5-1-2021

The Skill Assumption -Over-Reliance on Perception Skills in Hazard Assessment

M M Rene van Paassen

Jelmer R. Reitsma

Erik_Jan A M Hujibrechts

Clark Borst

Annemarie Landman

See next page for additional authors

Authors
M M Rene van Paassen, Jelmer R. Reitsma, Erik_Jan A M Hujibrechts, Clark Borst, Annemarie Landman, and Max Mulder

THE SKILL ASSUMPTION - OVER-RELIANCE ON PERCEPTION SKILLS IN HAZARD ASSESSMENT

M. M. (René) van Paassen†, Jelmer R. Reitsma†, Erik-Jan A. M. Huijbrechts‡,
Clark Borst†, Annemarie Landman* and Max Mulder†
† Aerospace Engineering – Delft University of Technology
Delft, The Netherlands
‡ Airline pilot, KLM, Amsterdam, The Netherlands
* TNO Human Factors, Soesterberg, The Netherlands
and Aerospace Engineering – Delft University of Technology
Delft, The Netherlands

In the analysis of human performance and human error, considerable attention is given to the cognitive processes of actors involved in error or success scenarios. Even with awareness of hindsight bias, it takes effort to understand the actions of agents in later inspection of error scenarios. One such topic of heated discussion was the perceived poor performance of pilots in the two 737 MAX MCAS-related crashes in applying the "memory item" checklist pertaining to a runaway trim. In this paper, we argue that it is not so much the reproduction of the checklist that was lacking in these scenarios, but the trigger for even starting the checklist. Not only trim run-away problems, but several other issues likewise require an instant reaction from pilots, designated as "memory items". Rasmussen's simplified schematic for the "skill, rule and knowledge" taxonomy already provides the tools for properly analyzing this. The skill to provide the triggers for these reactions relies on pattern extraction from the available sensory input, and, importantly, it can only be learned in a valid training context. It is argued that re-appraisal of these items is needed, addressing explicitly the validity of the training environments that enable pilots to learn the required pattern recognition skills.

Introduction

With improving technical reliability of aircraft, and improved training and team concepts for flight crew, aviation safety has slowly but steadily improved (Pasztor & Martin, 2020). Each remaining accident or incident is carefully analyzed to learn lessons for further safety improvements, and to watch for trends in aviation safety. The twin accidents of the new Boeing 737 MAX aircraft certainly received their share of the aviation community's attention. In these accidents, a faulty sensor feeding its data into the Maneuvering Characteristics Augmentation System (MCAS), led to the MCAS system applying repeated nose-down pitch trim inputs, producing high forces on the control column, so that the pilots were not able to keep the aircraft from entering a dive. Analyses led to conclusions on failed oversight, a degradation in safety culture at Boeing, questioning of the practice of co-pilots with relatively little flight experience, and the safety process of low-budget airlines.

Research on improving flight safety investigates issues such as the role of startle and surprise (Landman et al., 2017), and training with variability (Landman et al., 2018). This work focuses on preparing pilots for failure diagnosis and recovery, also in the presence of distracting startle conditions, and when operating with a possibly wrong situational frame with respect to concepts on the aircraft and environment state. Other work focuses on safety culture, (Dekker, 2006; Roelen & Klompstra, 2012), investigating an organization's ability to instigate safe practices. Research on specific topics, such as the language and communication (Tajima, 2004), is work that can lead to recommendations by certifying authorities.

It is important to understand that safety in flight operations depends in a large part on proper preparation. Time to diagnose problems is often short, and the time pressure in flight impacts diagnostic capabilities. In some conditions, with flight near aircraft limits such as stall or overspeed, or near terrain, the time available to the first action is counted in seconds. Most incidents are also not new and unique, like the erroneous MCAS activation was, but can be reasonably prepared for. Hazard analysis and operational experience have resulted in a vast body of possible failure cases, and pilots train for known cases such as engine failures, inadvertent stalls, wind shear conditions, cabin leaks leading to depressurization, and a whole slew of technical malfunctions. Reference materials and checklists are prepared for these cases, providing additional pre-programmed responses to foreseen trouble.

In the aftermath of the 737 MAX crashes, several authors claimed that, even though the trim activation by the MCAS system was inappropriate, the flight crew could have responded as taught, by performing the "memory items" from the "Runaway Stabilizer" checklist. Of this checklist, the first five items are memory items, these are actions that a pilot should study, learn by heart, and be able to perform without referring to the physical paper checklist. They are also common with the items in checklists for previous 737 models, making them familiar to pilots transitioning from previous models.

Boeing's training materials for 737 MAX pilots do not even mention the MCAS system as a separate topic. From later investigation, it appears that Boeing's official consensus was that any malfunction of the MCAS system would be handled by the pilots as if it were a "normal" trim system malfunction. It has now been recognized that Boeing overly relied on the capability of pilots to provide the necessary responses to malfunctions (Anon., 2019).

Angle of attack sensor failure on the original 737 MAX

Both accident scenarios with the 737 MAX were triggered by the same failure of an angle of attack sensor on the captain's side. Signals from the angle of attack sensor affect a wide range of the aircraft's systems:

- The pilot's and co-pilot's speed and altitude indications started to differ in both scenarios. This is due to the use of the angle of attack value from the sensor for the correction of placement errors of the static and total pressure sensors, correcting the output values of these sensors.[1]
- The measured angle of attack forms the basis for the activation of the captain's stick shaker and low airspeed warnings.
- The different values for angle of attack affect the feel systems for the captain's and co-pilots control column differently, leading to a "feel differential" message.
- The angle of attack sensor forms the input for the MCAS system, leading to repeated and aggressive pitch down trim actions, but only for certain aircraft configuration, that is, with the autopilot disconnected and flaps up.

This all produces a situation that does not resemble a normal "old school" runaway stabilizer trim incident. We will subsequently look at the incidents, and use Rasmussen's Skill, Rule and Knowledge taxonomy to understand the pilots' failure to start the trim checklist.

The Skill-Rule-Knowledge taxonomy inspected

Last week, I (the first author) opened the family's dishwasher with the intention to fill it with the dirty items that had collected in the sink. As I noticed humid warm air coming off the machine, I changed my activity and cleaned out the dishwasher, since it had finished the cleaning cycle initiated by one of the other members of the household.

Rasmussen (Rasmussen, 1983; Rasmussen, 1986), presents a three-level taxonomy for human behavior, allowing a rough categorization into skill-based, rule-based and knowledge-base behavior, see Fig. 1. This distinction is useful, because each level of human behavior has its own strengths and weaknesses, and understanding which behavior is needed in which work situation enables an analyst to assess training needs and possibilities for error. A typical example of skill-based behavior would be manual flying of an aircraft, rule-based behavior would be performing the checklist, and knowledge-based behavior would be assessing whether a flight can be performed safely, given pilot skills, aircraft equipment and weather forecast. However, these examples do not do justice to the richness of the concepts in the SRK taxonomy. Most rule-based behavior does not correspond to activities where explicit rules have been formulated, and may equally be termed "habit-based", or "routine-based". Thus activities like cleaning out the dishwasher, should be classified as rule-based behavior. In this case, the activity of cleaning out the dishwasher was triggered by a familiar perception of humid warm air.

[1] Judging from the very large altitude and speed differences presented to the pilot and co-pilot, it might be assumed that this algorithm does not cap the correction value, allowing unrealistic measured values to drastically influence instrument presentation.

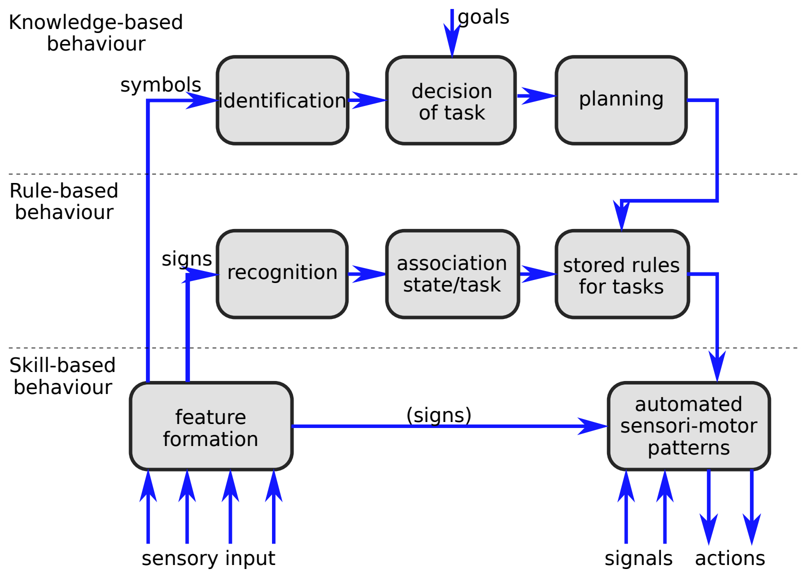

Figure 1: Simplified illustration of the skill-rule-knowledge taxonomy (Rasmussen, 1983)

Application to the MCAS incidents

To the reader or observer of transcripts, timelines, graphs and replays of aircraft accidents, it is often difficult to understand why the pilots did not initiate actions that would have led to a successful outcome. Of course, inspection of the incident after the fact is prone to hindsight bias; the knowledge of the outcome places all facts in a biased review, inviting the classification into wise and unwise decisions. It seems reasonable to assume that the pilots in both scenarios knew how to perform the memory items from the Runaway Stabilizer checklist, which makes it all the more curious why these actions were not performed swiftly or at all.

In Rasmussen's simplified model of SRK behavior, a significant clue is hidden in the connections between the blocks. Rasmussen labels three kinds of information: signal, sign and symbol. Signals are for perceptual communication. Signs are for triggering rule-based activities, and symbols are recognized morsels of information about the outside world that are used in cognition and knowledge-based behavior. In Fig. 1 Rasmussen's signs and symbol lines both originate in the "feature formation" block, explicitly depicted at the skill level. In the checklist for the 737 MAX (Fig. 2), the start of the procedure is simply listed as a condition: "Uncommanded stabilizer trim movement occurs continuously". This condition seems easy enough to detect; stabilizer trim movement is visible in the trim wheels and audible on the flight deck. Since memory items on the checklist must be performed quickly and skillfully, performing these items *should* be rule-based behavior. Rasmussen's simple model indicates that this rule-based behavior can be started by a "sign", which must be a recognized condition that triggers this pre-learned response. The recognition and production of the sign takes place *at the skill level* in Rasmussen's taxonomy. If we take the requirement that the Runaway Stabilizer checklist is executed swiftly and without delay, the pilots must therefore both have the skill to almost instinctively recognize the trigger, and the practice and routine to execute the required actions. The alternative option is to start these rule-based activities after knowledge-based evaluation, which is undesirable, since this is a possibly lengthy and effort-full process.

For the checklist trigger to occur, two other conditions are needed; the trim must be uncommanded and continuous. The Pilot Flying is supposed to recognize the condition, switch off autopilot and autothrottle, and, if needed, command the Pilot Not Flying to move the Trim Cutout Switches to Cutout by stating: "Runaway Stabilizer, memory items". To complicate matters, uncommanded trim input in the 737 (MAX as well as older models), is quite common. The Speed Trim System (STS), that is active in manual flight. can produce trim bursts of several second. As the third author has experienced, these trim actions can be counter-intuitive when the pilot wants to accelerate or decelerate.

Lion air 9.1
737 Flight Crew Operations Manual

Runaway Stabilizer

Condition: Uncommanded stabilizer trim movement occurs continuously.

1 Control column. . . . . Hold firmly

2 Autopilot (if engaged) . . . . .Disengage

Do **not** re-engage the autopilot.

Control airplane pitch attitude manually with control column and main electric trim as needed.

3 Autothrottle (if engaged). . . . . .Disengage

Do **not** re-engage the autothrottle.

4 **If** the runaway **stops** after the autopilot is disengaged:

■ ■ ■ ■

5 **If** the runaway **continues** after the autopilot is disengaged:

STAB TRIM CUTOUT switches (both) . . . . CUTOUT

If the runaway **continues**:

Stabilizer trim wheel . . . . Grasp and hold

6 Stabilizer. . . . . Trim manually

7 Anticipate trim requirements.

▼ Continued on next page ▼

June 15, 2017 D6-27370-MAX-MLI 9.1

Figure 2: First page from the "Runaway Stabilizer" checklist (Thanhjono, 2019)

To further compound matters, the trim activation was not the only activity at the flight deck at that time. As the Lion Air flight climbed out, pilots faced and became preoccupied with a "speed disagree" message, and differing airspeed on the captain's and co-pilot's speed instruments. In addition, the altitude indications on the left and right instruments disagreed. While confronted with the double problem, pilots were also questioning whether to turn back, and negotiating with air traffic control on where to enter a holding, and on the – erroneous – altitude indication. The co-pilot was focusing on stopping the climb and finding a waypoint to fly to when the MCAS system started its first ten-second nose down trim action. In response, the captain trimmed the aircraft nose up again. In between, a momentary "bank angle" warning was sounded by the Enhanced Ground Proximity Warning System (EGPWS), then "air speed low" warnings (while the air speed was actually high), and the stick shaker started to activate. On the captain's speed display, which was using data by the same faulty sensor that activated MCAS, the overspeed limits and the low speed limits merged. With navigation tasks, feel differential pressure, flight control low pressure, unreliable airspeed, stick shaker, the crew was busy on multiple tasks, and did not start the stabilizer runaway checklist.

Strictly speaking, the condition listed in the checklist was not even met. Common in older generations of aircraft, uncommanded stabilizer trim leading to a runaway was the result of either a trim switch on the control yoke that does not disengage, or of a solenoid that stays stuck. The event is usually after a trim commanded by the pilot, in the absence of other signals and events, and it is consistent; a trim that was initiated does not stop after the trim switch is released, but continues, steadily, in the original direction. This trigger is described in the checklist, and this scenario is implied in the checklist description. As a young trainee, the first author observed this phenomenon several times in the video recordings of a flight experiment with a Swearingen Metro. The Metro's pilots were accustomed to the event, quickly pulled out the circuit breaker disconnecting the trim, cycled the trim switches and turned the circuit breaker back on.

Because the Metro pilots were accustomed to the phenomenon, they instantly recognized the inappropriate trim actions of the aircraft. To the new Boeing 737 MAX pilots, conventional trim problems were a thing of the past, but the MCAS produced a-typical trim behavior in a non-typical context (i.e., not after one of the two pilots used the trim) and during a very busy time. The nominal description in the checklist, "continuous" activation, does not match what happened. The MCAS actions come in bursts, and untypically, in the opposite direction of the pilot's trim direction, instead of continuing in the same direction. The pilots in both accident scenarios only had received a tablet-based training before transitioning to the new aircraft variant, and before the first accident did not have any information on the MCAS system. It is thus clear that they did not have experience with the pattern and context in which the MCAS activates the trim system.

Considering Rasmussen's SRK taxonomy and the simplified model, it is clear that time-critical actions in response to a trim runaway should be implemented as a combination of skill and rule-based behavior; knowledge based behavior is too slow for this condition, and this justly motivates Boeing's choice to designate the checklist items as memory items. On the Lion Air flight preceding the accident flight with the same aircraft, the pilots did manage to switch off the electric trim and thereby stop the MCAS actions. However, an additional pilot was present on that flight in the observer's seat. Possibly due to not being in the thick of controlling the aircraft, he had time to detect the largest threat as coming from the uncommanded trim actions, and instructed the pilots to use the cutout switches.

Rasmussen's diagram places the triggering signals for a procedure clearly in the skill domain. But simply by not having experienced MCAS activation previously, there is no skill to recognize the start condition for the checklist. The pattern recognition is not possible, and the sign for activating a rule-based sequence of actions is not produced. With the short time available, and the multitude of sensory impressions, surviving an erroneous MCAS activation on the original 737 MAX design is only likely if the condition for action is trained and recognizable as a familiar pattern. This can only be achieved if detection and distinguishing of the trigger is practiced in the simulator using scenarios that present the trigger in the correct context, possibly with a multitude of masking signals and detractors as present in the accident scenario.

In this case, in the absence of the possibility to recognize a familiar pattern, the pilots are left with a large set of observations, from the various instruments and warnings, and with parallel checklists and actions that these indications require, and they lacked the time and capacity to determine which was the most urgent issue and attend to it. In Rasmussen's terms, in this scenario the pilots were forced to resort to knowledge-based behavior, which can be slow, error-prone and is always laborious.

Wider implications

In the aftermath of the two accidents, the design of the MCAS system was re-evaluated, and its main problems are addressed. The insight that oversight by certifying authorities and a degradation in safety culture at Boeing contributed to unsafe designs and practices has prompted a wider review, and the correction of other identified issues, e.g., the lack of capacity of one of the flight control computers. However, based on Rasmussen's proven insights, wider implications can be seen. Each time a (foreseeable) failure condition requires immediate or near immediate attention from pilots, we should ensure that not only the pilots have the opportunity to train and drill to execute their response, but that they are also trained in recognizing the sign pattern in the multitude of signals available on the flight deck, not just in isolated failure cases, but also in busy scenarios where multiple triggers for multiple checklists compete.

This recognition requires more than pictures on a tablet or paper materials. Signal timing, strength, noise, vibration, possibly heat, multiple cues need to be approximately correct. From people working on flight simulation devices, we know that in the aftermath of the two accidents the hardware behind the motors providing the trim feel in the 737 simulators had to be significantly upgraded, to approach the levels of forces present in the actual aircraft. Also with the airline pilot's job progression from an active, hands-on flight to supervisor and programmer of automation for flight, the need for simulator fidelity does not diminish. In all scenarios where quick action is required to avert danger, the recognition of pilots must be trained until the patterns become familiar and recognizable even in the presence of significant and distracting signals, and in unexpected situations. Constant innovation in training scenarios is also needed to prevent pilots from recognizing a situation in a simulator session on the basis of expectation ("*second scenario, likely an engine failure during take-off*"), because that would train for false patterns that are not present in real flights.

Conclusion

With the introduction of the 737 MAX and its MCAS system, which was intended to keep the aircraft from stall situations, Boeing assumed that the procedures for runaway trim would enable pilots to correct for potential malfunctions in the MCAS system. Boeing's assumption on the pilots' ability to perform the necessary actions might have been correct, but the assessment on whether pilots would readily recognize the condition for action was not. In this paper, we propose a more accurate perspective on pilot behavior based on Rasmussen's theory. This implies that pattern and trigger recognition for situations that require near-immediate response must be trained in relevant training environments, with the proper cues, and under a variety of conditions. This is not only needed for MCAS activation, but for all checklists with memory items, and for all other situations that require a quick response while triggering conditions may be ambiguous or vague. We should not force people at the sharp end to know when action is required through cognitive reasoning, but instead we should ensure that they can train their pattern recognition skills.

References

Anon. (2019, September). *Assumptions used in the safety assessment process and the effects of multiple alerts and indications on pilot performance* (Safety Recommendation Report ASR1901). National Transportation Safety Board. Washington, DC.

Dekker, S. (2006). *The field guide to understanding human error*. Ashgate.

Landman, A., Groen, E. L., van Paassen, M. M. (, Bronkhorst, A. W., & Mulder, M. (2017). Dealing with unexpected events on the flight deck: A conceptual model of startle and surprise. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *59*(8), 1161–1172.

Landman, A., van Oorschot, P., van Paassen, M. M. (, Groen, E. L., Bronkhorst, A. W., & Mulder, M. (2018). Training pilots for unexpected events: A simulator study on the advantage of unpredictable and variable scenarios. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *60*(6), 793–805.

Pasztor, B., & Martin, E. (2020, December). *Statistical summary of commercial jet airplane accidents*. The Boeing Company.

Rasmussen, J. (1983). Skills, rules, and knowledge; signals, signs, and symbols, and other distinctions in human performance models. *IEEE Transactions on Systems, Man and Cybernetics*, *SMC-13*(3), 257–266.

Rasmussen, J. (1986). *Information processing and human-machine interaction : An approach to cognitive engineering*. North-Holland.

Roelen, A. L. C., & Klompstra, M. B. (2012). The challenges in defining aviation safety performance indicators. *International Probabilistic Safety Assessment and Management Conference and the Annual European Safety and Reliability Conference. Helsink, Finland.*

Tajima, A. (2004). Fatal miscommunication: English in aviation safety [ISBN: 0883-2919 Publisher: Wiley Online Library]. *World Englishes*, *23*(3), 451–470.

Thanhjono, S. (2019, October). *PT. lion mentari airlines boeing 737-8(MAX); PK-LQPTanjung karawang, west java* (Aircraft Accident Investigation Report KNKT.18.10.35.04). Komite Nasional Keselamatan Transportasi. Jakarta (IN).