ANALYSES OF THE BOEING 737MAX ACCIDENTS:
FORMAL MODELS AND PSYCHOLOGICAL PERSPECTIVES

Immanuel Barshi
Human Systems Integration Division
NASA Ames Research Center, California
Asaf Degani
GM Advance Technology Center
Hertzeliya, Israel
Robert Mauro
Decision Research and University of Oregon
Eugene, Oregon
Randall J. Mumaw
San Jose State University Research Foundation
NASA Ames Research Center, California

Two fatal accidents involving the B737MAX resulted from the flight crews' inability to overcome the effects of the Maneuvering Characteristics Augmentation System (MCAS). MCAS was designed to mimic the control column feel pressure and pitching behavior of the B737NG, which was the certification basis for the B737MAX. We briefly describe the potential role of formally modeling different perspectives during system design, and how such modeling can reveal gaps and conflicts between perspectives. We also discuss some of the relevant human factors issues involved in these accidents and how the aircraft's behavior may have affected the pilots' psychological states. Implications for automation design are considered.

The Boeing 737 has been the most successful airliner model in the history of aviation. At any given moment, there are more B737s flying in the world than any other aircraft. In spite of its enviable safety record, the two fatal accidents of the B737MAX-8, one in Indonesia in October of 2018 (KNKT, 2019) and the second in Ethiopia in March of 2019 (EAAIB, 2022) shook the world and led to the unprecedented world-wide grounding of the MAX fleet.

In both accidents, the pilots were unable to understand what was happening to their aircraft. Although MCAS, the Maneuvering Characteristics Augmentation System, has been a major focus of numerous discussions of these accidents, the confusion that rendered the pilots unable to successfully diagnose and remedy the problems began before MCAS was activated. To understand why these accidents happened, one must consider the situations the pilots encountered from the pilots' perspectives.

The B737 was originally designed, in the 1960s, as a "federated" system. Separate and redundant aircraft avionics, on the right and left sides of the aircraft, supply data to a corresponding set of flight displays; left side for the Captain and right side for the First Officer. Although some comparators were added as the aircraft evolved, the fundamental federated design concept remained. Similarly, each side has separate flight control computers. Should a problem occur on one side, control can be transferred to the other side's computers and safely continue the flight. However, in this design, the burden is on the flightcrew to communicate about what is happening on each side, to determine which set of equipment is functioning properly. Accidents can occur when the flightcrew fails to understand which side has failed.

Formal Modeling

Every human-machine system can be viewed from different perspectives. These different perspectives can be characterized as "models," including the human's mental model of how an aircraft and its systems work (Degani et al., 2022). The design begins with a "conceptual model" that exists in the mind of the designer(s). This model—not necessarily fully detailed, accurate, or complete—portrays the thinking behind the system and is the vital first step. Next is the "machine model" which concretizes how the design team understands the conceptual model. The machine model is not necessarily complete, but the "system dynamics model" incorporates how the system works in its operational environment and is verified using system engineering tools and flight simulators. Other models such as requirement/specification models and software implementation models may also be produced.

An "interface model" represents the information the user is expected to need to operate the system. Thus, it necessarily abstracts the detailed behavior behind indications seen on displays; it is augmented with aircraft manual and training informaiton. The next model is called the "user model." This model, an abstracted version of the interface model, characterizes an individual user's "mental" model of the system and its workings. The user model is based on the information obtained and the user's understanding of it. It can decay with time and lack of recurrent experience. User models are also subject to degradation and loss due to fatigue and stress. Common examples of such degradation are cognitive tunneling and inattentional blindness (Levin & Baker, 2015), in which the user's attention is focused on one thing and ignores other potentially relevant data.

Formally modeling these different perspectives allows for the identification of gaps and potential conflicts between models. Such gaps and conflicts could lead users to confusion and to mistakes. Early research on pilot interactions with cockpit automation showed that the inability to understand what the automation was doing constituted the most critical concern (Billings, 1997; Parasuraman, et al., 2000). Cockpit observations by Earl Wiener when automated flight control systems were first introduced into commercial aviation showed that pilots wanted answers to four key questions: "what's it doing now, why is it doing it, how did I get here, and what will it do next" (Wiener & Curry, 1989). Albeit somewhat colloquial in nature, these types of questions are still being asked in modern-day cockpits.

For the sake of brevity, we focus here on one model and one accident (for a detailed analysis, see Barshi et al., 2023). We focus on the way in which a description of the machine model can expose a conflict with the user model. Exposing these conflicts while the aircraft is being designed could lead to solutions that could mitigate the risks associated with such conflicts. For a discussion of the pilots' experience, we focus on the first accident, Lion Air flight 610, because that crew did not know about MCAS (KNKT, 2019). The crew of the second accident, Ethiopian Airlines flight 302 supposedly knew about MCAS and was refreshed in its training of the proper procedure to disable it (EAAIB, 2022).

Figure 1 below presents a simplified version of a small portion of the machine model of the electric pitch trim system of the B737MAX (for a detailed analysis, see Barshi, et al. 2023). This system controls the movement of the horizontal stabilizer to trim the pitch attitude of the aircraft and includes a manual and an electric activation. The manual activation is performed using a hand-operated wheel in the cockpit that is physically connected with cables to the stabilizer and allows the flight crew to directly control the movements of the stabilizer. Electric activation is performed using an electric motor that can receive commands from the flight crew by use of thumb switches mounted on each yoke. The electric motor can also receive commands from the flight control computer, which houses three components that can activate the trim: the autopilot, the speed trim system, and MCAS (NTSB, 2019). The stabilizer can be moved to trim the aircraft nose up (ANU) or aircraft nose down (AND). The trim is used to relieve pressures from the control column for any given pitch attitude, power setting, and speed.

The stabilizer can be moved by automated systems (the autopilot and the speed trim system) that can fail and cause a runaway trim situation where the aircraft is forced into a dangerous pitch attitude (either too high leading to a stall, or too low leading to a dive). To stop a runaway trim, a mechanism is installed under the cockpit floor (known as the *floor switch*), at the base of the control column, that disengages the electric trim motor in case the column is moved in a direction opposite to the movement of the trim. For instance, if the autopilot fails and causes an excessive nose-up trim, pushing the control column forward stops the motor from moving the stabilizer. However, because MCAS is designed to produce forward pressure on the control column when the pilot is pulling the control column back, the floor switch is disabled when MCAS is active (MCAS_input = true in Fig. 1), leaving MCAS free to continue moving the stabilizer in an AND direction (NTSB, 2019).

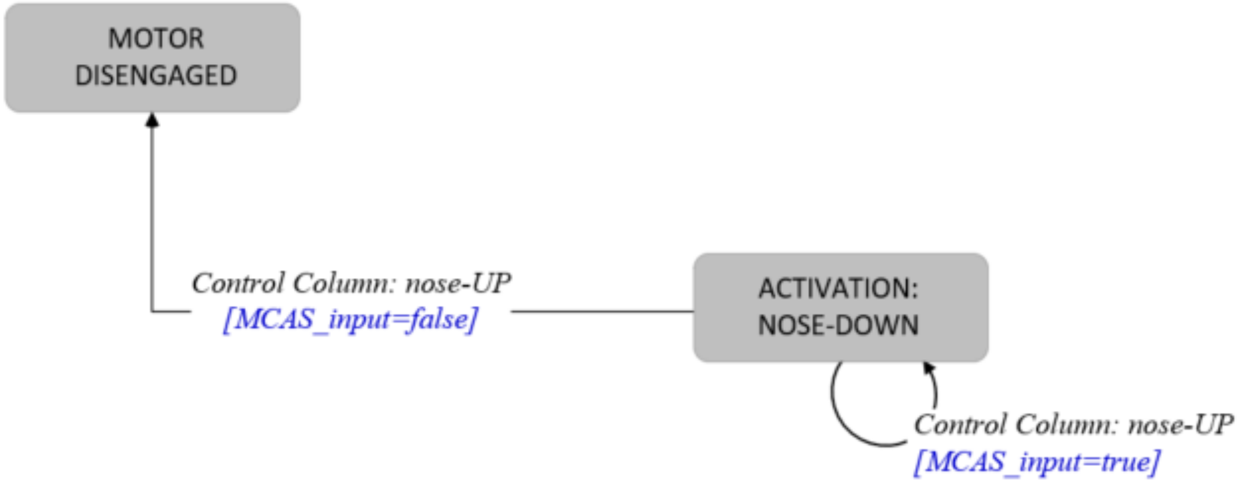

Figure 1. Modeling the behavior of the column-activated floor switch.

Presenting the machine model as seen in Figure 1 shows that the design creates a situation that to the pilot would appear *non-deterministic*; the pilot cannot predict the behavior of the system, even if it is completely predictable to the designer. If the autopilot is trimming AND, pulling back on the yoke stops the movement. If MCAS is trimming AND, pulling back on the yoke does nothing. Since the pilot may not know which system is causing the AND runaway trim, the pilot cannot predict whether pulling back on the controls is going to help or not. From the pilot's perspective the behavior of the system is unpredictable, and the pilot cannot answer the question of "what is it doing now?" nor the question of "what will it do next?" Thus, presenting the machine model, as in Figure 1, reveals this apparent non-determinism and provides the designers with an opportunity to develop a mitigation to resolve it.

This apparent non-determinism falls under the category of "mode error" (Woods et al., 1997; Sarter & Woods, 1995); the user is unable to determine what mode the system is in, and to predict the implications for ongoing control of the aircraft. Mode errors resulted in a long series of automation-related accidents starting in the 1980s (Mumaw, 2021). The method described here of detailing the different models involved in the human-machine system can expose specific gaps between the machine model and the user model and thus lead to specific design solutions.

Psychological Perspective

In the case of Lion Air flight 610 (KNKT, 2019), the problems on the flightdeck started as the aircraft began to rotate for its takeoff. At this point, the Captain's "stick shaker" activated. The stick shaker is designed to warn pilots that the aircraft is about to stall and cease flying. At any altitude, it demands immediate attention. When the aircraft is only a few feet above the ground, it warns of an impending disaster. The immediate emotional response is fear. The concomitant psychological and physiological changes would reduce the pilots' functional working memory, making it difficult to reason through complex problems (Davies & Parasuraman, 1982; Moran, 2016).

At this point in the takeoff, there is likely insufficient runway left to put the aircraft safely back on the ground. An attempt to abort the takeoff would likely cause the aircraft to overrun the runway, seriously damaging the aircraft and probably causing injuries to the passengers. Pilots are trained to continue the takeoff at this point. But if the aircraft stalls when near the ground, there is insufficient altitude to recover, and the aircraft will crash. The immediate actions that pilots must take when the stall warning is activated are well-rehearsed. Lower the nose and add power to break the stall. In the simulator, during training, a successful stall recovery is defined in part by a minimum loss of altitude. At takeoff, the aircraft is near full thrust, so there is little power to add. The aircraft is also so low that there is little altitude to lose. Thoughts of the Northwest 255 (NTSB, 1988) and Spanair 5022 (CIAIAC, 2011) accidents may jump to mind. In these accidents, the aircraft was improperly configured for takeoff and crashed immediately thereafter. One might expect that the pilots of Lion Air 610 immediately checked that their aircraft was properly configured. A glance at the cockpit indications would confirm that it was and furthermore the aircraft was flying and gaining altitude. Meanwhile, the stick shaker continued to shake the Captain's yoke and arms while making a loud racket. The only option at this point is to try to gain additional altitude and diagnose the problem or at least determine that the aircraft would be able to return to the airport and land safely.

There is no explicit indication from the Cockpit Voice Recorder (CVR) that either pilot noticed that only the Captain's stick shaker was activated, but it is hard to ignore. There could be a malfunction with the equipment on one side of the aircraft, but which side and what is the nature of the malfunction? Other alerts appeared seconds later: altitude disagree, airspeed disagree, and feel pressure differential. All indicate that some of the information calculated by the right-hand computers disagreed with the information calculated by the left-hand computers, but again, the burden in this federated architecture is on the flightcrew to determine which side is correct.

The immediate inference that one could make from these alerts is that something is seriously amiss with the aircraft systems. Airspeed is directly relevant to the potential stall problem, so determining which airspeed display (left or right) is correct would be a high priority after maintaining control of the aircraft. So the Captain called for the First Officer (FO) to carry out the memory items for the "Unreliable Airspeed" non-normal checklist. The FO failed to respond. A short while later, the Captain called for the checklist itself. The First Officer had trouble locating the checklist. These problems are likely symptoms of substantial stress and anxiety (Maloney et al., 2014; Moran, 2016). It is very unusual for the stick shaker to operate continuously. In normal operations, it rarely activates; when it does, it is only active momentarily, ceasing when the triggering condition is corrected. In addition to the noise, the constant reminder that the aircraft could cease flying at any moment could take a toll on the crew.

After the Unreliable Airspeed checklist was located it could have been used to effectively troubleshoot, locate the reliable airspeed indicator, and determine that the aircraft was not in danger of stalling. But at this point, the Captain asked the FO to request an Air Traffic Control (ATC) clearance to a holding point; he was likely looking for a safe space to troubleshoot the problem. This may indicate that the Captain had concluded that the aircraft was not in imminent danger of stalling and that he wanted to confirm that the aircraft could be safely operated before attempting to land, but this can't be confirmed from the CVR transcript available in the accident investigation report.

The FO complied and also suggested raising the flaps from 5 to 1. This action would be in line with normal procedures after takeoff but would have been a possible problem due to the loss of lift if the aircraft were on the verge of a stall. The Captain's agreement might further indicate that he thought the aircraft was not about to stall, despite the stick shaker.

The Captain then requested that the FO take over the controls, perhaps to allow him to be free to troubleshoot. The FO replied for him to standby and suggested raising the flaps the rest of the way.

Avoiding taking control of the aircraft and sticking to the normal procedures in a non-normal situation might be additional signs of narrowed attention and reduced cognitive functioning (Maloney et al., 2014; Moran, 2016). The Captain agreed to retract the flaps. Unbeknownst to the crew, this action armed MCAS. The MCAS program, like other programs running on the left-side aircraft computers were receiving and relying upon erroneous angle of attack information from the angle of attack sensor on the left side of the aircraft. The stick shaker and the various alerts were all symptoms of this malfunction.

Shortly thereafter, MCAS began to exert downward pressure on the controls through inputs to the stabilizer pitch angle. As Figure 1 shows, just pulling back on the controls was not going to stop MCAS. The Captain ordered a return to flap 1 and retrimmed, countering the effects of the previous MCAS command. Had the flaps remained deployed, MCAS would not have reactivated, and the flight could have landed safely.

But the flaps were raised again. The CVR transcript provided in the accident report (KNKT, 2019) does not include any discussion or commands to raise the flaps. There is no evidence that the airspeed unreliable checklist was ever completed. Yet, the Captain maintained appropriate pitch with trim, using the thumb switch, stopping MCAS and compensating for MCAS initiated nose down trim. Yet, he never verbalized what he was doing, and the FO may have had no awareness of these actions and the Captain's struggles. Perhaps, he was not completely conscious of it. With his hands shaking throughout the flight from the stick shaker, he might not have been fully aware of the forward pressure on the control column. In any case, the Captain managed to return the horizontal stabilizer to its climbing trimmed angle following each MCAS activation, and thus kept the aircraft flying safely.

While maneuvering for a return to a landing, and after 21 successive MCAS activations, but without making any reference to the extensive use of the trim, the Captain asked the FO again to take control of the aircraft. He might have wanted to take a break to prepare for the landing or do the trouble-shooting that the FO had been unable to conduct. He might have been saturated. When control was transferred, the aircraft was properly trimmed and flying. But the flaps were up and MCAS activated. The floor switch was disabled, the FO failed to compensate sufficiently with the use of the yoke-mounted thumb switch and eventually the MCAS' AND inputs overwhelmed the FO's attempts to manage pitch, leading to an unrecovered dive into the water.

Conclusion

Evaluating formal models during the design of a system can help identify gaps between models, such as the gap of apparent non-determinism between the machine model and the user model described above. Such a gap could be made visible to the crew, for instance, through a salient mode annunciation, alerting, or through education. Furthermore, understanding the operational context of use and some of the psychological aspects of the user can help elaborate the user model and possibly expose additional gaps, particularly between the user model and the interface model. Although the analysis presented here was done post-hoc, after the aircraft was already produced and after the accidents had occurred, the methodology can be applied during the design, testing and verification of systems and thus help prevent such accidents from happening again.

References

Barshi, I., Degani, A., Mauro, R., & Mumaw, R.J. (2023). *Analyses of the Boeing 737MAX accidents: Formal models and psychological perspectives*. Manuscript in preparation.

Billings, C. E. (1997). *Aviation Automation: The Search for a Human-Centered Approach. Mahwah*, NJ: Erlbaum.

CIAAIC (2011). Accident involving a McDonnell Douglas DC-9-82 (MD-82) aircraft registration EC-HFP, operaed by Spanair, at Madrid-Barajas Airport, on August 2008. Comision de Investigacion de Accidentes e Incidentes de Aviacion Civil, Report A-032/2008. Madrid, Spain.

Davies, D. R., & Parasuraman, R. (1982). *The Psychology of Vigilance*. London, Academic Press.

Degani, A., Shmueli, Y. & Bnaya, Z. (2022). Equilibrium of Control in Automated Vehicles: Driver Engagement Level and Automation Capability Levels. *Proceedings of the 4th IFAC Workshop on Cyber-Physical & Human-Systems*. Houston, TX: IFAC.

EAAIB (2022). Investigation Report on Accident to the B737- MAX 8, ET-AVJ, 10 March 2019. Ethiopian Aircraft Accident Investigation Bureau, Report No. AI-01/19. Addis Ababa, Ethiopia.

KNKT (2019). Aircraft Accident Investigation Report. PT. Lion Airlines Boeing 737 (MAX); PK-LQP Tanjung Karawang, West Java, Republic of Indonesia 29 October 2018. Komite Nasional Keselamatan Transportasi (National Transportation Safety Committee), Report No. 18.10.35.04. Jakarta, Indonesia.

Levin, D. & Baker, L. (2015). Change blindness and inattentional blindness. In. J. Fawcett et. Al. The Handbook of Attention, MIT Press.

Maloney, E. A., Sattizahn, J. R., and Beilock, S. L. (2014). Anxiety and cognition. *Wiley Interdisciplinary Reviews Cognitive Science, 5*, 403–411. doi: 10.1002/wcs.1299

Moran, T. P. (2016). Anxiety and working memory capacity: a meta-analysis and narrative review. *Psychological Bulletin, 142*, 831–864. doi: 10.1037/bul0000051

Mumaw, R.J. (2021). Plan B for eliminating mode confusion: An interpreter display. *International Journal of Human-Computer Interaction, 37* (6).

NTSB (1988). Aircraft Accident Report: Northwest Airlines, Inc. McDonnell Douglas DC-9-82, N312RC, Detroit Metropolitan Wayne County Airport, Romulus Michigan, August 16, 1987. National Transportation Safety Board. Washington, DC.

NTSB (2019). System Safety and Certification Specialist's Report. National Transportation Safety Board, NTSB ID No.: DCA19RA017. Washington, DC.

Parasuraman, R., Sheridan, T.B., & Wickens, C.D. (2000). A model for types and levels of human interaction with automation. *IEEE Transactions on Systems, Man, and Cybernetics, 30* (3), 286-297.

Sarter, N.B. & Woods, D.D. (1995). "How in the World Did We Ever Get into That Mode?" Mode Error and Awareness in Supervisory Control. *Human Factors, 37*, pp. 5-19.

Wiener, E. L., & Curry, R. E. (1989). Cockpit Automation and Crew Coordination. NASA Contractor Report 196099. NASA Ames Research Center, Moffett Field, CA

Woods, D., Sarter, N., & Billings, C. (1997). Automation surprises. In G. Salvendy (Ed.), *Handbook of Human Factors and Ergonomics*, pp. 1926-1943. New York: John Wiley.