COGNITIEF

Making sense of signals

Why it sometimes takes too long to understand a relevant cue

Numerous accidents and incidents in aviation and elsewhere are due to operators missing important signals from the environment. It seems that much more time is required to notice these signals compared to what we expect. These relevant cues are often perceptible in the physiological sense (the image reaches the retina) but are not perceived consciously. This is due to the enduring nature of mental models in our working memory that protect us from an overload of cues. A model is discussed that shows how our cognition interacts with reality to erratically reduce a mismatch between expectations and cues.

Robert J. de Boer

Aviation accidents

Numerous aviation accidents are a result of pilots missing relevant signs and warnings. On February 10th, 2010 a Boeing 737 inadvertently took off from a taxiway instead of runway 36C at Schiphol Amsterdam Airport, luckily missing other taxiing traffic by 300 meters. The dark and snowy circumstances were far from ideal, and some last-minute changes increased work pressure for the pilots, yet the taxiway lighting and markings were in compliance with international requirements and so should be perceptible enough (Dutch Safety Board, 2011b). Similarly, in the case of the Turkish Airlines Boeing 737 flight TK1951 that collided with the ground short of the runway in Amsterdam in 2009, the pilots missed the mode indication of the auto-pilot that is in the field of view of the flight crew (Dutch Safety Board, 2011a). Other examples of pilots missing signals and warnings include the Air Asiana flight in San Francisco in 2013 in which the speed indicator was not monitored, and not properly interpreting attitude and stall indications in the Air France 447 disaster over the South Atlantic in 2009. In both cases a severe accident resulted.

Yet these examples are exceptions, not the rule. In numerous other cases pilots have been able to perceive and interpret warnings and signals correctly. Pilots generally avoid taking off from taxiways despite snow and high work pressure. The exact same auto-pilot indication that surprised the Turkish Airlines crew had occurred at least 30 (!) times earlier in the same plane with the same cause (a faulty radio altimeter) without serious consequences (Dutch Safety Board, 2011a). And stall indications and iced pitot tubes occur not daily but certainly regularly.

One of the prime differences between a close call and an accident is the time available for recovery compared to the time that is needed to identify a problem. For instance, in flight TK1951 a single warning was available for 50 seconds and multiple other signals were available for 24 seconds thereafter (for a total of 74 seconds) before the shaking of the control column was triggered as a warning for an imminent stall (Silva & Hansman, 2015; Dutch Safety Board, 2011b, p. 92). This naturally triggered a pilot reaction but still the full extent of the problem (an auto-throttle stuck in retard mode) was not immediately understood, and so the remaining time was too short to prevent the aircraft hitting the ground a mile before the runway.

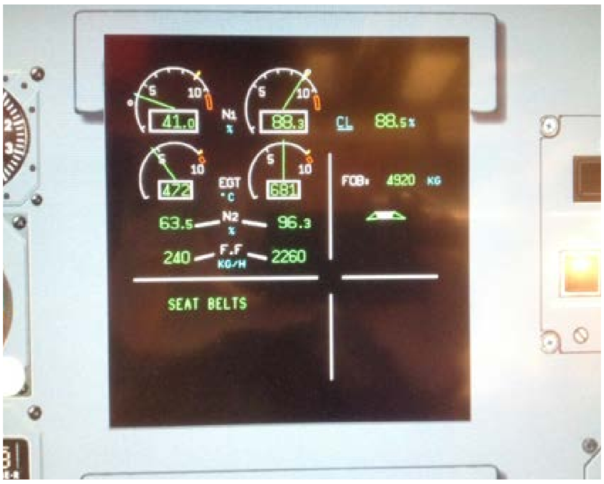

Figure 1. Indications of the Auto Throttle in idle showing four of the five cues: (1) a discrepancy between the N1 axis speed indications of the left and right engines; (2) deviations of the exhaust gas temperature (EGT); (3) rotation speed of the secondary axis N2; (4) deviation in fuel flow (F.F.) between the two engines.

The order of magnitude of more than a minute for TK1951 is consistent with the results of an analysis by Silva and Hansman (2015) of a number of different cases regarding auto-throttle mode confusion in aviation. The authors analysed the duration between the onset of a divergence between the ‘real’ aircraft state and the crew’s understanding of it. They found an average delay of 68 seconds between divergence and reconvergence of the crew’s mental models with the aircraft system state (2015, p. 322), and identified that in those cases where not sufficient time had been available to allow recovery, an accident occurred.

Laboratorium research

In a laboratory study using an Airbus A320 fixed base flight simulator, De Boer, Heems and Hurts (2014) found similarly long delays for pilots to identify a malfunction on the auto throttle even though multiple indications for the malfunction were in full view. In these experiments the participant was the Pilot Flying, and he or she was accompanied by a researcher who did not support the identification of the failure. The manipulation consisted of an auto throttle malfunction of the left engine that was fixed at idle power. Contrary to what the pilots expected, no warning messages appeared on the lower part of the upper Electronic Centralized Aircraft Monitor (ECAM). However, five indications were available to signal the malfunction: (1) a discrepancy between the primary axis speed indications of the left and right engines; (2) deviations of the exhaust gas temperature; (3) rotation speed of the secondary axis; (4) deviation in fuel flow between the two engines; and (5) rudder deflection indication, which was presented on the middle console next to the rudder trim knob. These cues were in direct sight of the pilots while the malfunction was present and are part of the standard scanning cycle. Four participants (11%) detected the failure within 45 sec, but 12 out of the 35 participants (34%) did not detect the failure within the total test time of 12 minutes. The time until failure detection was generally shorter for more experienced pilots.

Interestingly, this study confirmed the result of an earlier study (De Boer, 2012) in which it was found that the probability of detection of evasive cues follows a unimodal log-log probability distribution. Unimodal log-logistic probability density functions are characterized by a peak that occurs after a small delay and a long tail. In other words, there is a reasonable chance of a rapid response, but also a significant chance of a very delayed reaction. In the simulator study, the distribution for the more experienced pilots showed the peak *later* in time, but the tail was less pronounced. That means that extreme short or long delays are less expected with experienced participants compared to inexperienced pilots, and that there is more inter-subject consistency in the former group.

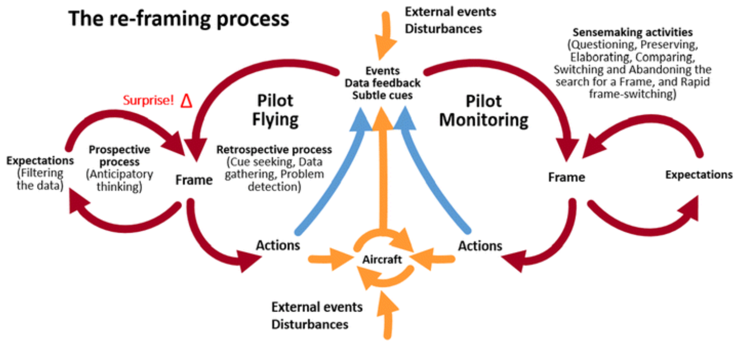

Figure 2. The crew-aircraft contextual control loop. Surprise (Δ) marks the cognitive realization that what is observed does not fit the current frame of thinking. Cues are ignored due to a previously existing frame or mental model, until a sudden awareness occurs of the mismatch between what is observed and what is expected. (Rankin, Woltjer, & Field, 2016, p. 633). Copyright © 2016 by the authors. Distributed under the terms of the Creative Commons Attribution 4.0 International License.

So why do pilots (or humans in general) not always perceive relevant cues despite the fact that these are in full view? According to Johnson-Laird (1983; 2006) humans are able to reason and solve problems because they construct simplified representations of the world around them in working memory. These mental models allow humans to operate effectively in routine situations by quickly grasping essential impressions with limited effort and saving cognitive resources for more complex tasks (Johnson-Laird, 1983; Kahneman, 2011). However, this phenomenon – as in the case of the taxiway take-off – may also lead to actions that in hindsight are incompatible with the requirements of the state of the world at that time. Relevant cues are often perceptible in the physiological sense (the image reaches the retina) but not are not perceived consciously.[1] Aspects of the divergence between mental model and system state have previously been described as fixation errors (De Keyser & Woods, 1990), inattentional blindness (Simons & Chabris, 1999; Mack & Rock, 1998), looking-but-not-seeing (Parasuraman & Manzey, 2010) cognitive lockup (Dekker, Cook, Johannesen, Sarter & Woods, 2010) or cognitive resistance (De Boer, 2012). This divergence plays a major role in many of the incidents in aviation described by Silva and Hansman (2015) and with automation in general (Manzey, Reichenbach & Onnasch, 2012).

Rankin, Woltjer and Field (2016) have suggested a sensemaking model that helps to explain this process. In this model *surprise* marks the cognitive realization that what is observed does not fit the current frame of thinking. Commencing with automation in aviation, the authors extend their model to include various operational issues between the crew and the aircraft. Other literature outside aviation (Schön, 1983) supports their suggestion that cues are ignored due to a previously existing frame or mental model, until a sudden awareness occurs of the mismatch between what is observed and what is expected. Their ‘crew–aircraft contextual control loop’ is represented in figure 2.

1 Sometimes the cues are brought into consciousness after the fact, as in a train driver passing a red sign and after the resultant collision remembering having seen it (Dutch Safety Board, 2013).

As Rankin, Woltjer and Field indicate, the contextual loop is a sensemaking model in that it assumes: '… perceiving and interpreting input from the environment after the fact (retrospective) [and] the continuous process of fitting what is observed with what is expected (anticipatory), an active process guided by our current understanding' (Rankin, Woltjer & Field, 2016, p. 625). The authors suggest that a mismatch between expectations and cues from reality will trigger a sudden surprise and effortful reframing. In the same vein Flach (2015) suggests that environment and mind interact to shape human experience, as a means for adapting to the functional demands of situations. Schön (1983, p. 69) suggests that the termination of the mental model is delayed but sudden, and followed by conscious effortful reasoning: 'The practitioner allows himself to experience surprise, puzzlement or confusion in a situation that he finds uncertain or unique. He reflects on the phenomena before him, and on his prior understandings which have been implicit in his behavior.'

De Boer and Dekker (2017) have been able to build support for the contextual control loop model through a field survey on automation surprise that was administered to a representative sample of 200 airline pilots. The data was used to empirically evaluate the crew–aircraft contextual control loop on three indicators: (1) a reduction in trust in the automated system following an automation surprise; (2) the way that the automation surprise was discovered, and (3) to which cause the automation surprise was attributed. The data the authors found fit well with the contextual control loop on all three points. (1) Despite experiencing an automation surprise, more than half of the respondents did not report a reduction in their trust of the automation. Only a small minority reported a strong reduction in trust. (2) The data showed that the discovery of the last automation surprise was predominantly (89%) by the respondent him/herself. This finding supports the contextual control loop model which suggests that re-framing may occur as a function of time without an external trigger. (3) The data indicate that a lack of understanding of the system, manual input issues, and *buggy*[2] knowledge (Dekker, 2014, p. 98-99) concerning aircraft systems were the predominant causes of automation surprise – as predicted by the contextual control loop.

Discussion

In the previous paragraphs I have proposed that even though we expect pilots (and operators in a supervisory role everywhere) to monitor their systems relentlessly, we cannot rely on any humans to perceive evasive cues immediately – even if they are in full view. Instead, the probability of detection varies rather a lot. This is due to the enduring nature of the mental models in our working memory that protect us from an overload of cues. The contextual control loop model shows how our cognition interacts with reality to erratically reduce a mismatch between expectations and cues through sudden surprise and effortful reframing. But the question that remains is: is it possible to reduce the negative effects of this phenomenon?[3]

From the literature we can indeed distill lessons to guard against the downside of cognitive resistance. These include announcing the unexpected, increased experience, more eyes, and a better mental model.

A simple, perhaps trivial solution that is often applied to counter *cognitive resistance* is to announce situations that otherwise might go unnoticed. Audible alarms, lights and even tactile warnings are often used to draw attention to cues that might otherwise be overlooked. Aviation examples include the Ground Proximity Warning (a voice saying 'pull up, pull up' or 'too low – terrain'), flashing lights to indicate reaching a beacon, and shaking of the control column as stall warning. Although each individual alarm may be useful, a combination of these might be confusing and could lead to cognitive overload (Hammer, 2010).

Experience has been shown to guard against unwanted cognitive resistance (De Boer, Heems & Hurts, 2014), but is hardly a quick fix. Gaining experience can be accelerated by training a wide range of scenarios (perhaps in a simulated environment) and allowing junior operators to accompany those more senior on unusual missions. An underutilized contribution to gaining experience is ensuring that operational feedback is available for all concerned: sharing experiences, considering alternatives retrospectively and discussing how work is actually done help to build experience more quickly than without.

De Boer and Soltani (2015, unpublished) found that for two pilots in comparison to a single pilot, the average time to perceive a malfunction were radically lowered by more than 50% and the success rate increased from 50% to 81% (figure 3). However, the same authors also found that performance of the dyad was dependent upon the amount of support and interaction of the second pilot (Pilot Monitoring), as determined through analysis of the frequency, duration and volume of his speech when compared to that of the Pilot Flying.

Finally, it is possible to guard against the down side of cognitive resistance through the eradication of *buggy* mental models. All too often cues are not perceived because the operators are not aware of the cues' importance in relation to the current task, as demonstrated by Turkish Airlines TK951 (Dutch Safety

2 A 'buggy' mental model occurs when there is a discrepancy between what the designer intended and what the crew thinks the automation should be doing; i.e. the system is working correctly but leading to operational issues.

3 Following (De Boer, 2012) we will use the term cognitive resistance.

Figure 3. Two pilots executing the experiment. All instruments are clearly visible on the touch screens. The pilots are wearing sensors on strings around their neck to determine relative frequency, duration and volume of their speech.

Board, 2011a) et cetera. The systems are not broken and they more or less operate as designed, yet the operators are insufficiently aware of the aircraft's state and behavior. Rather than treating these cases as a training deficiency, the sensemaking model suggests that our understanding of the interaction between humans and automation can be improved by considering systemic factors and the complexity of the operational context. Our sometimes-inconsistent relation with automation seems to be a manifestation of the system complexity and interface design choices, rather than the result of individual under-performance. Maybe it should be the designers that need to be trained, not the operators.

References

De Boer, R.J. (2012). *Seneca's Error: An Affective Model of Cognitive Resistance*. Delft, the Netherlands: Delft University of Technology.

De Boer, R.J., & Hurts, K. (2016). Automation Surprise: Results of a Field Survey of Dutch pilots. *Aviation Psychology and Applied Human Factors* (submitted).

De Boer, R.J., Heems, W., & Hurts, K. (2014). The duration of automation bias in a realistic setting. *The International Journal of Aviation Psychology, 24*(4), 287-299.

De Keyser, V., & Woods, D.D. (1990). Fixation errors: Failures to revise situation assessment in dynamic and risky systems. In *Systems reliability assessment* (pp. 231-251). Dordrecht: Springer.

Dekker, S. (2014). *The Field Guide to Understanding 'Human Error'*. Farnham, UK: Ashgate.

Dekker, S., Cook, D.R., Johannesen, D.L., Sarter, D.N., & Woods, P.D. (2010). *Behind human error*. Farnham: Ashgate Publishing, Ltd.

Dutch Safety Board. (2011a). *Crash near Schiphol, February 2009*. The Hague.

Dutch Safety Board. (2011b). *Take-off from Taxiway Amsterdam Airport Schiphol*. The Hague, the Netherlands.

Flach, J. (2015). Situation Awareness: Context Matters! A Commentary on Endsley. *Journal of Cognitive Engineering and Decision Making, 9*(1), 59-72. doi:10.1177/1555343414561087.

Hammer, J. (2010). Intelligent Interfaces. In J. Wise, V. Hopkin & D. Garland, *Handbook of Aviation Human Factors* (2 ed., Vol. Chapter 24). Boca Raton FL: Taylor & Francis.

Hollnagel, E., & Woods, D.D. (2005). *Joint Cognitive Systems: Foundations of Cognitive Systems Engineering*. Boca Ranton: CRC Press, Taylor & Francis Group.

Johnson-Laird, P.N. (1983). *Mental models: Towards a cognitive science of language, inference and consciousness*. Cambrdige, MA: Harvard University Press.

Johnson-Laird, P.N. (2006). *How we reason*. Oxford, UK: Oxford University Press.

Kahneman, D. (2011). *Thinking Fast, Thinking Slow*. New York, NY: Farrar, Straus, and Giroux.

Klein, G., Wiggins, S., & Dominguez, C. (2010). Team sensemaking. *Theoretical Issues in Ergonomics Science, 11*(4), 304-320.

Mack, A., & Rock, I. (1998). *Inattentional blindness*. Cambridge, MA: MIT press.

Manzey, D., Reichenbach, J., & Onnasch, L. (2012). Human performance consequences of automated decision aids: The impact of degree of automation and system experience. *Journal of Cognitive Engineering and Decision Making, 6*(1), 57-87.

Parasuraman, R., & Manzey, D.H. (2010). Complacency and bias in human use of automation: An attentional integration. *Human factors, 52*(3), 381-410.

Rankin, A., Woltjer, R., & Field, J. (2016). Sensemaking following surprise in the cockpit – a re-framing problem. *Cognition, Technology & Work, 18*(4), 623-642.

Schön, D. (1983). *The reflective practitioner: How practitioners think in action*. London: Temple-Smith.

Silva, S.S., & Hansman, R.J. (2015). Divergence between flight crew mental model and aircraft system state in auto-throttle mode confusion accident and incident cases. *Journal of Cognitive Engineering and Decision Making, 9*(4), 312-328.

Simons, D.J., & Chabris, C.F. (1999). Gorillas in our midst: Sustained inattentional blindness for dynamic events. *Perception, 28*, 1059-1074.

Soltani, P., & De Boer, R. (2015, unpublished). Understanding flight crew performance through the lens of honest signals. *Presentation at Human Factors & Ergonomic Society European Chapter*. Groningen, the Netherlands.

Samenvatting

Veel luchtvaartongelukken blijken het gevolg te zijn van het missen van een signaal ook al zijn die goed zichtbaar, zoals het opstijgen van een taxibaan of het ongeluk met Turkish Airlines bij Schiphol in 2009. Vaak blijkt al vaker iets soortgelijks zich te hebben voorgedaan, maar omdat er in die gevallen meer tijd beschikbaar was heeft dat niet tot een ongeluk geleid. Uit laboratoriumonderzoek blijkt dat het oppikken van onduidelijke signalen soms minuten kan duren, met een grote variatie tussen proefpersonen. In de literatuur is hiervoor een model beschreven die uitgaat van een steeds groter wordende kloof tussen realiteit en perceptie, totdat met een groot gevoel van verassing deze kloof plotseling wordt gedicht. De kloof kan eerder worden gedicht door waarschuwingssignalen, ervaring, meerdere waarnemers en het beter begrijpen van de systemen.

About the author:

Dr. R.J. de Boer
Amsterdam Campus of Northumbria University, Newcastle
Previously at the Aviation Academy, Amsterdam University of Applied Sciences
robert.deboer@northumbria.ac.uk