details such as control actions, process models, feedback, etc. were added and any conflicts or issues were resolved with the affected parties.

Once the appropriate level of detail was developed for the systems safety control structure, the one-on-one or small group discussions continued with STPA Steps 1 and 2. The results of this portion of the analysis were then discussed in another group setting to confirm the individual findings of the group and discuss potential mitigations for the system. While this process occurred over the course of three weeks, the personnel hours of group, one-on-one meetings, and individual facilitator work totaled approximately 300 hours.

STPA Results

The results of the STPA analysis uncovered a number of issues with the system design and provided many insights and areas for the team to further investigate. A few abstract examples from the AGV are provided to illustrate some of the hazardous control actions and causal scenarios that were found through STPA. These examples are also used to demonstrate how STPA compares to the analysis assessment criteria. With regard to feasibility for use within the company, STAMP and STPA meet all mandatory government and voluntary industry safety regulations and standards. Additionally, STPA is a formalized and documented process used in many applications across many industries that can be used within any company.

Analyzing the results of STPA using the quality criteria provided a number of interesting results. Due to identifying all possible control actions and conducting STPA Step 1 of determining if control actions are hazardous, the team identified a number of hazardous control actions throughout the system. During Step 2 of the analysis, the team was able to show how these hazardous control actions could occur in different scenarios, such as a

hardware failure, incorrect process models, system interactions, and the role of management and procedures combined with schedule pressures in the factory. One example comes from the system's heavy reliance on safety scanners to prevent the AGV and product transportation vehicle from colliding with surrounding objects, including workers. If an object is in the path of motion of the system and is not detected by the operator or scanners, there are no other controls designed into the system to prevent the hazard or mitigate the effects of an accident. Due to the massive size, length, and weight of the system and the force of the AGV drive motors, any unsafe control action could result in a severe accident in certain conditions. While operator and workers in the area staying alert to AGV movement around them and AGV safety scanners were the current design control, there are a number of causal scenarios that show them as insufficient.

First, as already demonstrated in other parts of the company, workers may bypass safety scanners if they malfunction and/or are not calibrated or designed properly, preventing the workers from completing their tasks and falling behind on the production schedule. The safety scanners are not integrated in a way that prevents the AGV from operating when the scanners are not in use and there is no notification to the operator or surrounding workers that the AGV is operating with a disabled safety system. While it may not be an issue for the user that turns the safety system off in order to get a job done, if they do not turn the safety system back on the next user will be operating the AGV without knowing that the scanners are disabled.

Another scenario is if an AGV were to encounter an obstacle that the scanners did not protect against. The drive motors for the vehicle do not provide any feedback to the drive controller about actual force being encountered. This would result in the AGV continuing to apply full force into the obstacle until the AGV drive controller received a different drive command from the operator. An example of this could be something that is outside of the scan volume of the scanners, such as an obstacle above the vehicle but still in the path of the product on top of the transportation system, a scanner adjusted with a resulting gap in coverage around the vehicle, scanner control logic about clearing faults or obstacles when an obstacle is still in the scan volume, multiple controllers managing multiple scan systems with the integrated product transportation system, etc.

AGVs are large and powerful mobile systems that people cannot simply put a fence around to keep workers safe. With the numerous ways that safety scanners can be and already have been bypassed or fail to function as intended, the team discussed multiple new controls to implement. One was a design change to enable the ability for workers to turn the scanners off and provide feedback to everyone in the surrounding area that the scanners were disabled, such as warning lights or an audible tone. Additionally, they focused on ways to provide feedback from the AGV to the operator and surrounding workers in general. Most of the focus in the original design seemed to operate under assumptions that the AGV would operate exactly as originally designed and as the operator intended, without helpful and readily displayed feedback to know what the system is actually doing, not just what the operator believes they commanded or what surrounding workers believe the robot should be doing.

Another discussion was to program motor feedback logic for the drive motors of the AGV as seen in many applications of collaborative robotics. Due to the number of scenarios that resulted in the safety scanners being the only control in place and the number of ways the scanners may not maintain proper control, a secondary method of control for the mobile functions of the system was considered. While this control may still allow minimum separation to be violated as defined in the hazards, the AGV would stop moving and hopefully avoid an accident or lessen the severity of the accident.

While these are just a few examples of the discussion that resulted from the process and results of STPA, the results for the rest of the system were very similar. In order to avoid the defined system-level hazards, there were design changes recommended for instances of missing process model variables between systems that required coordination and control, for system motor feedback logic for various moving parts, and for increased levels of feedback from the physical system to the human operators and surrounding workers. The team was focused on determining how the system could violate the safety controls in place, potentially resulting in an accident, and then how those existing controls needed to change or what new controls were needed to be put in place.

Because STPA is based in systems-theory and one of the first steps of the process is to define the system-level hazards that the team wants to prevent through the enforcement of

safety constraints, the specific and overarching recommended safety control structure changes were used to compare the STPA results to the hierarchy of control. The recommended changes categorized according to the hierarchy of controls is show in Figure 4-17. The most frequent category of recommended changes to the system was engineering controls, which was expected since STPA is a methodology to ensure adequate control of system safety constraints to avoid hazards. The next highest category of recommended changes was in the administrative control classification due to numerous recommendations of additional feedback to the operator about the actual state of the system versus their commanded or perceived state of the system and to the surrounding workers who are not controlling the system. The lowest category of recommended changes remained design out.

STPA		Risk Reduction Measures	Examples	Influence on Risk Factors	Classification
14%	Most Preferred	Elimination or Substitution	<ul style="list-style-type: none"> Eliminate pinch points (increase clearance) Intrinsically safe (energy containment) Automated material handling (robots, conveyors, etc.) Redesign the process to eliminate or reduce human interaction Reduced energy Substitute less hazardous chemicals 	<ul style="list-style-type: none"> Impact on overall risk (elimination) by affecting severity and probability of harm May affect severity of harm, frequency of exposure to the hazard under consideration, and/or the possibility of avoiding or limiting harm depending on which method of substitution is applied. 	Design Out
			<ul style="list-style-type: none"> Barriers Interlocks Presence sensing devices (light curtains, safety mats, area scanners, etc.) Two hand control and two-hand trip devices 	<ul style="list-style-type: none"> Greatest impact on the probability of harm (Occurrence of hazardous events under certain circumstances) Minimal if any impact on severity of harm 	
29%	Least Preferred	Awareness Devices	<ul style="list-style-type: none"> Lights, beacons, and strobes Computer warnings Signs and labels Beeper, horns, and sirens 	<ul style="list-style-type: none"> Potential impact on the probability of harm (avoidance) No impact on severity of harm 	Administrative Controls
		Training and Procedures	<ul style="list-style-type: none"> Safe work procedures Safety equipment inspections Training Lockout / Tagout / Verify 	<ul style="list-style-type: none"> Potential impact on the probability of harm (avoidance and/or exposure) No impact on severity of harm 	
		Personal Protective Equipment (PPE)	<ul style="list-style-type: none"> Safety glasses and face shields Ear plugs Gloves Protective footwear Respirator 	<ul style="list-style-type: none"> Potential impact on the probability of harm (avoidance) No impact on severity of harm 	

Figure 4-17: STPA Risk Reduction Measures Hierarchy of Controls Classifications

Because the assessment process starts by building and detailing a safety control structure for the system, the team was forced to develop a thorough understanding of the system and expand the boundaries of their analysis to the factors that could cause the unsafe control actions, to include hardware failures, software and component interaction, the people involved and why they could cause an unsafe control action, work processes, management,

and the factory environment. The assessment focused on how to change the entire socio-technical system instead of accepting the physical system as is and trying to change the people, processes, and procedures around the physical system.

4.4 Comparison of the Analyses

While the currently used Task-based Risk Assessment and the proposed Systems-Theoretic Process Analysis are both intended to assess and improve workplace safety, they are based on two fundamentally distinct models that approach the problem using drastically different methodologies.

The Task-based Risk Assessment's iterative methodology and underlying assumptions resulted in repeatedly looking for multiple events sequenced as a forward chain over time, searching for a root cause solution, discounting systemic factors, and blaming accidents on some type of failure or human error without fully understanding the hazards to be mitigated. Additionally, the Task-based Risk Assessment methodology used the subjective criteria of severity, exposure, and avoidance to determine risk levels and determine appropriate levels of minimum mitigations allowed.

The examples from the assessment showed how the process brainstormed possible hazards for a task, such as crushing, pinching or impact; determined the hazard to be caused by unexpected movement of the robotic system; selected levels for the risk determination criteria of severity, exposure, and avoidance; mitigated the hazard through risk reduction measures, such as standard shop practices, procedures, and training, situational awareness, the use of safety scanners, and standing clear of the unexpected movement; and finally reevaluated the risk determination criteria for the hazard with these new measures in place to determine that the hazards were mitigated. As applied to assessment evaluation criteria, the traditional assessment fully met all of the feasibility criteria. On the other hand, the traditional assessment only met a few of the quality criteria fully and met the majority of the criteria marginally or not at all.

Systems-Theoretic Process Analysis's systems thinking approach to safety resulted in defining the accident and hazards the designers wanted to control against, defining system

boundaries and building a detailed functional control structure to analyze the control of the system hazards, and systematically analyzing the entire developed model using systems- and control-theory principles to determine potential causal factors for insufficient or missing controls and/or feedback. The examples discussed from the assessment detailed if a hazard such as violating minimum separation between an AGV and a person could happen; ways it could potentially occur in the current design, such as bypassed scanners, obstacles outside of the scanner field of view, or scanner control logic flaws; and recommended potential design changes to improve the control of the hazard, such as designing in the ability to turn scanners off, motor feedback logic, and helpful and accurate feedback to operators and surrounding workers. As applied to assessment evaluation criteria, the proposed assessment met all of the feasibility criteria except for the company internal standard that specifies the use of the ANSI recommended assessment. Furthermore, the proposed assessment met all of the applicable quality criteria fully.

A visual depiction of the assessment analysis can be seen in Figure 4-18 and 4-19.

Criteria	TRA	STPA
Feasibility		
Compliant to Government Safety Regulations and Standards	G	G
Compliant to Industry Safety Regulations and Standards	G	G
Compliant to Company Safety Regulations and Standards	G	Y
Documented Analysis Process	G	G
Quality		
Includes hardware failure accidents	G	G
Includes technological factors in accidents beyond hardware failures, such as system design and requirements flaws (software and component interactions)	Y	G
Includes the role of management, operations, and procedures in accidents	R	G
Includes the role of the environment in accidents	R	G
Goes beyond specifying what humans did wrong to explain why they did what they did (includes sophisticated human factors in analysis)	R	G
Creates thorough understanding of the problem before implementing controls IAW the hierarchy of controls	R	G
Cost - Time required for case study assessment (approximate total personnel hours)	1000	300

Figure 4-18: Assessment Analysis Criteria Results


	Risk Reduction Measures	Examples	Influence on Risk Factors	Classification	TRA	STPA
	Elimination or Substitution	<ul style="list-style-type: none"> Eliminate pinch points (increase clearances) Intrinsically safe (energy containment) Automated material handling (robots, conveyors, etc.) Redesign the process to eliminate or reduce human interaction Reduced energy Substitute less hazardous chemicals 	<ul style="list-style-type: none"> Impact on overall risk (elimination) by affecting severity and probability of harm May affect severity of harm, frequency of exposure to the hazard under consideration, and/or the possibility of avoiding or limiting harm depending on which method of substitution is applied 	Design Out	4%	14%
Guards, Safeguarding Devices, and Complementary Measures	<ul style="list-style-type: none"> Barriers Interlocks Presence sensing devices (light curtains, safety mats, area scanners, etc.) Two hand control and two-hand trip devices 	<ul style="list-style-type: none"> Greatest impact on the probability of harm (Occurrence of hazardous events under certain circumstance) Minimal if any impact on severity of harm 	Engineering Controls	34%	57%	
Awareness Devices	<ul style="list-style-type: none"> Lights, beacons, and strobes Computer warnings Signs and labels Bespeem, horns, and sirens 	<ul style="list-style-type: none"> Potential impact on the probability of harm (avoidance) No impact on severity of harm 	Administrative Controls	62%	29%	
Training and Procedures	<ul style="list-style-type: none"> Safe work procedures Safety equipment inspections Training Lockout / Tagout / Verify 	<ul style="list-style-type: none"> Potential impact on the probability of harm (avoidance and/or exposure) No impact on severity of harm 				
Personal Protective Equipment (PPE)	<ul style="list-style-type: none"> Safety glasses and face shields Ear plugs Gloves Protective footwear Respirators 	<ul style="list-style-type: none"> Potential impact on the probability of harm (avoidance) No impact on severity of harm 				

Figure 4-19: Risk Reduction Measures Hierarchy of Controls Classifications Comparison

While the ultimate goals of the methodologies are identical, the process and results from the two methodologies are drastically different. Task-based Risk Assessments result in a group of people iteratively mitigating individual hazards to prove that the system is safe. On the other hand, STPA results in a group of people developing a deeper understanding of what are accidents to avoid and hazards for the system and actively analyzing how they could possibly occur. Simply put, people using STPA are trying to identify how the system could be unsafe rather than trying to prove it is safe. This shift in mindset for the people involved results in drastically different results for the assessment. Feedback from the team involved in the assessment processes was that instead of accepting hazards as having unexpected or unintentional causes, the process of STPA helped them gain an understanding of the causes and the entire system while developing the control structure and determining what changes were possible to adequately control the hazards.[24] Due to this increased understanding of the system, the possible mitigations developed were now being considered at the system level, from each level of the hierarchy of control, within the specific situation, and in accordance with other company factors such as cost. This is increasingly important for the complex socio-technical systems that are being introduced to today’s manufacturing environment.

While the execution of a Task-based Risk Assessment may vary from team to team, it is a time-consuming process that requires the entire team to sit in a room and brainstorm as a group to complete the process as designed. It relies heavily on previous experience and knowledge of the process and results in a long list of subjective results that would be difficult for anyone to understand that was not involved in the assessment. This makes the assessment extremely difficult to verify or update over time as people change roles within a company.

STPA is still a labor-intensive process, but it puts much less strain on the organization and the people, promoting detailed and in-depth analysis of work that people in the organization are already responsible for. For example, the engineers responsible for the AGV engage in discussion about how the AGV interacts with the rest of the system and then only focus on the specifics of the AGV unless there is another system or component interaction discussion that involves their expertise. This allows the assessment to remain efficient and effective within a fast paced and highly complex organization, where employees have numerous competing tasks to complete on any given day.

THIS PAGE INTENTIONALLY LEFT BLANK

Chapter 5

Conclusions and Recommendations

The use of the Systems-Theoretic Accident Model Process (STAMP) tool called Systems-Theoretic Process Analysis (STPA) proved to be more effective and efficient than the traditional industry methods to conduct a hazard analysis for the semi-automated manufacturing process, as discussed in this thesis.

In order to reach this conclusion, the thesis began by outlining some of the traditional methods and respective issues currently in use to address risks and hazards, the traditional accident causation models and their underlying assumptions, and the traditional analysis techniques that stem from previous accident models. Then the thesis introduced a new accident model based on systems and control theory, called STAMP, and a hazard analysis methodology, called STPA. The current industry standard process was then outlined and a case study of a semi-automated manufacturing process at an aerospace company was used to apply the traditional, industry standard assessment and the STAMP STPA assessment. These assessments were analyzed using a list of criteria developed by company stakeholders and determined to be crucial for any workplace safety assessment to meet company safety goals.

Because of the changing nature of hazards being introduced and encountered on modern manufacturing floors with the introduction of automated systems, the safety and hazard analysis tools being used within a company provide a significant opportunity for improving workplace safety. Although some companies continue to use updated versions of traditional risk analysis methods that have historically proven to be effective for electro-mechanical

systems of the past, those analysis methods are beginning to reveal their limitations as they are applied to modern socio-technical systems. While the technology and engineering involved to make these systems possible has advanced and changed over the years, the processes and methodologies used in these companies to analyze and control safety risks have not. The STAMP model and its tools, such as STPA, are based upon a new framework of concepts applied to safety to address these complex risks in modern socio-technical systems.

Bibliography

- [1] E.E. Adams. Accident Causation and the Management System. *Professional Safety*, 21:26–29, 1976.
- [2] A. Amir and S. Weiss. *Aerospace Industry*. Encyclopedia Britannica, Inc., 2014.
- [3] American National Standards Institute (ANSI). *Introduction to ANSI*. <https://www.ansi.org/>, 2017.
- [4] W.R. Ashby. *An Introduction to Cybernetics*. Chapman & Hall Limited, 1957.
- [5] Robotic Industries Association. *American National Standard for Industrial Robots and Robot Systems - Safety Requirements (ANSI/RIA R15.06-2012)*, 2012.
- [6] Robotic Industries Association. *Task-based Risk Assessment Methodology (RIA TR R15.306-2014)*, 2014.
- [7] M.S. Herring B.D. Owens and N.G. Leveson. Safety-Driven Model-Based System Engineering Methodology Part II: Application of the Methodology to an Outer Planet Exploration Mission. Technical report, Massachusetts Institute of Technology, 2007.
- [8] F.E. Bird and R.G. Loftus. *Loss Control Management*. Institute Press, 1976.
- [9] J.C. Burnham. *Accident Prone : A History of Technology, Psychology, and Misfits of the Machine Age*. Chicago ; London : The University of Chicago Press, 2009.
- [10] L.A.T. Cox Jr. What’s wrong with Risk Matrices? *Risk Analysis*, 28(2):497–512, 2008.
- [11] US DoD. *MIL-STD-882E, Department of Defense Standard Practice System Safety*. U.S. Department of Defense, 2012.
- [12] H.A. Duckworth and R.A. Moore. *Social Responsibility : Failure Mode Effects and Analysis*. Industrial innovation series. Boca Raton, FL : CRC Press/Taylor & Francis, 2010.
- [13] C.H. Fleming. *Safety-driven Early Concept Analysis and Development*. PhD dissertation, Massachusetts Institute of Technology, Department of Aeronautics and Astronautics, February 2015.
- [14] M.C. Goodspeed. Job Analysis Reveals the Accident Causes. *National Safety News*, 22:32, 1930.

- [15] S.O. Hansson. Risk. In E.N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, Spring 2014 edition, 2014.
- [16] H.W. Heinrich. *Industrial Accident Prevention; A Scientific Approach*. McGraw-Hill Insurance Series. R. H. Blanchard, editor. New York, London, McGraw-Hill, 1931.
- [17] U. Khan and D.M. Kupor. Risk (Mis)Perception: When Greater Risk Reduces Risk Valuation. *Journal of Consumer Research*, 43(5):769–786, 2017.
- [18] N.G. Leveson. *SafeWare : System Safety and Computers*. Reading, Mass. : Addison-Wesley, 1995.
- [19] N.G. Leveson. A New Accident Model for Engineering Safer Systems. *Safety Science*, 42:237–270, 2004.
- [20] N.G. Leveson. *Engineering a Safer World : Systems Thinking Applied to Safety*. Engineering systems. Cambridge, Mass. : MIT Press, 2011.
- [21] R.R. Lutz. Analyzing Software Requirements Errors in Safety-critical, Embedded Systems. *Proceedings of the IEEE International Symposium on Requirements Engineering*, 1993.
- [22] M. Mesarovic and Y. Takahara. *Theory of Multilevel Hierarchical Systems*. Academic Press, New York, NY, 1970.
- [23] National Institute of Occupational Safety and Health (NIOSH). *Hierarchy of Control*. <https://www.cdc.gov/niosh/topics/hierarchy/>.
- [24] N.A. Peper. Interviews with company personnel, 2016.
- [25] J. Rasmussen, K.D. Duncan, and J. Leplat. *New Technology and Human Error*. New Technologies and Work. Chichester ; New York : J. Wiley, 1987.
- [26] N.C. Rasmussen. Reactor Safety Study: An Assessment of Accident Risks in U.S. Commercial Nuclear Power Plants. Technical Report WASH-1400; NUREG-75/014, Nuclear Regulatory Commission, Washington, D.C. (USA), 1975.
- [27] M. Rausand and A. Høyland. *System Reliability Theory: Models, Statistical Methods, and Applications*. Wiley series in probability and statistics. Applied probability and statistics. Hoboken, N.J. : Wiley-Interscience, 2004.
- [28] J.T. Reason. *Human Error*. Cambridge [England] ; New York : Cambridge University Press, 1990.
- [29] R. Griego S. Friedenthal and M. Sampson. IncoSe Model Based Systems Engineering (MBSE) Initiative. In *INCOSE 2007 Symposium*, 2007.
- [30] Occupational Safety and Health Administration. *OSHA 3071: Job Hazard Analysis*. U.S. Department of Labor, 2002.

- [31] N.B. Sarter and D.D. Woods. How in the World Did We Ever Get into That Mode? Mode Error and Awareness in Supervisory Control. *Human Factors*, 37(1):5–19, 1995.
- [32] W.R. Spiegel, C.E. Myers, F.B. Gilbreth, and L.M. Gilbreth. *The Writings of the Gilbreths*. Irwin Series in Industrial Engineering and Management. Homewood, Ill., R. D. Irwin, 1953.
- [33] W.E. Vesely. *Fault Tree Handbook*. Washington, D.C. : Systems and Reliability Research, Office of Nuclear Regulatory Research, U.S. Nuclear Regulatory Commission, 1981.
- [34] Watson, H. A. *Launch Control Safety Study*. Bell Laboratories: Murray Hill, NJ, 1961.
- [35] R.F. Zammuto and E.J. O'Connor. Gaining Advanced Manufacturing Technologies' Benefits: The Roles of Organization Design and Culture. *Academy of Management Review*, 17(4):701–728, 1992.