ENGINEERING FOR HUMANS: A NEW EXTENSION TO SYSTEMS THEORETIC PROCESS ANALYSIS

Megan France & John Thomas
Massachusetts Institute of Technology
Cambridge, MA

Systems Theoretic Process Analysis (STPA) is a new hazard analysis method developed at MIT to address a broad range of accident causal factors including dysfunctional interactions among components, design flaws, and requirements problems. This paper presents a new extension for analyzing human interactions with automation and understanding why unsafe behaviors may appear appropriate in the operational context. The extension is demonstrated by applying it to pilot control of aircraft pitch control during stall recovery using scenarios from the Air France 447 accident.

On May 31, 2009, Air France flight 447 was scheduled to fly to Paris from Rio de Janeiro. Tragically, the A330's pitot tubes became clogged with ice during the transatlantic flight, causing inconsistent airspeed indications and the disconnection of the autopilot system. These events in the cockpit, combined with environmental conditions, led to intense pilot confusion. The ensuing interactions between the pilots and the aircraft sent the plane into an aerodynamic stall, which went undetected until the plane plunged into the ocean. Two hundred sixteen passengers and twelve crewmembers were killed.

Following the accident, the French Accident Investigation Bureau (BEA, 2012) released a thorough investigation into the causes of the accident, including a human factors perspective into the pilots' behavior. This analysis provided the aviation industry with valuable information; however, it cannot undo the tragedy that occurred. In order to effectively prevent future accidents, it is necessary to perform both hazard analyses and human factors investigations during design and early development. In this paper, we demonstrate a method for incorporating human factors into the hazard analysis process by expanding upon an existing technique.

Systems-Theoretic Process Analysis (STPA) is a hazard analaysis technique designed to capture not only accidents which result from component failures, but also accidents which result from design flaws and unsafe interactions (Leveson, 2012). STPA is well-suited for analyzing complex systems, but it does not provide guidance specific to humans. A new extension to the STPA method, "Engineering for Humans", was recently developed to provide guidance early in the design process and address human interactions in the system (Thomas and France, 2016). This paper demonstrates how the new extension can be applied in an aviation context to understand pilot behavior. To demonstrate the relevance of this method to the aviation domain, we apply STPA to the process of aircraft pitch control during an aerodynamic stall and show how the Engineering for Humans extension could be used to identify factors involved in the fatal Air France 447 accident.

Systems Theoretic Process Analysis

STPA begins with the identification of relevant high-level *system accidents:* any undesired or unplanned event that results in a loss. Next, high-level *system hazards* are identified: the system states or set of conditions that, together with a particular worst-case environmental conditions, will lead to an accident (loss). For example "aircraft collides with terrain" is the system accident we will be concerned with in this paper, and the system hazard which could lead to that accident is "loss of lift during flight," which may occur due to inadequate speed or excessive angle of attack.

Once the system accidents and hazards are identified, the analyst must draw the *safety control structure.* Figure 1 below shows the safety control structure for this system, which includes the control actions and feedback between entities in the system.

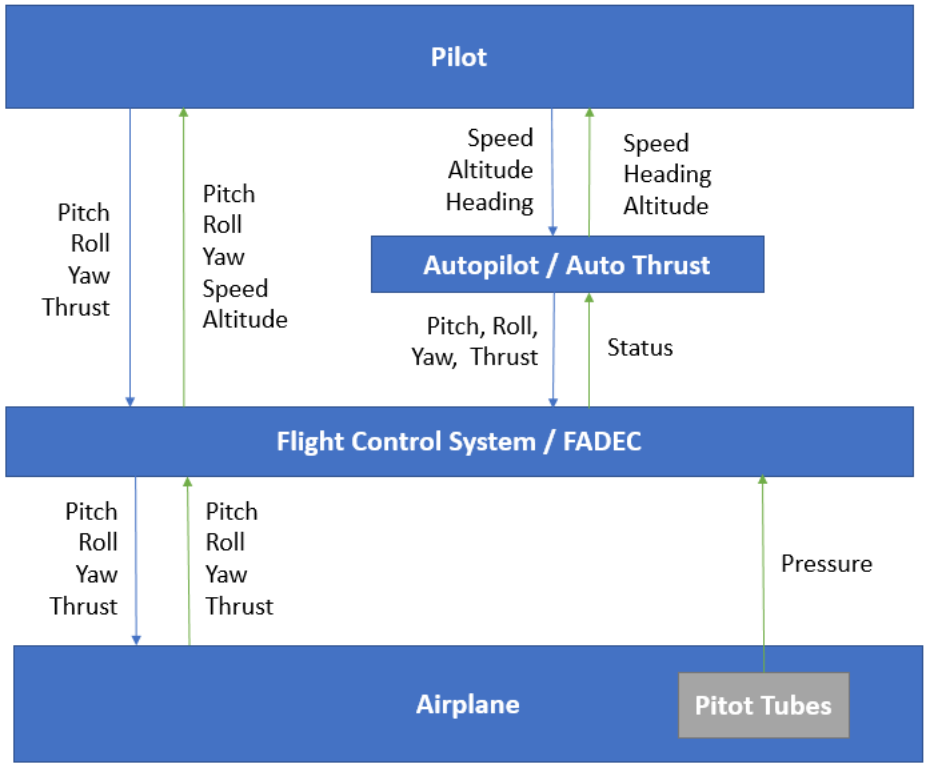

Figure 1.
Safety control structure for pilot interactions with aircraft. Note that for the purposes of this paper, a simplified and abstracted depiction of basic pilot controls can be used.

STPA has two main analysis steps. The first step examines how each control action in the system could cause a hazard. These *unsafe control actions* or UCAs, shown in Table 1, must fall under one of the four categories included as column headings.

Table 1.
Examples of Unsafe Control Actions (UCAs) for the Control Action "Increase Pitch."

Control Action	Not Providing Control Action Causes Hazard	Providing Control Action Causes Hazard	Wrong Timing or Order	Stopped Too Soon or Applied Too Long
Increase Pitch	UCA-1: PF does not increase pitch when aircraft is at risk of collision with terrain.	UCA-2: PF increases pitch while the aircraft is in a stall or approaching a stall.	-	UCA-3: PF increases pitch, but stops too soon before reaching the target pitch. UCA-4: PF continues to increase pitch too long when doing so exceeds the safe flight envelope.

Notes. "PF" refers to the pilot flying. UCA-2 is used for additional examples later in this paper.

The second step of STPA is where causal scenarios related to the UCAs are identified. The new STPA extension, Engineering for Humans, provides a process to anticipate and explain why humans might provide these unsafe control actions. This process is summarized in the next section.

Method

The new extension uses the human controller model depicted in Figure 2 below.

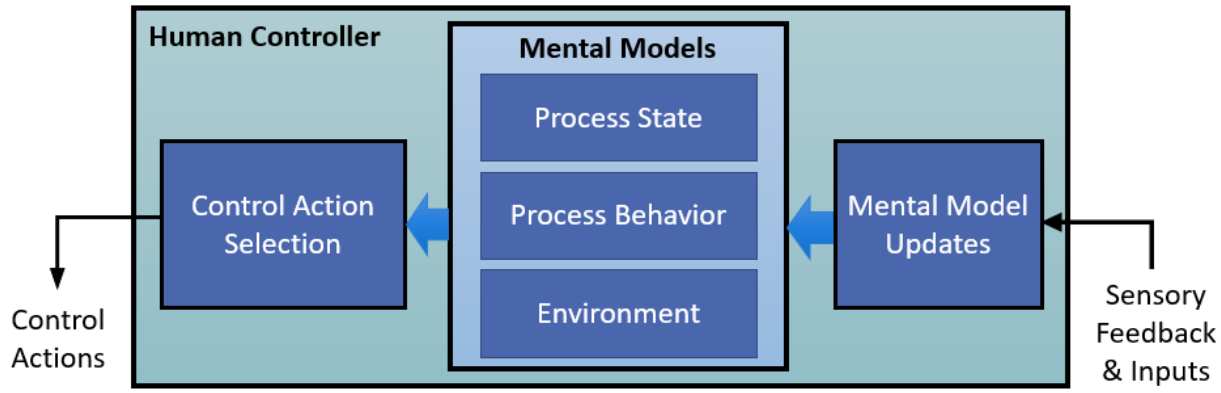

Figure 2.
Extended model of the human controller.

In this model, the *Control Action Selection* stage explains why a particular control action may be chosen by considering factors such as the operator's goals and other tasks that may compete for priority. Whether an action is skill, rule, or knowledge-based (Rasmussen, 1982) is also an important aspect of this stage of the model.

The *Mental Models* stage captures various human beliefs about the outside world. First, the *Mental Model of Process State* reflects the operator's awareness of software modes and the current state of operation. Incorrect mental models of process state may result from automatic mode changes, or progression of a controlled process to the next stage without feedback to the operator. The *Mental Model of Process Behavior* describes the operator's expectations for how the system will behave in a particular mode or stage of operations, and includes cause and effect relationships between the operator's actions and the system's behavior. Lastly, the *Mental Model of the Environment* includes factors outside the operator's control, including the behavior of other controllers and the novelty or variability of the environment.

Finally, this model requires the analyst to consider the source of mental models, including both how they are formed and how they are updated in response to change. Factors such as the salience of the change and the operator's expectations influence how likely that change is to be sensed (Wickens, Helleberg, Goh, Xu, & Horrey, 2001). In this stage we may also consider how factors such as time pressures and limitations of human attention may lead to the formation of incomplete or incorrect mental models.

Results

The output of the new extension is a set of scenarios for each UCA that explain why that action may have appeared logical to the human operator in context. These scenarios can be written in a paragraph or outline format and summarize the systemic factors that could contribute to the operator's behavior. The new extension also provides a way to illustrate the scenarios in a graphical format, at shown in Figure 4.

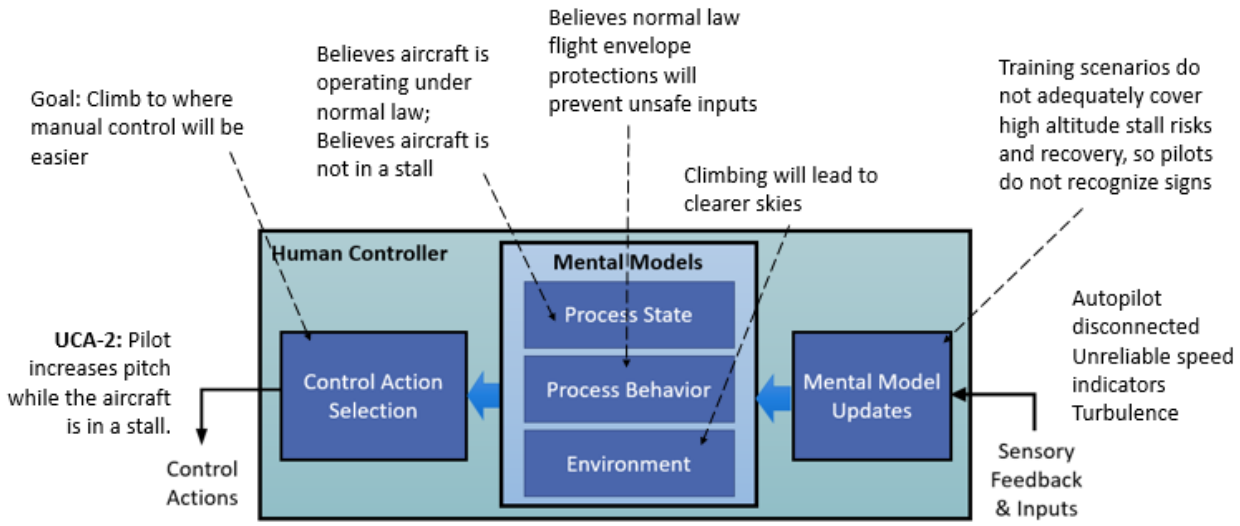

Figure 3.
Graphical depiction of a scenario based on events of the Air France 447 crash.

In Figure 3 above, a scenario is depicted graphically to show its relationship to each aspect of the new human controller model. The Unsafe Control Action in this figure is UCA-2: PF increases pitch while the aircraft is in a stall. Why in the world would a pilot do that? The first part of the model, the Control Action Selection, explores factors like pilot goals—as explained in the previous section. For example, the pilot may believe manual control will be

easier at a higher altitude, where he knows the skies are clearer. The next stage, the Mental Models, explores pilot beliefs. For exmaple, the pilot may believe it is safe because he thinks the aircraft is not in a stall, is operating under normal law, and flight envelope protections will prevent him from stalling the plane. The last stage, the Mental Model Updates, explores how the pilot interprets information and updates (or doesn't update) their mental model. For example, due to inadequate training in high altitude stalls, he does not expect one to occur, and thus does not recognize the turbulence he is experiencing as part of a stall.

In the case of Air France flight 447, the pilots were faced with the sudden disconnection of the autopilot system, as shown in the scenario above. Lacking accurate speed information, they did not realize that they were at risk for a stall, proceeding to climb and even decrease speed. They realized too late that normal law, which provides flight envelope protections, was no longer in effect and the aircraft was operating under alternate law, which permitted the unsafe pitch inputs. The pilots may have been able to recover from an incident involving any of these factors alone, but it was the combination of the cockpit and stimuli, their beliefs about the system, and the circumstances influencing their decisions that ultimately led to the accident.

The scenario shown in Figure 3 is just one of many potential ways that an accident could occur. Other scenarios related to pitch input could lead to different accidents, and the method demonstrated here provides a systematic way of identifying such scenarios so that proactive efforts can be made to eliminate the factors that contribute to accidents. For example, the scenario in Figure 3 may suggest a need to more conspicuously indicate shifts from normal law to alternate law, or to improve training in high altitude stall procedures. Using STPA as a design tool, rather than a means of understanding causes of an accident after the fact, allows engineers and company management to make proactive decisions to improve safety.

The advantage of this extended STPA method is that it prompts the analyst to consider not only the different components of the operator's mental model, but also how that model is formed and what impact it has on decisions. While other models may provide a more nuanced view into human cognition (eg. Rasmussen, 1982; Endsley, 1995; Wickens, 1992), this model is deliberately simplified so that it can be used by industry practitioners without an extensive background in psychology or human factors. Those who do have such a background can also use this method to elicit their knowledge and experience in detailed topics such as sensation, perception, learning, and decision making while ensuring their explanations are structured in a way that is accessible to engineers and practitioners of all backgrounds. This model can thus be used as a common framework to talk about human automation interactions during system design and early development efforts.

Conclusions

STPA is a valuable hazard analysis technique with applicability across domains, and the new Engineering for Humans extension proposed in this paper provides additional guidance for analyzing human-automation interactions within the context of the larger system. When used to examine aircraft pitch control during an aerodynamic stall, the new extension can be used to model scenarios from the tragic Air France 447 crash. This new model provides a framework for

understanding and anticipating human behavior during the design process, which is necessary to prevent such tragic accidents in the future.

Acknowledgements

This work is based on original research sponsored by General Motors (Thomas and France, 2016). The authors thank General Motors for their ongoing support in this area.

References

BEA (2012). *Final report on the accident on 1st June 2009 to the Airbus A330-203 registered F-GZCP operated by Air France, flight AF 447 Rio de Janeiro – Paris*. Paris, France: Bureau d'Enquêtes et d'Analyses pour la sécurité de l'aviation civile.

Endsley (1995). Toward a theory of situation awareness in dynamic systems. *Human Factors, 37(1),* 32-64.

Leveson, N. G. (2012). *Engineering a safer world: Systems thinking applied to safety.* Cambridge, MA: The MIT Press.

Rasmussen, J. (1982). Human errors. A taxonomy for describing human malfunction in industrial installations. *Journal of Occupational Accidents, 4,* 311-333.

Thomas, J., France, M. (2016). *Engineering for Humans: STPA Analysis of an Automated Parking System*. Fifth MIT STAMP Conference, Cambridge, MA.

Wickens, C. D., Helleberg, J., Goh, J., Xu, X., & Horrey, W. J. (2001). Pilot Task Management: Testing an Attentional Expected Value Model of Visual Scanning (Technical Report No. ARL-01-14/NASA-01-7). NASA Ames Research Center. Retrieved from http://www.aviation.illinois.edu/avimain/papers/research/pub_pdfs/techreports/01-14.pdf.

Wickens, C.D. (1992). Engineering Psychology and Human Performance, New York: Harper Collins.