Dealing With Unexpected Events on the Flight Deck
A Conceptual Model of Startle and Surprise

Landman, Annemarie; Groen, Eric L.; van Paassen, M. M.(René); Bronkhorst, Adelbert W.; Mulder, Max

DOI
10.1177/0018720817723428

Publication date
2017

Published in
Human Factors: the journal of the human factors and ergonomics society

Citation (APA)
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. https://doi.org/10.1177/0018720817723428

Dealing With Unexpected Events on the Flight Deck: A Conceptual Model of Startle and Surprise

Annemarie Landman, Delft University of Technology, Delft, The Netherlands, **Eric L. Groen**, TNO, Soesterberg, The Netherlands, **M. M. (René) van Paassen**, Delft University of Technology, Delft, The Netherlands, **Adelbert W. Bronkhorst**, TNO, Soesterberg, The Netherlands, and **Max Mulder**, Delft University of Technology, Delft, The Netherlands

Objective: A conceptual model is proposed in order to explain pilot performance in surprising and startling situations.

Background: Today's debate around loss of control following in-flight events and the implementation of upset prevention and recovery training has highlighted the importance of pilots' ability to deal with unexpected events. Unexpected events, such as technical malfunctions or automation surprises, potentially induce a "startle factor" that may significantly impair performance.

Method: Literature on surprise, startle, resilience, and decision making is reviewed, and findings are combined into a conceptual model. A number of recent flight incident and accident cases are then used to illustrate elements of the model.

Results: Pilot perception and actions are conceptualized as being guided by "frames," or mental knowledge structures that were previously learned. Performance issues in unexpected situations can often be traced back to insufficient adaptation of one's frame to the situation. It is argued that such sensemaking or reframing processes are especially vulnerable to issues caused by startle or acute stress.

Conclusion: Interventions should focus on (a) increasing the supply and quality of pilot frames (e.g., though practicing a variety of situations), (b) increasing pilot reframing skills (e.g., through the use of unpredictability in training scenarios), and (c) improving pilot metacognitive skills, so that inappropriate automatic responses to startle and surprise can be avoided.

Application: The model can be used to explain pilot behavior in accident cases, to design experiments and training simulations, to teach pilots metacognitive skills, and to identify intervention methods.

Keywords: aviation, mental models, pilot performance, resilience, training

---

Address correspondence to Annemarie Landman, Faculty of Aerospace Engineering, TU Delft, Kluyverweg 1, 2629 HS, Delft, The Netherlands; e-mail: h.m.landman@tudelft.nl.

HUMAN FACTORS
Vol. 59, No. 8, December 2017, pp. 1161–1172
DOI: 10.1177/0018720817723428

INTRODUCTION

The increased use of automated systems has greatly improved aviation safety; however, it has also created some new challenges. Situations that cannot be handled by automated systems and that require human intervention are typically unforeseen and complex, demanding quick judgment and decision making (Militello & Hutton, 1998). Such situations may arise after long periods of automated flight, making it difficult to suddenly switch to an active role (Endsley, 1996; Young & Stanton, 2002). At the same time, automation may decrease the transparency of the flying process to the flight crew, which may lead to automation surprises (de Boer & Hurts, 2017; Sarter, Woods, & Billings, 1997). Furthermore, high reliability of automated systems may decrease active monitoring due to complacency (Parasuraman & Riley, 1997), and extensive use of automation may erode pilots' manual flying skills (Haslbeck & Hoermann, 2016).

In several recent flight safety events, such as those involving loss of control during flight, the unexpectedness of the situation is thought to have induced a "startle factor," complicating the crew's troubleshooting (Belcastro & Foster, 2010; Bürki-Cohen, 2010; Kochan, Breiter, & Jentsch 2004; Martin, Murray, Bates, & Lee, 2016; Shappell et al., 2007). In response to these events, new regulations include recommendations to incorporate startle and surprise in training programs to prepare flight crews for unexpected events (European Aviation Safety Agency [EASA], 2015; Federal Aviation Administration [FAA], 2015; International Civil Aviation Organisation [ICAO], 2013). The current paper is aimed at conceptualizing these processes in order to better understand and prevent potential pilot incapacitation in startling or surprising situations.

Differences Between Startle and Surprise

The terms *startle* and *surprise* are often used interchangeably in aviation operational practice (Rivera, Talone, Boesser, Jentsch, & Yeh, 2014). Nevertheless, several authors have pointed out that startle and surprise are different responses, with different causes and effects (e.g., Bürki-Cohen, 2010; Martin, Murray, Bates, & Lee, 2015; Rivera et al., 2014). A startle is a brief, fast, and highly physiological reaction to a sudden, intense, or threatening stimulus, such as the sound of a pistol shot (Ekman, Friesen, & Simons, 1985; Martin et al., 2015; Thackray, 1988). Measurable aspects of startle include eye blinks, contraction of facial and neck muscles, arrest of ongoing behaviors, increased physiological arousal, and reports of fear or anger. Although unexpectedness increases the response, anticipated stimuli were shown to be startling as well (Damasio, 1999; Ekman et al., 1985; Hagemann, Levenson, & Gross, 2006; Roberts et al., 2004). A typical example of an event in aviation that is startling but not very surprising would be a lightning strike when flying in stormy weather.

Surprise is an emotional and cognitive response to unexpected events that are (momentarily) difficult to explain, forcing a person to change his or her understanding of the situation (Foster & Keane, 2015; Meyer, Reisenzein, & Schützwohl, 1997; Schützwohl, 1998; Teigen & Keren, 2003). Surprise may occur in the absence of startle when an event is appraised as odd and curious. Although surprise, like startle, increases arousal and draws attention to its cause, it does so in a more orienting manner (i.e., the orienting response) and less in a defensive or "flinching" manner (Bradley, 2009). Examples of highly surprising events in aviation include subtle technical failures or automation surprises that are "baffling" and difficult to explain. Studies indicate that surprises occur quite frequently in aviation, but most of them remain inconsequential (Hurts & de Boer, 2014; Kochan et al., 2004). However, in extreme cases, surprise may impair the crew's troubleshooting capabilities, as we will describe next using our conceptual model.

A CONCEPTUAL MODEL OF STARTLE AND SURPRISE

The differences between surprise and startle raise questions regarding ground-based training to prepare flight crew for unexpected events in flight. Would a sudden and loud noise in the simulator be sufficient to simulate difficulties associated with in-flight emergencies (Thackray, 1988)? Or should training scenarios primarily involve unexpectedness (Bürki-Cohen, 2010)? To answer these questions, some authors have focused on the causes and effects of surprise (e.g., Kochan et al., 2004; Rankin, Woltjer, & Field, 2016), and others have described those of startle (Martin et al., 2015). In the current paper, we present a conceptual model (Figure 1) that brings the existing knowledge about startle and surprise together. The model is a synthesis of elements of the cognitive-psychoevolutionary model of surprise (Meyer et al., 1997), the perceptual cycle model (Neisser, 1976), the data/frame theory of sensemaking (Klein, Phillips, Rall, & Peluso, 2007), and literature on startle and acute stress.

Elements of the Model

The perceptual cycle. The bold lines in the model represent the perceptual cycle: A person perceives stimuli, interprets these stimuli, assesses the situation (appraisal), and selects and executes actions, which may generate new data. Appraisal is modeled in such a way that it can be fast and highly automatic in some cases, or it may also involve a more slow, effortful, and knowledge-based processing (Kahneman, 2003; Rasmussen, 1983). Action selection (decision making) is modeled so that it is an integral part of the perceptual cycle, which thus represents a continuous process of hypotheses generation and testing (Flach, Feufel, Reynolds, Parker, & Kellogg, 2017). For simplicity, the model does not discern different levels of control at which perceptual cycles may occur in parallel, such as in Hollnagel's extended control model (Hollnagel & Woods, 2005).

Startle. On the left side of this perceptual cycle, the startle response is pictured. This response results from a fast, sometimes reflexive, appraisal of a stimulus as threat–related (Globisch, Hamm, Esteves, & Öhman, 1999).

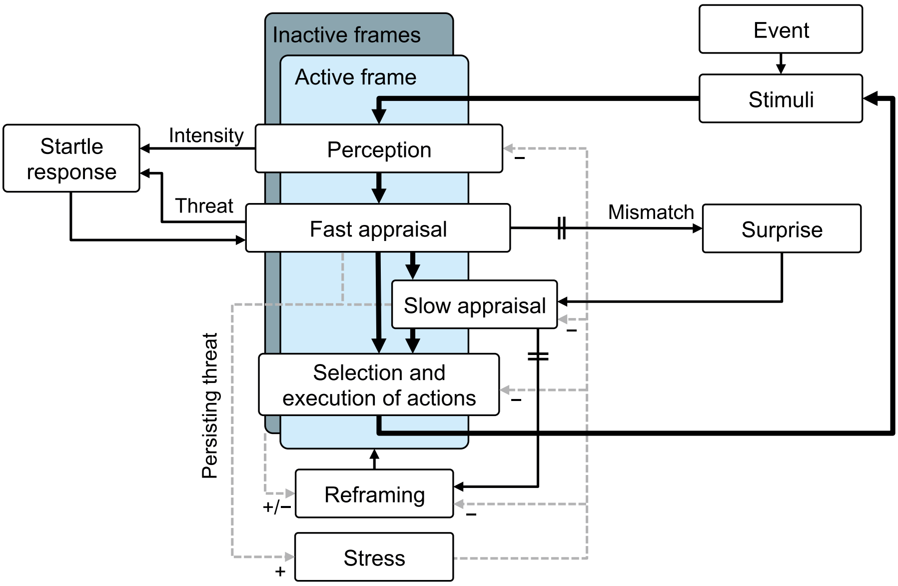

Figure 1. Conceptual model of startle and surprise. Solid lines indicate sequenced events. Dashed lines indicate potential influences, with plus signs indicating an increasing effect and minus signs indicating an impairing effect. Double lines indicate thresholds.

Startle is modeled to cause a closer examining of the triggering stimulus, which may lead to further increase of stress (dashed line; Martin et al., 2015). If startle occurs in the absence of surprise, only the left loop (*startle response*) is activated, and the appraisal process will remain relatively fast. However, if the appraisal of a startling stimulus brings momentarily unexplainable information to light, the right loop (*surprise*) will subsequently be activated. The perceptual cycle then continues, either with actions in response to the threat or by resuming as before in case of a false alarm.

Frames. In order to explain the causes and effects of surprise, the concept of frames is useful. A *frame* is defined as an explanatory structure, such as a story, map, or plan, which links perceived individual data points together and gives them meaning (Klein et al., 2007). Frames synthesize concepts, such as schemata, mental models, scripts, and other types of knowledge structures in long-term memory, that describe generic or specific situations, how things work, how events are sequenced, and which actions are appropriate (Bartlett, 1932; Brewer & Nakamura, 1984; Neisser, 1976; Piaget, 1976; Schmidt, 1975; Weick, 1995). Frames are created based on previous experiences (i.e., bottom up) so that understanding of a new situation or concept can be achieved and stored in memory (the supply of inactive frames in Figure 1). If a situation occurs in which the frame-related knowledge can be applied, a corresponding frame may be activated and applied (see Stanton, Salmon, Walker & Jenkins, 2009). Frames are thus instrumental for the achievement of higher levels of situation awareness (i.e., comprehension and projection) based on a lower level of situation awareness (i.e., perception) in the terms of Endsley's (1995) model.

Besides being shaped based on incoming data (bottom–up), frames are thought to actively select, filter, and provide meaning to incoming data (i.e., top–down; Neisser, 1976). They are thought to play a significant role in skilled performance, as frames structure complex stimuli and action sequences into manageable "chunks" based on the existing constraints (see Flach, Mulder, & Van Paassen, 2004; Klein, 1993). This is why, for instance, expert chess players are able to perceive and reproduce chess positions very quickly and accurately, as long as the positions make sense in terms of the game. In our model, we have illustrated the influence of the frame on perception, appraisal, and action by placing it behind these elements of the perceptual cycle, rather than making it an integral part of the perceptual cycle (Neisser, 1976). This way, we indicate that perception and action are still possible—although difficult—when there is no fitting frame activated. The model is simplified in that it represents merely one active frame, distinct from other frames. In reality, people are thought to use a number of frames at once, which are highly interconnected or nested and have no clear boundaries.

The use of frames to explain performance during surprise events in aviation has recently gained interest (e.g., Kochan, 2005; Rankin et al., 2016). In the latter study, pilot performance is modeled as the interaction of a crew with the aircraft and the environment using frames, anticipatory thinking, and expectations. The authors discuss an extensive list of sensemaking activities following surprise event cases in aviation. In our current model, we aim to add to their model by illustrating how the frame interacts with the perceptual cycle and how or why certain performance issues may occur.

Surprise. In the perceptual cycle, hypotheses based on the active frame are continually applied and tested with regard to their practical consequences (abduction; see Flach et al., 2017). As long as the results are consistent with the hypotheses, the active frame becomes strengthened in memory. However, a mismatch between feedback and the active frame will induce a surprise (Meyer et al., 1997), given that the mismatch exceeds a certain assumed threshold (double intersecting lines before *surprise* in Figure 1; e.g., Senders, 1964). This threshold indicates a form of confirmation bias, as events of low salience are more easily missed when they are deemed unlikely within the active frame (see, e.g., Wickens, Hooey, Gore, Sebok, & Koenicke, 2009).

Slow appraisal: Sensemaking. Appraisal of a surprise event involves sensemaking activities, or efforts to understand the cause of the mismatch between the encountered data and the active frame (Klein et al., 2007). Sensemaking is an explorative process that is active, analytical, conscious, and potentially effortful, characterized by top-down or goal-directed processing (Kahneman, 2003). Due to its active nature, it may be particularly problematic when pilots are not mentally prepared, for example, after a long period of automated flight (Young & Stanton, 2002). Sensemaking activities can be categorized into three groups (Klein et al., 2007; Weick, 1995). First, if the surprising data are determined to be the result of a misperception, the active frame can be preserved. Second, if the surprising data are being judged as correct, the active frame may not be detailed enough to account for them, in which case it can be elaborated (i.e., assimilation; Piaget, 1976). Third, if the data are being judged as correct, and they are fundamentally inconsistent with the active frame (i.e., a fundamental surprise; Lanir, 1986), a paradigm shift is required and a new frame should replace the active frame (i.e., accommodation; Piaget, 1976). This sensemaking activity is modeled as the element *reframing* being connected to the (transformation of the) active frame in Figure 1. People were shown to avoid considering a fundamental surprise as being the causal factor for mismatches, perhaps as a mechanism to reduce unnecessary efforts (i.e., frame fixation; Chinn & Brewer, 1993; De Keyser & Woods, 1990), indicated by a threshold toward *reframing* in Figure 1.

Reframing. A frame switch, or reframing, occurs when one restructures the way in which a situation is represented. Previously perplexing information may suddenly “fall into place,” and the appropriate responses become obvious. In contrast, the adoption of an inappropriate frame or the loss of a fitting frame may lead to a complete “loss of grip” on the situation, as there is no frame in place to guide perception, appraisal, and action. This may negatively affect the pilot’s ability to track what is going on (loss of Level I situation awareness; Endsley, 1995) or lead to information overload. Data can no longer be appraised in relation to other data and therefore lose meaning. The selection and execution of actions become reactive and sequential (bottom-up controlled) instead of anticipatory and proactive (top-down controlled), which may lead to tunnel vision or cognitive lockup (Sheridan, 1981). The involvement of acute stress may be even more deteriorative, as we will discuss next.

Stress. Both startle and surprise may cause acute stress, which constitutes the appraisal of present demands as taxing or exceeding one's resources and endangering one's well-being (Lazarus & Folkman, 1984; dashed lines with plus signs in Figure 1). Startle may increase stress very briefly and rapidly at first, and subsequent appraisal of the startling stimulus as threatening may cause a further increase in stress (Martin et al., 2015). Surprise may also cause stress, as it may pose, on the one hand, an increase in task demands to solve the situation and, on the other hand, a perceived decrease of available resources when one becomes aware of the inadequateness of the active frame.

The function of stress is to facilitate the recruitment of additional resources to respond effectively to demanding circumstances. However, aspects of stress, such as impaired top-down and increased stimulus-driven attentional control, emotions of fear and frustration, excessive physiological arousal, or performance rigidity, may also impair a pilot's cognitive and motor performance (Dismukes, Goldsmith, & Kochan, 2015; Eysenck, Derakshan, Santos, & Calvo, 2007; Nieuwenhuys & Oudejans, 2012; Wickens, Stokes, Barnett, & Hyman, 1993). Stress is thus modeled to impair perception, appraisal, action, and reframing (dashed lines with minus signs in Figure 1). Stress can be expected to particularly impair slow appraisal and reframing, as these are relatively more analytical, top-down or goal-directed processes. Stress is thought to cause a shift from analytical skills toward intuitive judgment, making one susceptible to biases (Kowalski-Trakofler, Vaught, & Scharf, 2003). This bias may, for instance, cause the incorrect application of a partially fitting frame that is easily retrieved from memory due to recent experiences.

Influencing Factors and Intervention Methods

In this section, several factors, which have previously been identified as affecting pilot performance in surprising or startling situations, are described and related to our model.

Domain expertise. One of the factors that facilitate pilot performance in surprising situations is domain expertise, or accumulated knowledge and skills through practice and experience. By applying and testing hypotheses based on frames in a large number of situations, these frames become more accurate and more fixed in memory (see Kochan, 2005), which allows one to easier relate new situations to those that have previously been encountered and to make decisions in a quick manner (Klein, 1993). In the literature, some results indeed indicate beneficial effects of pilot expertise on problem assessment and flexibility in unfamiliar scenarios (Gillan, 2003; McKinney & Davis, 2003), whereas other results suggest no effects or even somewhat detrimental effects (Kochan, 2005; McKinney & Davis, 2003), perhaps due to counterintuitive actions being more difficult to perform when certain frames have become tightly fixed through experience (Kochan, 2005).

Judgment skills. Domain-independent judgment skills, such as decision-making skills, cognitive flexibility, and metacognitive skills, were found to improve pilot performance following surprise in one study (Kochan, 2005). Such skills could be tested in the selection process, and certain judgment skills are thought to be trainable as well (see Kochan, 2005). Decision-making skills involve capabilities of problem analysis (sensemaking) and action selection. Cognitive flexibility involves reframing abilities. Our model may in particular be useful to increase metacognitive skills in pilots, which include the recognition of frame mismatches and potential reframing issues. By recognizing such situations, pilots can apply learned coping strategies, such as taking a moment to "breathe" and reflect or returning to more transparent and understandable configurations or autopilot modes.

Variable training. Researchers and aviation safety organizations emphasize the need for training with a variety of situations or scenarios (e.g., Bürki-Cohen, 2010; Casner, Geven, & Williams, 2013; EASA, 2015; FAA, 2015; ICAO, 2013; Kochan et al., 2004; Rankin et al., 2016). Training variability can be applied to reduce predictability so as to stimulate sensemaking activities and to improve reframing skills. Training variability is also thought to increase the number and elaborateness of available frames (e.g., Van Merriënboer, 1997). A more elaborate frame is thought to discriminate better between situations, aiding the generation of accurate hypotheses, the detection of data/frame mismatches, and the selection of an appropriate frame based on the available data (see the plus sign on the line from the inactive frames toward *reframing* in Figure 1; Gioia & Poole, 1984; Phillips, Klein, & Sieck, 2004). Experiencing examples of a concept in a variety of situations may improve one's understanding of the concept, facilitating the transfer of the knowledge and skills to new situations (Klein, 1993). In contrast, one-sided training of a small number of situations or (combinations of) failures may increase the risk of an inappropriate selection of these frames in stressful situations (the minus sign on the line toward *reframing* in Figure 1; Kowalski-Trakofler et al., 2003).

Practical training. Literature indicates that theoretical training should be enhanced with practical experience and feedback on performance so that the frame-related knowledge is linked to other knowledge, environmental cues, and actions (Phillips et al., 2004). Our model indicates that action selection in operational practice is an inherent part of the perceptual cycle, meaning that mere theoretical training is likely insufficient. For instance, scenario-based training (Summers, 2007) is based on the concept that knowledge cannot be fully understood independent from its context. This means that training should not be focused on specific maneuvers that are laid out in advance, but on the pilot's own decisions in response to a situation that is presented. Practical training may also be used in combination with exposure to a manageable amount of stress or startle, to make skills more robust to the effects of stress (Driskell, Salas, Johnston, & Wollert, 2008). This would decrease the detrimental effects of stress on other elements in our model (dashed lines with minus signs in Figure 1).

Fatigue. Fatigue is known to degrade logical reasoning and accurateness of performance, as well as to increase inattentiveness and the tendency toward preservation (Caldwell, 1997). Fatigue can thus be expected to increase confirmation biases (increase the thresholds; Figure 1), as well as to impair mentally taxing activities of sensemaking and reframing.

Flight deck design. Display designs that enhance situation awareness (Endsley, 1995) may aid in quicker recognition of anomalies by making mismatching data more salient. Our model suggests that the interpretation of a display system may be straightforward when the appropriate frame is already activated but not when a surprise occurs. Thus, interfaces designed for use in surprising situations (e.g., upset recovery display aids) should be tested in conditions in which surprise is sufficiently accurately simulated (see Implications for Experimental Design and Simulation section). Transparent automated systems (Endsley, 1996; Sherry, Feary, Polson, & Palmer, 2001) that aim to keep the pilot in the loop may help to update the active frame when a situation changes. Displays can also be designed to aid the sensemaking process (e.g., Muhren & Van de Walle, 2010). For instance, ecological interface design is intended to structure complex relationships between information in such a way that constraints become self-evident, decreasing the need for the pilot to construct frames for these relationships (e.g., Borst, Sjer, Mulder, Van Paassen, & Mulder, 2008).

IMPLICATIONS FOR EXPERIMENTAL DESIGN AND SIMULATION

As outlined in the model, startle and surprise have different causes and different effects, which means that different factors should be manipulated depending on whether the aim is to induce mainly startle or mainly surprise. The key element for inducing surprise is to set up a situation that mismatches with a previously activated frame. A mismatch that is not immediately understood would increase the effort required to reframe the situation, which may be useful for training purposes. Surprise and reframing can thus be elicited, for instance, through explicit misinformation, by presenting a number of similar scenarios followed by one that is subtly different, by presenting a situation that is subtly different from one that is well known to pilots, or through variation or novelty.

Although a surprising stimulus can be subtle, a startling stimulus should be highly salient (see Differences Between Startle and Surprise section). A startling stimulus can be a loud and abrupt sound or a sudden, uncommanded motion of the aircraft. Unexpectedness may increase salience and perceived threat, but in contrast to the manipulation for surprise, a startling event does not need to require sensemaking or reframing (e.g., in the case of a lightning strike). For an extensive list of surprising or startling flight scenarios, see Martin et al. (2015).

PREVIOUS EXPERIMENTAL STUDIES ON SURPRISE AND STARTLE

To date, few experimental studies focusing specifically on surprise and startle in the cockpit have been published. The studies indicate that pilot performance may decrease significantly, even when skills and procedures were practiced shortly beforehand. In the concise review that follows, we link the experimental studies to our model. As the reports do not always explicitly mention whether the participating pilots were surprised, startled, or both, we have tried to infer this reaction from the manipulations used.

In two studies, pilots had to detect, recognize, and respond to unannounced problems, such as aerodynamic stalls, wind shears, or automation failures (Beringer & Harris, 1999; Casner et al., 2013). The results showed that response times were longer after surprising compared with nonsurprising events, with some participants responding exceptionally late. Similar results were found in a simulator study by Martin et al. (2016), in which pilots were tasked with flying the same missed approach, once with and once without an unexpected fire alarm and a loud explosion sound. Although the startling stimulus did not require a change of plans, the stimulus resulted in a delayed initiation of the missed approach in one third of the pilots. In regard to our model, the frame-incongruent information in these experiments likely caused a surprise, and the highly salient stimulus in the experiment by Martin et al. (2016) was likely startling as well. Our model explains such later responses as being caused by inattentiveness to frame-incongruent information, or by slow appraisal processes delaying or interfering with actions.

Some studies also showed impairments of performance in terms of the incorrect or incomplete application of procedures. Pilots in the study by Casner et al. (2013) displayed difficulty with recognizing and responding correctly to an unexpected wind shear compared to an expected one. Schroeder, Bürki-Cohen, Shikany, Gingras, and Desrochers (2014) actively misled pilots into expecting a different upcoming event. During final approach, an unexpected aerodynamic stall, induced by a sudden tailwind, was inserted in the scenario. The results indicated that 78% of the pilots made errors in executing the stall recovery template, even though they had applied it many times beforehand. A check of the subjective impact of the manipulation confirmed that all pilots were highly surprised by the event. Whether they were also startled or stressed is not clear. The study did not include a control condition to confirm whether the performance degradation was attributable to the surprise. For this reason, we recently performed a simulator study in which pilots were exposed twice to an aerodynamic stall: once in a surprise condition and once in an anticipation (control) condition (Landman, Groen, Van Paassen, Bronkhorst, & Mulder, in press). The results showed that, compared to the control condition, the proportion of pilots adhering to the recovery template decreased by around 25% in the surprise condition, whereas measures of surprise, startle, and mental workload increased significantly. According to our model, this performance impairment would result from reframing efforts, as a frame switch is needed before one can respond accurately to the unanticipated event.

APPLYING THE MODEL TO FLIGHT SAFETY INCIDENTS

In this section, we will evaluate four aviation incidents or accidents in the context of our model (see Figure 2). These four cases were selected because they seem to demonstrate several different aspects of our model. We focus in particular on potential causes of reframing issues and on the effects of reframing issues on perception, appraisal, and action (see also Rankin et al., 2016).

Case 1

The accident of Flash Airlines Flight 604 in 2004 (Ministry of Civil Aviation, 2004) suggests that pilot spatial disorientation (Previc & Erco-

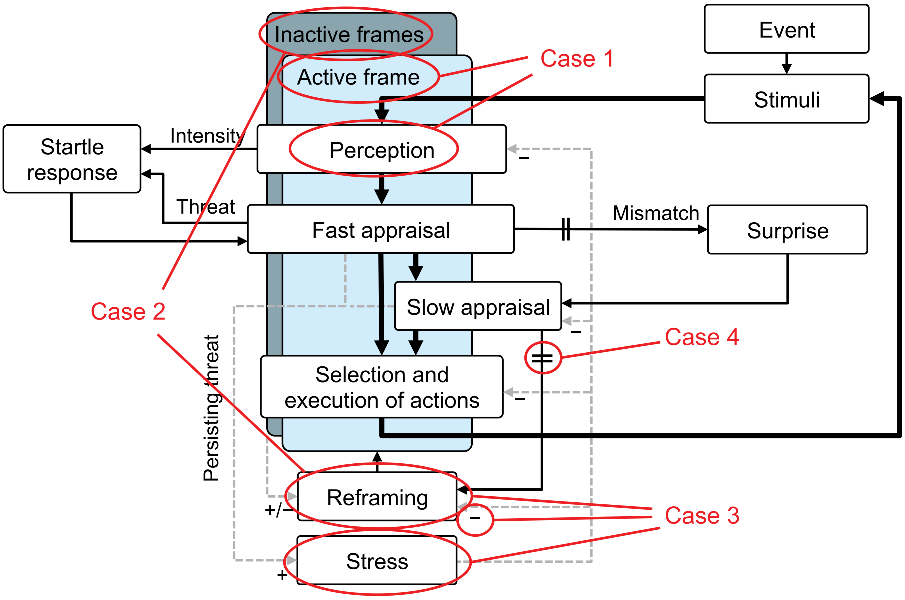

Figure 2. Estimated causal factors in the four cases as mapped onto the conceptual model of startle and surprise.

line, 2001) of the captain (pilot flying) played a significant role in the development of the event, although other causes of the accident have not been ruled out by all investigating parties. The captain had initiated a long, left climbing turn, during which the aircraft transitioned from a left bank to a right bank at a rate below the detection threshold of the vestibular system (Mumaw et al., 2016). When the first officer alerted the captain to the right turn ("Aircraft turning right, sir"), the captain expressed surprise ("How turning right?"). Next, he seemed to recognize that the attitude was indeed off ("Ok, come out"). According to our model, there was at that moment likely a mismatch between the captain's frame (aircraft turning left) and the first officer's assertion of the aircraft turning right. Next, instead of rolling to wings level, the captain gave further roll inputs to the right, leading to an overbank and ultimately to loss of control. This suggests that reframing did not occur following the surprise, and that the incorrect frame of a left bank remained active. Because the active frame influences perception (Case 1 in Figure 2), this frame of banking left may have induced an incorrect perception of the artificial horizon (Previc & Ercoline, 1999). A similar sequence of events may have occurred in the Crossair Flight 498 accident in 2000, indicating that hazardous frame-induced misperceptions of displays may occur more often.

Case 2

The incident with a B-737 near Brisbane, Australia, in 2013 ("B738, En-Route," 2013) may be an example in which an inactive frame influences the reframing process with negative consequences. While approaching the glide slope beam of Brisbane airport at night, the aircraft unexpectedly began to climb due to an earlier unintended selection of an autopilot mode. The crew quickly noticed this and disconnected the autopilot mode. Later, during the descent, the aircraft began to bank to the left due to a residual rudder deflection that was previously corrected for by the autopilot. This motion was again detected, but the crew incorrectly assumed that it was induced by the autopilot. After 80 s, the crew realized that the autopilot was not engaged, and they corrected the deviation manually. In our model (Case 2 in Figure 2), these actions are explained as caused by an influence of the previously activated frame on the reframing process. Because of the recent events in the incident, the frame of unintended explaining the events as caused by autopilot activation was perhaps most easily retrievable from memory, such that it was incorrectly applied again to the new situation.

Case 3

The accident of Air France Flight 447 in 2009 (Bureau d'Enquêtes et d'Analyses pour la Sécurité de l'Aviation Civile [BEA], 2012) seemed to involve a negative spiral of reframing issues and high stress (Case 3 in Figure 2). The accident report indicates that there were several signs that the crew were unable to identify an aerodynamic stall situation (BEA, 2012, pp. 179–180), which followed unreliable airspeed indication and autopilot disengagement during cruise. Cues indicating stall, such as buffeting and the auditory stall warning, were not appraised as such, and potentially led to incorrect reframing to an overspeed situation. The report reads that a lack of exposure to aerodynamic stall situations, in contrast to the well-known dangers of overspeed, may have caused the crew to fixate on the overspeed explanation of events. As was also described in the previous case (Case 2), it could be that the frame of an overspeed situation was more easily retrievable from memory, such that it influenced the reframing process. The accident report also reads that there were signs of excessive stress, which may have exacerbated the pilots' inability to reframe correctly. Strong initial pitch and roll inputs immediately following the autopilot disengagement suggest that the pilot flying was not only surprised but perhaps also startled by the sudden autopilot disconnect.

Case 4

West Caribbean Airways Flight 708 in 2005 ("JIAAC-9-058-2005," 2005) seems to be an example of frame fixation following a switch toward an inappropriate frame (Case 4 in Figure 2). Leading up to the accident, the aircraft's anti-icing systems were turned on at too high an altitude, so that sufficient engine performance could not be maintained. Subsequent loss in airspeed, loss in engine power, and autopilot-induced changes in attitude went unnoticed. An aerodynamic stall ensued, causing a further decrease in engine power due to variations of airflow into the engines. According to the voice recorder, the captain (pilot flying) misdiagnosed the problem as an engine flameout (reframed to an incorrect frame) and gave nose-up inputs. It seems that the captain then fixated on this incorrect frame, and disregarded the first officer's two callouts of an aerodynamic stall as well as the stall warnings of the system. It also seems that these reframing issues were not preceded by startle. In contrast, the pilots seemed to underestimate the gravity of the situation as they mentioned no checklists and declared no emergency despite making contact with air traffic control.

CONCLUSION

We propose an integrated model, which explains the effects of both startle and surprise responses to unexpected events in the cockpit. Examples of flight safety events show that inappropriate crew responses do not always involve startle but can often be traced back to surprise, which indicates a mismatch between what is being perceived and the pilot's active frame. The model explains such inappropriate responses as resulting from reframing issues following the mismatch, issues that can be exacerbated by startle, acute stress, fatigue, or unclear and complex interface designs. Information mismatching with an active frame may also remain unnoticed or be incorrectly interpreted, meaning that a loss of situation awareness may occur.

By explaining inappropriate or absent responses to unexpected situations as reframing issues, we emphasize that intervention methods should be focused on instilling a supply of sufficiently elaborate frames. Toward this end, we suggest using a variety of training scenarios to increase pilots' frame supply and elaborateness, using unpredictable and practical training to practice reframing skills, and using transparent interface designs—tested for effectiveness in surprising situations—to aid in framing or reframing. Finally, our model provides an aid to increase pilots' metacognitive skills of recognizing and understanding the hazards involved in frame mismatches and reframing issues.

KEY POINTS

- Pilot performance is described as taking place within *frames*, or structures of learned knowledge with regard to systems or situations.
- A conceptual model is presented in which surprise is related to an adaptation or switch of one's active frame, a process that is particularly vulnerable to effects of startle or acute stress.
- The model proposes that pilot performance in surprising and startling circumstances depends on frame supply and frame adaptation skills, which may be improved, for instance, through variable and unpredictable training.

REFERENCES

B738, en-route, south south west of Brisbane Australia. (2013). Retrieved from http://www.skybrary.aero/index.php/B738,_en-route,_south_south_west_of_Brisbane_Australia,_2013

Bartlett, F. C. (1932). *Remembering: An experimental and social study*. Cambridge, UK: Cambridge University.

Belcastro, C. M., & Foster, J. V. (2010). Aircraft loss-of-control accident analysis. In *Proceedings of AIAA Guidance, Navigation and Control Conference* (Paper No. AIAA 2010-8004). Reston, VA: AIAA.

Beringer, D. B., & Harris, H. C., Jr. (1999). Automation in general aviation: Two studies of pilot responses to autopilot malfunctions. *International Journal of Aviation Psychology*, *9*, 155–174.

Borst, C., Sjer, F. A., Mulder, M., Van Paassen, M. M., & Mulder, J. A. (2008). Ecological approach to support pilot terrain awareness after total engine failure. *Journal of Aircraft*, *45*, 159–171.

Bradley, M. M. (2009). Natural selective attention: Orienting and emotion. *Psychophysiology*, *46*, 1–11.

Brewer, W. F., & Nakamura, G. V. (1984). The nature and functions of schemas. In R. S. Wyer & T. R. Srull (Eds.), *Handbook of social cognition* (pp. 119–160). Hillsdale, NJ: Lawrence Erlbaum.

Bureau d'Enquêtes et d'Analyses pour la Sécurité de l'Aviation Civile. (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*. Le Bourget, France: Author.

Bürki-Cohen, J. (2010). Technical challenges of upset recovery training: Simulating the element of surprise. In *Proceedings of the AIAA Modeling and Simulation Technologies Conference* (Paper No. AIAA 2010-8008). Reston, VA: AIAA.

Caldwell, J. A., Jr. (1997). Fatigue in the aviation environment: An overview of the causes and effects as well as recommended countermeasures. *Aviation, Space, and Environmental Medicine*, *68*, 932–938.

Casner, S. M., Geven, R. W., & Williams, K. T. (2013). The effectiveness of airline pilot training for abnormal events. *Human Factors*, *55*, 477–485.

Chinn, C. A., & Brewer, W. F. (1993). The role of anomalous data in knowledge acquisition: A theoretical framework and implications for science instruction. *Review of Educational Research*, *63*, 1–49.

Damasio, A. R. (1999). *The feeling of what happens: Body and emotion in the making of consciousness*. New York, NY: Harcourt Brace.

de Boer, R. J. & Hurts, K. (2017). Automation surprise: Results of a field survey of Dutch pilots. *Aviation Psychology and Applied Human Factors, 7*(1), 28–41.

De Keyser, V., & Woods, D. D. (1990). Fixation errors: Failures to revise situation assessment in dynamic and risky systems. In A. G. Colombo & A. Saiz de Bustamante (Eds.), *Systems reliability assessment* (pp. 231–251). Dordrecht, Netherlands: Springer.

Dismukes, R. K., Goldsmith, T. E., & Kochan, J. A. (2015). Effects of acute stress on aircrew performance: literature review and analysis of operational aspects. (NASA/TM-2015-218930). Moffett Field, CA: National Aeronautics and Space Administration.

Driskell, J. E., Salas, E., Johnston, J. H., & Wollert, T. N. (2008) Stress exposure training: An event-based approach. In P. A. Hancock & J. L. Szalma (Eds.), *Performance under stress* (pp. 271–286). Hamphire, UK: Ashgate.

Ekman, P., Friesen, W. V., & Simons, R. C. (1985). Is the startle reaction an emotion? *Journal of Personality and Social Psychology*, *49*, 1416–1426.

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

Endsley, M. R. (1996). Automation and situation awareness. In R. Parasuraman & M. Mouloua (Eds.), *Automation and human performance: Theory and applications* (pp. 163–181). Boca Raton, FL: CRC Press.

European Aviation Safety Agency. (2015). *Loss of control prevention and recovery training: Notice of Proposed Amendment 2015-13*. Cologne, Germany: Author.

Eysenck, M. W., Derakshan, N., Santos, R., & Calvo, M. G. (2007). Anxiety and cognitive performance: Attentional control theory. *Emotion*, *7*, 336–353.

Federal Aviation Administration. (2015). *Advisory circular* (No. 120/111). Washington, DC: Author.

Flach, J. M., Feufel, M. A., Reynolds, P. L., Parker, S. H., & Kellogg, K. M. (2017). Decisionmaking in practice: The dynamics of muddling through. *Applied Ergonomics*, *63*, 133–141.

Flach, J. M., Mulder, M., & Van Paassen, M. M. (2004). The concept of the situation in psychology. In S. Banbury & S. Tremblay (Eds.), *A cognitive approach to situation awareness* (pp. 42–60). Oxford, UK: Ashgate.

Foster, M. I., & Keane, M. T. (2015). Why some surprises are more surprising than others: Surprise as a metacognitive sense of explanatory difficulty. *Cognitive Psychology*, *81*, 74–116.

Gillan, C. (2003). Analysis of multicrew decision making from a cognitive perspective. In *Proceedings of the 12th International Symposium on Aviation Psychology* (pp. 427–432). Mahwah, NJ: Lawrence Erlbaum.

Gioia, D. A., & Poole, P. P. (1984). Scripts in organizational behavior. *Academy of Management Review*, *9*, 449–459.

Globisch, J., Hamm, A. O., Esteves, F., & Öhman, A. (1999). Fear appears fast: Temporal course of startle reflex potentiation in animal fearful subjects. *Psychophysiology*, *36*, 66–75.

Hagemann, T., Levenson, R. W., & Gross, J. J. (2006). Expressive suppression during an acoustic startle. *Psychophysiology*, *43*, 104–112.

Haslbeck, A., & Hoermann, H.-J. (2016). Flying the needles: Flight deck automation erodes fine-motor flying skills among airline pilots. *Human Factors*, *58*, 533–545.

Hollnagel, E., & Woods, D. D. (2005). *Joint cognitive systems: Foundations of cognitive systems engineering*. Boca Raton, FL: CRC Press.

International Civil Aviation Organisation. (2013). *Manual of evidence-based training* (Nr. 9995). Montreal, Canada: Author.

JIAAC-9-058-2005 [English translation of Venezuelan accident report]. (2005). Caracas, Venezuela: Civil Aviation Accident Investigation Board.

Kahneman, D. (2003). A perspective on judgment and choice: Mapping bounded rationality. *American Psychologist*, *58*, 697–720.

Klein, G. A. (1993). A recognition-primed decision (RPD) model of rapid decision making. In G. A. Klein, J. Orasanu, R. Calderwood, & C. E. Zsambok (Eds.), *Decision making in action: Models and methods* (pp. 138–147). Norwood, NJ: Ablex.

Klein, G. A., Phillips, J. K., Rall, E. L., & Peluso, D. A. (2007). A data-frame theory of sensemaking. In *Expertise out of context: Proceedings of the Sixth International Conference on Naturalistic Decision Making* (pp. 113–155). London, UK: Psychology Press.

Kochan, J. A. (2005). *The role of domain expertise and judgment in dealing with unexpected events* (PhD thesis). University of Central Florida, Orlando.

Kochan, J. A., Breiter, E. G., & Jentsch, F. (2004). Surprise and unexpectedness in flying: Database reviews and analyses. In *Proceedings of the Human Factors and Ergonomics Society 48th Annual Meeting* (pp. 335–339). Santa Monica, CA: Human Factors and Ergonomics Society.

Kowalski-Trakofler, K. M., Vaught, C., & Scharf, T. (2003). Judgment and decision making under stress: an overview for emergency managers. *International Journal of Emergency Management*, *1*(3), 278-289.

Landman, A., Groen, E. L., Van Paassen, M. M., Bronkhorst, A., & Mulder, M. (in press). The influence of surprise on upset recovery performance in airline pilots. *International Journal of Aerospace Psychology*.

Lanir, Z. (1986). *Fundamental surprise*. Eugene, OR: Decision Research.

Lazarus, R. S., & Folkman, S. (1984). *Stress, appraisal, and coping*. New York, NY: Springer.

Martin, W. L., Murray, P. S., Bates, P. R., & Lee, P. S. Y. (2015). Fear-potentiated startle: A review from an aviation perspective. *International Journal of Aviation Psychology*, *25*, 97–107.

Martin, W. L., Murray, P. S., Bates, P. R., & Lee, P. S. Y. (2016). A flight simulator study of the impairment effects of startle on pilots during unexpected critical events. *Aviation Psychology and Applied Human Factors*, *6*, 24–32.

McKinney, E. H., & Davis, K. J. (2003). Effects of deliberate practice on crisis decision performance. *Human Factors*, *45*, 436–444.

Meyer, W.-U., Reisenzein, R., & Schützwohl, A. (1997). Toward a process analysis of emotions: The case of surprise. *Motivation and Emotion*, *21*, 251–274.

Militello, L. G., & Hutton, R. J. B. (1998). Applied cognitive task analysis (ACTA): A practitioner's toolkit for understanding cognitive task demands. *Ergonomics*, *41*, 1618–1641.

Ministry of Civil Aviation. (2004). *Final report of the accident investigation, Flash Airlines Flight 604*. Cairo, Egypt: Author.

Muhren, W. J., & Van de Walle, B. (2010). A call for sensemaking support systems in crisis management. In R. Babuška & F. C. A. Groen (Eds.), *Interactive collaborative information systems* (pp. 425–452). Berlin, Germany: Springer.

Mumaw, R. J., Groen, E., Fucke, L., Anderson, R., Bos, J., & Houben, M. (2016). A new tool for analyzing the potential influence of vestibular illusions. *ISASI Forum*, *49*(1), 6–12.

Neisser, U. (1976). *Cognition and reality: Principles and implications of cognitive psychology*. New York, NY: W. H. Freeman.

Nieuwenhuys, A., & Oudejans, R. R. D. (2012). Anxiety and perceptual-motor performance: Toward an integrated model of concepts, mechanisms, and processes. *Psychological Research*, *76*, 747–759.

Parasuraman, R., & Riley, V. (1997). Humans and automation: Use, misuse, disuse, abuse. *Human Factors*, *39*, 230–253.

Phillips, J. K., Klein, G., & Sieck, W. R. (2004). Expertise in judgment and decision making: A case for training intuitive decision skills. In D. J. Koehler & N. Harvey (Eds.), *Blackwell handbook of judgment and decision making* (pp. 297–315). Oxford, UK: Blackwell.

Piaget, J. (1976). Piaget's theory. In B. Inhelder, H. H. Chipman, & G. Zwingmann (Eds.), *Piaget and his school* (pp. 11–23). New York, NY: Springer.

Previc, F. H., & Ercoline, W. R. (1999). The "outside-in" attitude display concept revisited. *International Journal of Aviation Psychology*, *9*, 377–401.

Previc, F. H., & Ercoline, W. R. (2001). Trends in spatial disorientation research. *Aviation, Space, and Environmental Medicine*, *72*, 1048–1050.

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

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*, *3*, 257–266.

Rivera, J., Talone, A. B., Boesser, C. T., Jentsch, F., & Yeh, M. (2014). Startle and surprise on the flight deck similarities, differences, and prevalence. In *Proceedings of the Human Factors and Ergonomics Society 58th Annual Meeting* (pp. 1047–1051). Santa Monica, CA: Human Factors and Ergonomics Society.

Roberts, N. A., Beer, J. S., Werner, K. H., Scabini, D., Levens, S. M., Knight, R. T., & Levenson, R. W. (2004). The impact of orbital prefrontal cortex damage on emotional activation to unanticipated and anticipated acoustic startle stimuli. *Cognitive, Affective, & Behavioral Neuroscience*, *4*, 307–316.

Sarter, N. B., Woods, D. D., & Billings, C. E. (1997). Automation surprises. *Handbook of Human Factors and Ergonomics*, *2*, 1926–1943.

Schmidt, R. A. (1975). A schema theory of discrete motor skill learning. *Psychological Review*, *82*, 225.

Schroeder, J. A., Bürki-Cohen, J., Shikany, D. A., Gingras, D. R., & Desrochers, P. (2014). An evaluation of several stall models for commercial transport training. In *AIAA Modeling and Simulation Technologies Conference* (Paper No. AIAA 2014-1002). Reston, VA: AIAA.

Schützwohl, A. (1998). Surprise and schema strength. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *24*, 1182–1199.

Senders, J. W. (1964). The human operator as a monitor and controller of multidegree of freedom systems. *IEEE Transactions on Human Factors in Electronics*, *1*, 2–5.

Shappell, S., Detwiler, C., Holcomb, K., Hackworth, C., Boquet, A., & Wiegmann, D. A. (2007). Human error and commercial aviation accidents: An analysis using the human factors analysis and classification system. *Human Factors*, *49*, 227–242.

Sheridan, T. B. (1981). Understanding human error and aiding human diagnostic behaviour in nuclear power plants. In J.

Rasmussen & W. B. Rouse (Eds.), *Human detection and diagnosis of system failures* (pp. 19–35). New York, NY: Plenum.

Sherry, L., Feary, M., Polson, P., & Palmer, E. (2001, March). *What's it doing now? Taking the covers off autopilot behavior*. Paper presented at the 11th International Symposium on Aviation Psychology, Columbus, OH.

Stanton, N. A., Salmon, P. M., Walker, G. H., & Jenkins, D. (2009). Genotype and phenotype schemata and their role in distributed situation awareness in collaborative systems. *Theoretical Issues in Ergonomics Science*, *10*, 43–68.

Summers, M. M. (2007). *Scenario-based training in technically advanced aircraft as a method to improve risk management*. Retrieved from: http://www.faa.gov/education_research/training/fits/research/media/SBT_for_RM.pdf

Teigen, K. H., & Keren, G. (2003). Surprises: Low probabilities or high contrasts? *Cognition*, *87*, 55–71.

Thackray, R. I. (1988). *Performance recovery following startle: A laboratory approach to the study of behavioral response to sudden aircraft emergencies* (No. DOT/FAA-AM-88/4). Washington, DC: Federal Aviation Administration, Office of Aviation Medicine.

Van Merriënboer, J. J. G. (1997). *Training complex cognitive skills: A four-component instructional design model for technical training*. Englewood Cliffs, NJ: Educational Technology.

Weick, K. E. (1995). *Sensemaking in organizations*. Thousand Oaks, CA: Sage.

Wickens, C. D., Hooey, B. L., Gore, B. F., Sebok, A., & Koenicke, C. S. (2009). Identifying black swans in NextGen: Predicting human performance in off-nominal conditions. *Human Factors*, *51*, 638–651.

Wickens, C. D., Stokes, A., Barnett, B., & Hyman, F. (1993). The effects of stress on pilot judgment in a midis simulator. In A. Maule & O. Svenson (Eds.), *Time pressure and stress in human judgment and decision making* (pp. 271–292). New York, NY: Springer.

Young, M. S., & Stanton, N. A. (2002). Malleable attentional resources theory: A new explanation for the effects of mental underload on performance. *Human Factors*, *44*, 365–375.

Annemarie Landman received her MSc in movement sciences from the VU University Amsterdam in 2011. She is currently a PhD candidate in the Control and Simulation Division, Aerospace Engineering, at the Delft University of Technology.

Eric L. Groen earned his PhD in biology in 1997 from Utrecht University. He is currently is a senior scientist in the Multimodal Interface Research Group, Human Interfaces Department, at TNO Human Factors.

M. M. (René) van Paassen is associate professor in the Control and Simulation Division, Aerospace Engineering, at the Delft University of Technology, where he received his PhD in aerospace engineering in 1994.

Adelbert W. Bronkhorst received his PhD in audiology in 1990 from the medical faculty of the VU University Amsterdam. He is currently a principal scientist at TNO Human Factors, and professor by special appointment of Applied Cognitive Psychology at the VU University Amsterdam.

Max Mulder is professor in the Control and Simulation Division, Aerospace Engineering, at the Delft University of Technology, where he received his PhD in aerospace engineering in 1999.

Date received: November 14, 2016
Date accepted: July 7, 2017