Ioana V. Koglbauer
Sonja Biede-Straussberger
(Editors)

With a foreword by
Gunnar Steinhardt
(President of EAAP)

Aerospace Psychology and Human Factors

Applied Methods and Techniques

hogrefe

Aerospace Psychology and Human Factors

Aerospace Psychology and Human Factors

Applied Methods and Techniques

Ioana V. Koglbauer and
Sonja Biede-Straussberger

(Editors)

hogrefe

Library of Congress of Congress Cataloging in Publication information for the print version of this book is available via the Library of Congress Marc Database under the Library of Congress Control Number 2024940701

Library and Archives Canada Cataloguing in Publication
Title: Aerospace psychology and human factors : applied methods and techniques / Ioana V. Koglbauer and Sonja Biede-Straussberger (eds.).
Names: Koglbauer, Ioana V., editor. | Biede-Straussberger, Sonja, editor.
Description: Includes bibliographical references.
Identifiers: Canadiana (print) 20240403231 | Canadiana (ebook) 20240403274 | ISBN 9780889376472 (softcover) | ISBN 9781616766474 (PDF) | ISBN 9781613346471 (EPUB)
Subjects: LCSH: Aeronautics—Human factors. | LCSH: Aviation psychology.
Classification: LCC TL553.6 .A38 2024 | DDC 629.1301/9—dc23

www.hogrefe.com

PUBLISHING OFFICES

USA:	Hogrefe Publishing Corporation, 44 Merrimac St., Newburyport, MA 01950 Phone 978 255 3700; E-mail customersupport@hogrefe.com
EUROPE:	Hogrefe Publishing GmbH, Merkelstr. 3, 37085 Göttingen, Germany Phone +49 551 99950 0, Fax +49 551 99950 111; E-mail publishing@hogrefe.com

SALES & DISTRIBUTION

USA:	Hogrefe Publishing, Customer Services Department, 30 Amberwood Parkway, Ashland, OH 44805 Phone 800 228 3749, Fax 419 281 6883; E-mail customersupport@hogrefe.com
UK:	Hogrefe Publishing, c/o Marston Book Services Ltd., 160 Eastern Ave., Milton Park, Abingdon, OX14 4SB Phone +44 1235 465577, Fax +44 1235 465556; E-mail direct.orders@marston.co.uk
EUROPE:	Hogrefe Publishing, Merkelstr. 3, 37085 Göttingen, Germany Phone +49 551 99950 0, Fax +49 551 99950 111; E-mail publishing@hogrefe.com

OTHER OFFICES

CANADA:	Hogrefe Publishing Corporation, 82 Laird Drive, East York, Ontario, M4G 3V1
SWITZERLAND:	Hogrefe Publishing, Länggass-Strasse 76, 3012 Bern

Format: PDF

ISBN 978-0-88937-647-2 (print) • ISBN 978-1-61676-647-4 (PDF) • ISBN 978-1-61334-647-1 (EPUB)
https://doi.org/10.1027/00647-000

Dedication

To my family, Reinhard, Alina and Dan, for their love and support

Ioana V. Koglbauer

To my daughter, Mina, who inspires with courage and strength

Sonja Biede-Straussberger

Acknowledgments

The editors would like to thank the following organizations and individuals:

For Supporting This Book Project

Gunnar Steinhardt, President of the European Association for Aviation Psychology (EAAP), and the Board of Directors of the EAAP – Renée Pelchen-Medwed, Mickaël Causse, Julia Behrend, Jennifer Eaglestone, Robert Bor, and Jóhann Wíum

For Expertly Reviewing Parts of This Book

René Amalberti, Ciprian Baciu, André Droog, Renée Pelchen-Medwed, and Peter Sandl

Each chapter of this book was reviewed by three independent peer reviewers.

For Their Assistance From Contracting to Publishing and Marketing

The dedicated team at Hogrefe Publishing

Contents

Foreword

Our professional community is dedicated to developing and sharing knowledge beyond our own domains, enriching disciplines and ensuring we anticipate future challenges to get it right from the onset when crafting new designs or procedures. As the aerospace sector encounters new technological and societal challenges affecting operators and passengers, it is essential for practitioners and scientists to collaborate, refining the integration of the human element into the overarching sociotechnical system. This book, co-edited by Ioana Koglbauer and Sonja Biede-Straussberger and featuring contributions from experts in academia, industry, and international agencies, marks a significant step in advancing the human aspect of aerospace. It delves into both present and future methodological trends in aviation psychology and human factors. This volume fosters interdisciplinary learning and collaboration, essential for effective human performance management. It does so by offering discussions on research methods, practical "how to do it" guidance, insights from past experiences, and projections of future trends. The ethos of this book portrays the spirit of the European Association for Aviation Psychology (EAAP), which aims at promoting applied psychology and human factors in aviation, ensuring the dissemination of information and experience. Everyone, whether newcomers or seasoned experts from academia, industry, or government, interested in the human-centric approach in aerospace systems design and operation, will find invaluable insights and guidance in these chapters.

Gunnar Steinhardt
President of the European Association for Aviation Psychology

Preface

How to Put People First in the Design and Operation of Aerospace Systems

Nearly 80 years ago, human factors became an area of interest in aviation. Since then, the world has rapidly evolved: Changes have occurred, knowledge has improved, experience has grown, society has changed, and new technologies have been invented. As people and organisations involved in aviation and space dream bigger and as the technical possibilities develop at a fast pace, focus needs to be maintained on integrating the human element in the system. More than ever, the maintenance of and even the increase in the current level of safety are of utmost importance.

Despite all these advances, human factors and aerospace psychology professionals still need to strive for the integration of the human element throughout business, development and operations. Especially in such a complex system as that of aerospace, we need to ask ourselves whether we are solving the right problems. Once the right problems are identified, the next question is how to solve them in the right way. Which industry standards are suitable and applicable? Where are the gaps? Which scientific methods can help bridge the gap between the status quo and future performance expectations? Is the human element appropriately addressed in each stage of a system's life cycle? Are interdisciplinary perspectives convergent and harmonised?

Thus, as psychologists and human factors specialists who drive and enable these innovations, we are often confronted with new questions that cannot be answered by conventional means. Sometimes we need to adapt or develop new methods or tools to address them. In this book experts working in the industry and in academia share methods and techniques of aerospace psychology and human factors that are currently used in research and development. Thus, our intention with this new book is to provide a wide range of methods, techniques and tools for promoting the application of aerospace psychology and human factors. All of this serves to build better products for operators. These operators want to be efficient in their tasks. Their objective is to deliver safe and efficient operations.

Several chapters of this book try to grasp the role of human factors from a more global perspective, such as describing the current practice in specific organisations, whereas others zoom in on addressing specific problems, such as how to capture human performance. At the beginning, design methods are addressed. The first chapter provides a systems-theoretic perspective and a method for modelling emergent system properties in existing systems or in those that are in development.

Professionals looking for a powerful and efficient tool to identify the right problems in a system and to address them will find it here. In the next chapter, the use of a cockpit philosophy is highlighted to support the transition from initial concepts to detailed designs that justify certification requirements. A special chapter is dedicated to human factors challenges in cabin design for commercial aircraft. New questions on how to design assistance are addressed in a different chapter, a topic which has rapidly spread in multiple industries over recent years.

The design environments have a number of different aerospace psychology and human factors topics in common that are addressed in the following sections of the book. The introduction of new artificial intelligence technology poses new challenges and requires new solutions on how a system can explain information to operators. Combined measures of workload and situation awareness are integrated in a model to support the assessment of performance from a team's perspective. Operators' attention and awareness are addressed in the context of pilot monitoring, and the benefits and drawbacks of the current eye-tracking technology are analysed for both design and training. Furthermore, techniques to improve fatigue risk management systems by adding additional parameters to identify and monitor risks are presented. These are expanded with a chapter on hazards related to human space flight and methods to analyse them. Another chapter is dedicated to the development of a free flying virtual companion for an astronaut and methods to implement various humanlike features in the area of tension between the machine and the "uncanny valley."

Another section of the book is dedicated to the use of virtual, augmented, or mixed reality technologies that found their way into daily aviation business. They are studied in depth to investigate how they may better support operations and training, in application fields from cockpit to air traffic control or even maintenance. A number of chapters cover human factors methods and techniques related to this issue. A special chapter is dedicated to virtual reality applications for developing and testing the Argonaut Lunar Lander. An additional chapter addresses applications of extended reality for studying human behaviour in immersive conditions, manipulating mental workload, prototyping, and evaluating complex interfaces. In addition, challenges of virtual reality and techniques to overcome them are presented in the context of pilot training. This section is rounded up by a chapter on methods to prioritise and implement augmented reality-based innovations for pilot training in a sustainable manner.

The final section of the book includes methods and techniques that provide a broader view of how to systematically learn from past research and to plan future developments. Thus, the method for conducting a meta-analysis is explained, an approach that will be more frequently used to gain knowledge by aggregating results of a large number of studies. In a different chapter, a method for integrating the assessment of human readiness level in the Single European Sky Air Traffic Management Research (SESAR) is presented. The final chapter takes a look at where a major aircraft manufacturer stands in the process of integrating human

and organisational factors throughout the organisation along key principles to be taken into account (e.g., competencies) and anticipates the impact of new technologies and a changing society.

The sociotechnical aerospace system has rapidly evolved and continues to change, as we see in the current sociopolitical context. New challenges will emerge that are far from being anticipated today. Whatever those challenges will be, our strongly connected and interdisciplinary community of professionals will strive to put the human at the centre and do their best for society.

Sonja Biede-Straussberger & Ioana V. Koglbauer

Chapter 1

Integrating Human Factors Into the System Design Process

Brittany Bishop, Pauline Harrington, Nancy Leveson, and Rodrigo Rose

Abstract

Hazard analysis is the basis of engineering for safety. However, in such analyses, human factors are often oversimplified as simply "human failure," disregarding the systemic issues that lead to flawed decisions. A new, more powerful hazard analysis technique, called "system-theoretic process analysis" (STPA), combines sophisticated human factors, hardware design, software design, and even social systems in one integrated model and analysis. STPA can be used to identify conditions and events that can lead to an accident or mission loss so that designs can prevent or minimize losses. Safety assurance is typically carried out separately from system design and in later stages of development. By the time these assurance processes are used, it is often too late to effectively modify a system to address any safety issues that are found. STPA assists in overcoming these problems when used by an integrated team of engineering specialists, including human factors experts, to identify potential scenarios leading to unsafe behavior starting from the beginning of the design process.

Keywords

aviation psychology, human factors in system engineering, system safety engineering, STPA

The Goal

Hazard analysis is the foundation of engineering for safety. It is used to identify the hazards, which are defined as system states or sets of conditions that, together with a particular worst-case environment, will lead to a loss (Leveson, 2012). Once identified, this information can be used in system development and operations to eliminate these hazards or, if that is not possible, to reduce their likelihood or to minimize their potential impact. Unfortunately, the complex software-intensive systems being built today cannot be fully analyzed using traditional hazard analysis techniques. In addition, human contributions to risk have traditionally been

oversimplified by engineers in the hazard analysis process, thus limiting the usefulness of the hazard analysis process in reducing overall system risk.

The role of humans is changing as our systems become increasingly automated. Rather than directly controlling a potentially dangerous system, operators today are more often supervising automation and taking over in the cases where automation is not able to cope. It is no longer useful to only look at simple human mistakes in reading a dial or operating controls. The cognitively complex activities in which operators are now engaged do not lend themselves to simple failure analyses.

At the same time, some systems are designed such that a human error is inevitable, and then the loss is blamed on the human rather than on the system design (Leveson, 2019). Hazards may result from automation design that induces erroneous or dangerous operator behavior. Sometimes interface changes can alleviate these human errors, but often interface design fixes alone are not enough.

Human–machine *interactions* are greatly affected by the design of both the software and the hardware in concert with the design of the activities and functions provided by the operator. Changing the software, hardware, and human activities is the most direct and effective way to eliminate interaction problems as opposed to simply changing the interface between the human operator and the rest of the system. To reduce risk most effectively, the design or redesign of the functionality of the software and hardware and of the activities assigned to the operator and to the automation is needed rather than merely the design or redesign of the displays and controls.

In addition, today's complex, highly automated systems argue for the need for integrated system analyses and design processes. In the analysis and design of complex systems, it is not enough to separate the efforts in hardware design, software design, and human factors. Successful system design can only be achieved by engineers, human factors experts, and application experts working together. Obstacles to this type of collaboration stem from limitations in training and education, the lack of common languages and models among different specialties, or an overly narrow view of one's responsibilities. These obstacles need to be overcome to successfully build safer systems. This chapter presents an approach involving new modeling and analysis tools that will allow all the engineering specialties to use common tools and work more effectively together.

An overriding assumption in this chapter is the systems theory principle that human behavior is impacted by the design of the system in which it occurs. If we want to change operator behavior, we have to change the design of the system in which the operator is working. For example, if the design of the system is confusing the operators, (1) we can try to train the operators not to be confused, which will be of limited usefulness, (2) we can try to fix the problem by providing more or better information through the interface, or (3) we can redesign the system to be less confusing. The third approach will be the most effective.

Simply telling operators to follow detailed procedures that may turn out to be wrong in special circumstances or relying on training to ensure they do what they "should" do – when that may only be apparent in hindsight – will simply guarantee that unnecessary accidents will occur. The alternative is to ask how we can design to reduce operator errors or, conversely, identify what design features induce human error. In other words, we must design to support the operator.

A New Foundation for Integrated System Analysis

Achieving this goal will require new modeling and analysis tools. Traditional hazard modeling and analysis techniques do not have the power to handle complex systems today. They are based on a very simple model of causality that assumes accidents are caused by component failures. A new model of accident causality, called the "system-theoretic accident model and process" (STAMP), comprises more complex types of causal factors, including interactions among system components and including the operators (Leveson, 2012). In this enhanced model of causality, accidents may result from unsafe interactions among components that may not have "failed." In other words, each system component satisfies the specified requirements but the overall system design is unsafe. For example, the software and hardware satisfy their specified requirements and the operators correctly implement the procedures they were taught to use.

As an example, consider the crash of a Red Wings Airlines Tupolev (Tu-204) aircraft that was landing in Moscow in 2012. A soft touchdown made runway contact a little later than usual. There was also a crosswind, which meant that the weight-on-wheels switches did not activate. Because the software did not think that the aircraft was on the ground and because it was programmed to protect against activation of the thrust reversers while in the air (which is hazardous), the command of the pilot to activate the thrust reversers was ignored by the software. At the same time, the pilots assumed that the thrust reversers would deploy as they always do, and quickly engaged high engine power to stop sooner. Instead, the pilot command accelerated the aircraft forward, eventually colliding with a highway embankment (Leveson & Thomas, 2018).

Note that nothing failed in this accident. The software satisfied its requirements and behaved exactly the way the programmers were told it should. The pilots had no way of knowing that the thrust reversers would not activate. There were no hardware failures. The software performed exactly as it was designed to do. The humans acted reasonably. In complex systems, human and technical considerations cannot be isolated.

These types of accidents are enabled by the inability of designers and operators to completely predict and understand all the potential interactions in today's tightly

coupled and complex systems. That is, the error is in the overall system design and how the system components interact and not in the individual components. These types of accidents, which are increasingly occurring in today's complex systems, cannot be handled with the traditional linear causality model and hazard analysis techniques.

STAMP, by contrast, treats safety as a control problem rather than a failure problem. In other words, accidents result when the system design does not control hazardous system states. Those hazardous states may result from component failures, but they may also arise from overall system design flaws.

In this chapter, we describe and illustrate a new hazard analysis technique, called "system-theoretic process analysis" (STPA), which is built on STAMP and is more powerful than the traditional hazard analysis techniques. STPA is a structured step-by-step process for identifying the ways that hazards can occur in a system. It integrates hardware, software, and human factors into one modeling and analysis process and enhances the design process by allowing for shared modeling and analysis efforts (Leveson, 2012; Leveson & Thomas, 2018).

The Concept of Control in Safety

As noted, STAMP treats safety as a control problem. The system design must control both the component failures and the unsafe interactions among the components.

STPA uses a simple model of control in the form of feedback control loops. Such a loop is illustrated in Figure 1.1. At the top of the figure is the operator, who provides commands to automation as well as, in some cases, directly to the controlled process. For example, the driver of the vehicle may issue acceleration and braking commands. The operator gets feedback about the state of the controlled process directly (e.g., by feeling or seeing the vehicle slow down or accelerate) or through electronic displays. Even with highly automated systems, human operators often get feedback in addition to that provided by the displays, such as from sound, vibration, etc., which cannot easily be communicated through an electronic interface.

Figure 1.1 shows two components within the operator box. A human mental model contains the information the operator uses to make control or monitoring decisions. The mental model contains what the operator *thinks* is the current state of the automation (e.g., the brakes are being activated), the controlled process (the aircraft is slowing or accelerating), and relevant parts of the environment. The mental model is updated by various means, but primarily from feedback. Other information operators may use to update this mental model include beliefs they have about how the process can change and inferences about the effect of previous commands the operator issued to the automation – and assumes were executed

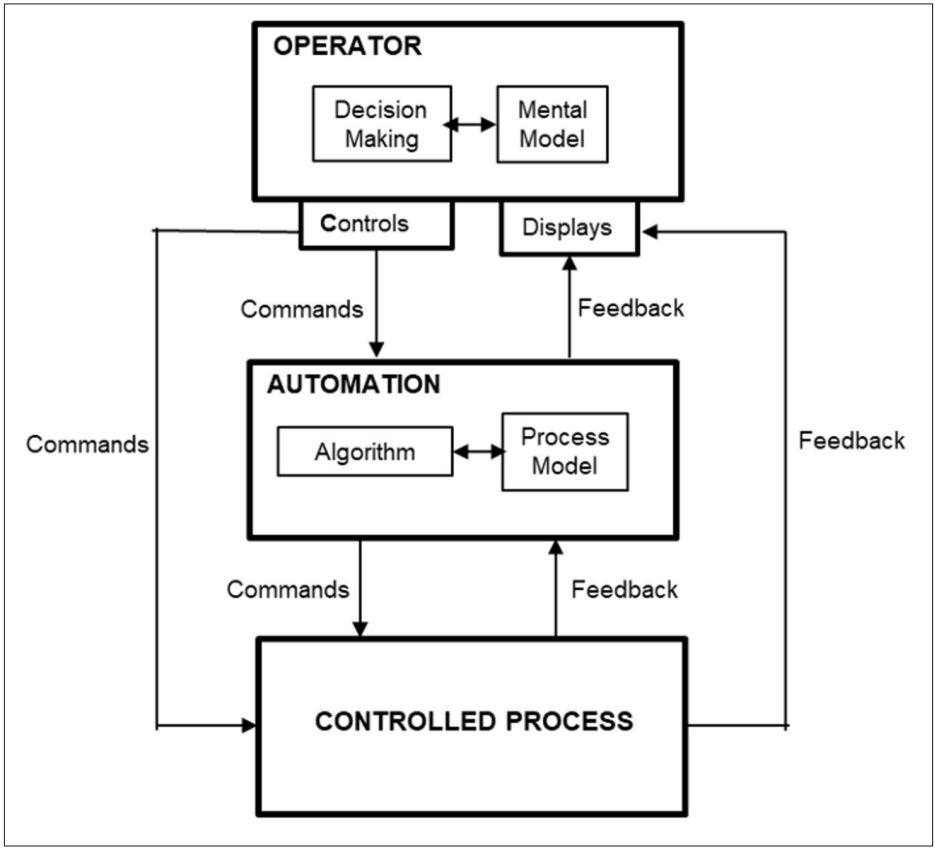

Figure 1.1. A basic model of system control.

correctly. An example of the latter is the belief that the thrust reversers would activate on the Red Wings aircraft mentioned earlier because the operator had commanded the software to activate them.

Note that computer automation also has a model of the state of the process. This model is usually much simpler than human mental models and may simply be represented as a few variables in the memory of the computer or in the software algorithm.

The automated controllers update their process models through direct feedback from sensors in the system and through human controller input. For example, an altimeter tells the automated controller the altitude of an aircraft, and a pilot may tell the automated controller what the desired altitude is.

Human controllers update their mental models of the controlled process, the environment, and the automation through direct feedback they receive from the system (i.e., displays, alerts, observed system behavior, etc.). Updates to the mental model of the environment can also occur through direct stimuli from visual, auditory, or vestibular systems, such as a pilot seeing clouds through the

windshield. Any of these models can also be updated by information from another human controller in the system, such as the copilot in an aircraft. In the control structure, these relationships will be modeled as arrows flowing into the human and automated controllers.

In basic feedback control loops like the one shown in Figure 1.1, feedback from the controlled process is used by the controller of that process to adjust the system's behavior to achieve the system goals and avoid hazards. In this way, the feedback received by the controller is used to guide decision-making for future control actions.

In a feedback control loop, the actions available to controllers to manage the process are termed "control actions" (commands) and are represented by downward arrows (see Figure 1.1). *Feedback*, which is used to inform the decisions about these actions, is represented by upward arrows. For example, an automated cruise control system on a car might have "accelerate" and "brake" as control actions. The car determines which action to take based on feedback from sensors about the car's current speed and from commands by the human operators about the desired speed.

Each controller uses their process model to make decisions about the changes they need to enact on the controlled process. To ensure that each controller's decision-making process is adequately informed, the process model(s) of the human controller needs to match the process model(s) of the automation, both of which need to match the reality of the system and environment. If these models do not match, the control actions coming from any of these controllers may become unsafe. If the pilots think the aircraft is not in a stall, they will not behave properly regarding the stall.

The safety of control actions depends on the context in which the actions occur, namely, the state of the overall system and its environment. Mismatched process models between controllers or misunderstanding of the context for a particular action can lead to unsafe control actions. Therefore, the human controller should understand what control actions are safe or unsafe in each context. This understanding may come from the human controller's prior experience, training, or any additional resources that they can consult, such as manuals.

Accidents often happen when the operator's mental model or the automation's process model become inconsistent with the real state of the controlled process and the environment. For example, the driver or the vehicle automation thinks that the lane to the left is clear, when it is not, and moves into that lane. Another example is that the human operator or the automation thinks that the helicopter state is fine when, in fact, some equipment is overheating and a control action is required to prevent an accident.

This chapter uses the example of mode confusion to illustrate the new STAMP-based design tools. In mode confusion, the human controller's process model about the mode of the automation and/or the controlled process does not match the actual mode. Two potential examples are:

- The human operator believes the system is in mode A, when the system actually is in mode B.
- The human operator knows the system is in mode A, but does not know the implications of mode A on the state of the system.

Frequently, accidents are related to such mode confusion, that is, well-trained controllers believe that they are making the right decision to maintain safe operations because they are confused about the current mode of the aircraft or automation. One reason such confusion may occur is that the automation changes the aircraft mode without any inputs to do so by the pilots. As an example of such indirect mode changes, an A320 crashed while landing at Bangalore, India, in 1990. The pilot selected a lower altitude while the automation was in the *altitude acquisition* mode. This command resulted in the activation of the *open descent* mode, where speed is controlled only by the pitch of the aircraft and the throttles go to idle. In that mode, the automation ignores any preprogrammed altitude constraints. To maintain the pilot-selected speed without power, the automation had to use an excessive rate of descent, which led to the aircraft crashing short of the runway (Sarter & Woods, 1995).

How could this happen? There are several different ways to activate *open descent* mode without the pilot directly commanding it. The investigators suspected that the inaccurate pilot mental model resulted from the automation design that activates *open descent* mode when a pilot selected a lower altitude while in *altitude acquisition* mode. The pilot must not have been aware the aircraft was within 200 ft of the previously entered target altitude, which triggers *altitude acquisition* mode and thereafter *open descent* mode. He therefore may not have expected selection of a lower altitude at that time to result in a mode transition and did not closely monitor his mode annunciations during this high workload time. He discovered what happened 10 s before impact, but that was too late to recover with the engines at idle (Sarter & Woods, 1995).

Accident investigators often blame operators in such cases for poor decision-making or blame a loss of situational awareness (Leveson, 2012). However, neither of these bring us closer to preventing future incidents. Redesign of the interface is also often not the right solution. Instead, redesign of the automation may be more effective.

In the STAMP model terminology, mode confusion occurs when one or more controllers have different models of system status and behavior (Leveson & Palmer, 1997). This is occurring more often as operators move from active control roles to monitoring (Leveson et al., 1998), as is common in new systems with primarily automated controls. Previous studies by Sarter and Woods (1995), Leveson et al. (1998), and Bredereke and Lankenau (2005) have examined and described the cognitive processes that underlie mode confusion. This chapter shows how the STPA process enables the analyst to identify sources of mode confusion in a specific system design and to generate recommendations for improvement.

The rest of this chapter presents an example of the use of STPA in the identification and prevention of potential mode confusion in the autopilot design of a Boeing 777 aircraft. This example is adapted from Bishop et al. (2023).

STPA and an Example of Its Use

STPA consists of four basic steps to identify why a particular system might behave in a hazardous manner and what requirements should be implemented to prevent losses (see Figure 1.2). Leveson and Thomas's *STPA Handbook* provides a detailed guide on how to properly follow these steps (Leveson & Thomas, 2018). A brief overview will be provided here to explain the basic process and illustrate its use with respect to mode confusion.

The first step in STPA is to define the purpose of the analysis.

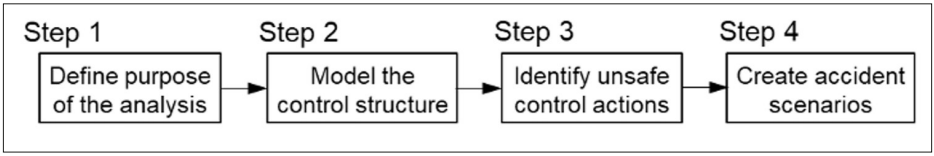

Figure 1.2. The four steps in STPA.

Step 1: Identifying the Goals of the Analysis

The first step in any engineering activity is to identify the goals or purpose. This step involves identifying the system and its boundaries, the potential system-level losses and hazards, and the necessary constraints of system behavior to avoid those hazards.

A *system* is the set of components that work together to accomplish specific objectives. The system and its boundaries must be defined to clearly understand which design aspects can be controlled to prevent hazards. The *boundary* separates the environment, which is *not* under the control of the designers, from the entities within the system, which *are* under the control of the designers (Leveson, 2012). For instance, an aircraft designer might define their aircraft as the system but consider airport infrastructure as part of the environment and thus not under the control of the designer.

In safety, the goal is to prevent losses. A *loss* involves anything of value to stakeholders. Examples of losses include loss of life or injury to people, loss of or damage to the system, loss of or damage to objects outside of the system, loss of mission, and even loss of reputation.

The Boeing 777 features an autopilot (A/P) with various pitch, roll, and thrust modes to manage the speed and direction of flight. Selected pitch and roll modes may impact the set thrust mode and vice versa. Thus, it is imperative for pilots to know the current mode of the autopilot and the consequences of changing that mode. The pitch modes are managed by the autopilot flight director system (AFDS). Additionally, pilots can engage autothrottle (A/T) to have the autopilot manage thrust. The current modes for the AFDS and A/T are displayed in the flight mode annunciator (FMA), a rectangle at the top of the primary flight display.

For commercial aircraft, the highest priority losses to be prevented commonly include *loss of life involving passengers or crew, destruction or damage to the aircraft,* and *loss of mission.*

After defining the unacceptable losses, the system-level hazards are identified. The *system hazards*, as defined previously, are the system states that will lead to a loss given a particular set of worst-case environmental conditions. Hazards refer to the overall system and not to individual components. Hazards identified by an aircraft designer could include the aircraft coming too close to terrain or losing controllability. To narrow the example to one that can be included in this chapter, we select the hazard as *H-1: Loss of control of the aircraft*.

After generating losses and hazards, the safety constraints for the system are defined. Safety constraints are simply statements of what the system should not do. Traceability remains a key component throughout as constraints are linked to hazards that are connected to losses. The safety constraint here is simply that the aircraft must always be controllable.

Step 2: Creating a Model of the Control Structure

Figure 1.3 shows a simplified control structure for the Boeing 777 autopilot system. In this case the human controller is the pilot, the automated controller is the autopilot, and the controlled process/system is the aircraft itself. The control actions and feedback lines identified in this figure are not meant to be exhaustive but are sufficient to generate the UCAs and scenarios in the example shown in the following sections. Note that this model does not contain design details and could be constructed early in the system design process. That would make it possible to generate a safe design from the start without having to undo earlier design decisions.

Steps 3 and 4: Identifying Unsafe Control Actions and Scenarios

In the third step of STPA, users identify *unsafe control actions* (UCAs). A UCA is a control action that will lead to a hazard given specific worst-case conditions.

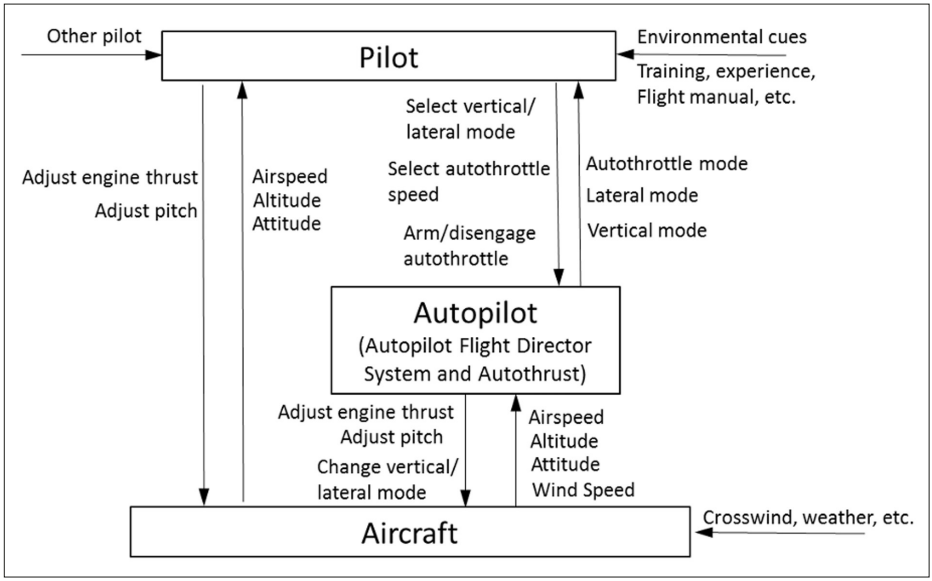

Figure 1.3. Simplified control structure for the Boeing 777 autoflight system.

There are four ways in which a UCA can occur: (1) not providing the control action leads to a hazard; (2) providing the control action leads to a hazard; (3) providing a control action too early, too late, or in the wrong order leads to a hazard; and (4) the control action lasts too long or is stopped too soon, which leads to a hazard. For example, a control action might be applying the brakes in a car. A driver could: (1) not apply the brakes when an obstacle is in front of the car; (2) apply the brakes when there are cars close behind; (3) apply the brakes too late to fully stop; or (4) apply the brakes for too short of a time to decelerate to a safe speed.

The fourth and final step of STPA is to identify potential loss scenarios by analyzing the causal factors that would lead to UCAs. In other words, identify the reasons that a UCA might be taken. Among other things, this step involves asking why a controller would reasonably take a UCA. One possible reason (involving mode confusion) is that they misunderstand the true mode of the controlled process or automated controller and issue a UCA as a result.

Within the STPA framework, controllers' choices of control actions are understood through their process/mental models. Inadequate mental/process models may occur when controllers receive incorrect feedback; they receive feedback but interpret it incorrectly; they do not receive feedback when needed; or the necessary feedback does not exist (Leveson & Thomas, 2018). For example, in a plane, if the altimeter sensor is broken, the pilot will get incorrect feedback from the display. If the altimeter is in a different mode than expected, the pilot may misinterpret the altitude. These unsafe control actions involving feedback can be captured when identifying loss scenarios.

By providing a systematic method to identify hazards and potential loss scenarios, STPA allows users to efficiently analyze the system architecture, generate effective requirements for safety and reliability, and ultimately identify gaps where changes need to be implemented. When applied in a particular way, STPA can be leveraged to effectively identify sources of mode confusion and generate recommendations to improve system design in that regard.

UCAs are identified when using STPA by applying a rigorous process described by Leveson and Thomas (2018), which requires more space to describe than is possible in this chapter. Two examples and some potential scenarios that could lead to UCAs are shown instead.

The first example involves the control action of a Boeing 777 pilot to "engage autothrottle THR REF mode." In THR REF mode, thrust is set to the reference thrust limit displayed on the engine indication and crew alerting system (EICAS; National Transportation Safety Board [NTSB], 2014). An example UCA for the pilot related to this control action is that:

UCA-1: The pilot does not engage autothrottle TRH REF mode when the pilot intended to do so.

The causal scenarios by which this unsafe control action could develop are diverse, and to conduct an exhaustive search the analyst should consider all contexts by which the THR REF mode would not engage despite a pilot's intent to engage it. One potential scenario is the following:

- **Scenario 1**
  The pilots do not engage autothrottle THR REF mode when they had the intent to engage it because they press the incorrect button and do not verify engagement of the mode by checking the FMA. The pilot may not verify the engagement of the mode due to task saturation or expectation that the mode will engage when a button is pressed because their prior experience has always been that the mode engages at the press of a button.

This scenario explains the unsafe control action in terms of a non-update to the pilot's process model of the automation. Recommendations stemming from scenarios like this will relate to ensuring *appropriate* feedback, rather than including *new* feedback, because the feedback was not perceived by the pilots despite it being available to them.

A second potential scenario is the following:

- **Scenario 2**
  Pilots do not engage autothrottle THR REF mode when they had the intent to engage it because a single press of the takeoff go around (TO/GA) switch will not engage THR REF mode if the aircraft is in a landing configuration with go-around mode armed. Engaging THR REF mode during a go-around requires a double push of the TO/GA switch (Air Accident Investigation Sector [AAIS], 2020). The pilot may not verify THR REF engagement in the FMA after the first click due to task saturation or expectation that the mode will engage when a single button is pressed. Or the pilots may not be aware that TO/GA is engaged.

This scenario also explains the unsafe control action in terms of a non-update to the pilot's process model of the automation, but also involves a misunderstanding of the implications of a control action during a particular mode. Both of these scenarios could ultimately lead to pilot mode confusion, as they all involve inadvertent activation (or non-activation) of modes coupled with an opposing belief.

The generation of UCAs for automated controllers is aided by a thorough understanding of mode transition logic and criteria, but it is also necessary that each potential mode transition be analyzed for unsafe interactions with different states of the system and environment, rather than in isolation. Specifically, in the Boeing 777 example, an available control action to the autopilot is "change vertical mode." There are various vertical modes, two of which are TO/GA mode and ALT mode. In TO/GA mode, the autopilot acquires and maintains a takeoff speed reference after liftoff, or a go-around speed reference after initial go-around rotation (NTSB, 2014). In ALT mode, the autopilot adjusts the pitch of the aircraft to stay on a target altitude (Federal Aviation Administration [FAA], 2022). An example UCA for this mode transition is:

UCA-2: A/P changes default TO/GA mode to ALT mode too early when the aircraft is still on ground without adequate feedback to flight crew.

This action involves a specific mode transition coupled with a context in which this mode transition becomes unsafe. Causal scenario identification for this UCA should consider how the mode transition logic would allow the transition to happen in this unsafe context. One potential scenario for this is:

- **Scenario 1 for UCA-2**
  The autopilot changes the pitch mode from default TO/GA mode to ALT mode because a realignment of the air data inertial reference system was initiated when the flight director was ON and the MCP selected altitude was within 20 ft of the barometric altitude (FAA, 2022). The pilot may not perceive this transition because they are expecting the default TO/GA mode. Thus, they become mode confused when they take off and upon liftoff, A/P commands nose-down pitch to obtain the set altitude for ALT mode (sea level).

This scenario explains the unsafe control action in terms of an update to the automation's process model by an action (realignment of the air data inertial reference system) but also involves a miscommunication of that process model update to the flight crew. Once again, this could lead to pilot mode confusion by the inadvertent activation of a mode that a pilot would not expect in a particular context.

Once the potential scenarios are identified, recommendations can be derived from them to design the system to eliminate or mitigate them.

Summary and Outlook

As modern systems grow increasingly complex, it becomes impossible to simply "train out" pilot and automaton behavior that can lead to hazards. Instead, the sources of such unsafe behavior should be identified and designed out of the system. STPA provides a methodology to do this by abstracting systems in terms of functional control feedback loops. The results of the STPA analysis can be used by designers to identify effective design requirements and reduce hazardous behavior.

In addition, the use of a modeling and analysis technique, such as STPA, makes it easier for hardware, software, and human factors engineers to work together to create safer design and operational procedures.

References

Air Accident Investigation Sector. (2020). *Runway impact during attempted go-around. Aircraft accident final report AIFN/0008/2016*. https://reports.aviation-safety.net/2016/20160803-0_B773_A6-EMW.pdf

Bishop, B., Harrington, P., Rose, R., & Leveson, N. (2023). System theoretic process analysis for identification of sources of mode confusion. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 67*(1), 2397–2403. https://doi.org/10.1177/21695067231192457

Bredereke, J., & Lankenau, A. (2005). Safety-relevant mode confusions – modelling and reducing them. *Reliability Engineering & System Safety, 88*(3), 229–245. https://doi.org/10.1016/j.ress.2004.07.020

Federal Aviation Administration. (2022). *Special airworthiness information bulletin: Autopilot flight director system: ALT HOLD engaged on takeoff on Boeing Model 777/787 Common Fleets*.https://ad.easa.europa.eu/blob/AIR-22-09R1.pdf/SIB_AIR-22-09R1_1

Leveson, N. G. (2012). *Engineering a safer world: Systems thinking applied to safety*. The MIT Press. https://doi.org/10.7551/mitpress/8179.001.0001

Leveson, N. G. (2019). *CAST handbook: How to learn more from incidents and accidents*. http://sunnyday.mit.edu/CAST-Handbook.pdf

Leveson, N. G., & Palmer, E. (1997). Designing automation to reduce operator errors. *Computational Cybernetics and Simulation 1997 IEEE International Conference on Systems, Man, and Cybernetics, 2*, 1144–1150. https://doi.org/10.1109/ICSMC.1997.638104

Leveson, N. G., Pinnel, L. D., Sandys, S. D., Koga, S., & Reese, J. D. (1998). Analyzing software specifications for mode confusion potential. *Safety Analysis of FMS/CTAS Interactions During Aircraft Arrivals*. https://ntrs.nasa.gov/citations/19990063822

Leveson, N. G., & Thomas, J. P. (2018). *STPA handbook*. http://psas.scripts.mit.edu/home/get_file.php?name=STPA_handbook.pdf

National Transportation Safety Board (NTSB). (2014). *Descent below visual glidepath and impact with seawall Asiana Airlines flight 214 Boeing 777-200ER, HL7742 San Francisco, California July 6, 2013* [Aircraft accident report NTSB/AAR-14/01, PB2014-105984]. https://www.ntsb.gov/investigations/accidentreports/reports/aar1401.pdf

Sarter, N., & Woods, D. (1995). How in the world did we ever get into that mode? Mode error and awareness in supervisory control. *Human Factors, 37*(1), 5–19. https://doi.org/10.1518/001872095779049516