A New Tool for Analyzing the Potential Influence of Vestibular Illusions

by

Randall J. Mumaw1 (Associate Technical Fellow, Human Factors, Boeing), Eric Groen (Senior Scientist, TNO), Lars FucKee (Lead Engineer, Boeing), Richard Anderson (Senior Accident Investigator, Boeing), Jelte Bos (Senior Scientist, TNO; Professor, VU University) & Mark Houben (Research Scientist, TNO)

Dr. Randy Mumaw, a cognitive scientist, has worked 15 years in aviation safety and accident investigation at Boeing. Recently, he has focused on loss of control events, especially those involving loss of situation awareness regarding airplane state. Dr. Eric Groen, aerospace physiologist, Prof. Dr. Jelte Bos, physicist, and Dr. Mark Houben, biomedical engineer, work as scientists at TNO and share research interest in the areas of pilot spatial disorientation, upset recovery, motion sickness, and flight simulator technology. Lars FucKee is an R&D Lead with Boeing (Madrid) working in the areas of system and flight deck operations safety. Richard Anderson has been a Boeing accident investigator since 1997.

Role of Spatial Disorientation in Aviation Accidents

In January of 2004, a 737-300 crashed minutes after take-off from Sharm el-Sheikh, Egypt. The departure was on a dark night, over the Red Sea, and there were few, if any, visible cultural landmarks that could be used to orient to the horizon. The Captain (the Pilot Flying, PF) had initiated a long left climbing turn, but partway through that turn the airplane had actually made a slow transition from a left bank to a right bank (20° and increasing slowly). The First Officer (FO) informed the Captain that they were turning right in this exchange:

FO: Turning right, sir
Captain: What?
FO: Aircraft turning right
Captain: Turning right? How turning right?

At this point, the Captain was making control inputs on the wheel to roll further to the right, and continued doing so. The airplane eventually rolled to about 110° to the right before substantial control inputs in the opposite direction were made, which was too late to avoid the crash into the Red Sea. During this event, the Captain seemed unable to determine which way to roll the airplane to restore it to wings-level—at one point trying to engage the autopilot to get assistance in recovering from the upset. The investigation reached the conclusion that the pilot was spatially disoriented2.

This event and the findings of the investigation were surprising for many safety experts studying commercial jet transports: spatial disorientation (SD) was not considered a significant hazard in airline operations. SD was known to be a risk in high-speed, highly maneuverable military jets, but in the relatively stable world of commercial jet transports, SD was not considered a threat. At this point, the findings from Flash Air seemed to be a “one-off” event. Unfortunately, two more similar 737 accidents

1 Dr. Randy Mumaw is now at NASA Ames
2 Of course, virtually every accident is the result of a chain of events and failures, and, in Boeing’s analyses, no accident was judged to be caused solely by a pilot’s spatial disorientation.
happened in 2007 (Adam Air at Sulawesi; Kenya Airways in Douala, Cameroon). In each case, the PF made control inputs away from wings level, resulting in a loss of control (LOC) and fatal crash.

In 2008, Boeing took a closer look at the influence of SD in commercial transport accidents. We established a clear definition of SD for this context and searched for accidents and major incidents that fit that definition. In some cases, accident reports, especially reports from before 1990, did not provide sufficient detail to place the event conclusively in the SD category. However, this extensive search identified 16 SD-related accidents and one major incident in the period of 1991-2007; roughly one event/year (see Figure 1). Also, 2008 produced another 737 accident (Aeroflot Nord at Perm, Russia) that had the same signature of the PF rolling away from wings level. Further, since 2008, other accidents and incidents around the world have been linked to SD—e.g., Afriqiyah A330 at Tripoli, Libya in May, 2010 and the Scat CRJ-200 near Almaty, Kazakhstan in January, 2013.

One important finding of this review of accidents and incidents was the identification of two very different SD phenomena that were contributing to these accidents:

- **Sub-threshold roll** – In these cases, the pilot had an understanding or expectation of the airplane’s orientation; typically, it was wings-level. Then, for various reasons, the airplane rolled away from that orientation at a rate less than 5°/second. Roll rates this slow fall below the vestibular system’s ability to detect; hence the name sub-threshold. Further, load factors during the roll were less than 1.2 G, indicating both that pilots were not loading up the airplane during an intentional turn and that their somatosensory (or “seat-of-the-pants”) input would not have been significantly different from level flight. Pilots were unaware of the change in orientation and then suddenly found their airplane banked at 35° or beyond. In this situation, these pilots were apparently confused about which direction to roll back to wings-level, and they rolled the airplane in the wrong direction. These inappropriate pilot inputs were key because the airplane was not initially in an unrecoverable attitude. Note that it is possible that a post-roll illusion could also have influenced the progressively inappropriate control inputs [1,2].

- **Somatogravic illusion** – This illusion is quite different from the first. Sub-threshold roll relies on the vestibular system failing to detect a change to the airplane. The somatogravic illusion, on the other hand, is the result of a misinterpretation of a very noticeable sensation related to linear acceleration. This illusion typically occurs on a go-around when the airplane transitions from a slowing down to a rapid acceleration and pitch-up. The vestibular system cannot distinguish between an inertial acceleration and a component of gravity, and the rapid acceleration can be misinterpreted as a further pitching up moment. Again, IMC and/or darkness contribute by removing valid visual inputs. As the airplane begins the go-around, the pilot perceives that the airplane is pitching up considerably and starts to push the nose downward to compensate. This can result in an actual nose-down attitude and descent into the ground [3,4].
Figure 1. Identified SD events

Commercial Aviation Safety Team

These insights from the 2008 work led Boeing to engage the larger aviation safety community. In 2009, we approached the Commercial Aviation Safety Team\(^3\) (CAST) to share our findings on SD events. CAST takes the role of bringing together government and industry to analyze safety issues, generate potential solutions, assess the feasibility of those solutions, and adopt the solutions that are both effective and feasible. These solutions become the official CAST Safety Enhancements, which are then implemented by CAST members.

CAST agreed to study this issue beginning in 2010, and combined it with another group of LOC events tied to energy state. The larger theme for CAST was the pilot’s loss of awareness regarding airplane state: loss of attitude awareness (SD) and loss of energy state awareness. More generally, it was called airplane state awareness. The group given this charge included members from Boeing, Airbus, Embraer, Bombardier, Honeywell, Rockwell Collins, MITRE, Airlines for America, the Regional Airline Association, the National Air Carrier Association, FAA, NASA, and pilots’ unions (ALPA and APA).

This group conducted a detailed analysis of the following SD-related events\(^4\):
- Gulf Air, Airbus A320, August 23, 2000, Bahrain.
- Flash Air, Boeing 737-300, January 3, 2004, Sharm el-Sheik, Egypt.
- Armavia Airlines, Airbus A320, May 3, 2006, Sochi, Russia.
- Aeroflot-Nord, Boeing 737-500, July 14, 2008, Perm, Russia.

\(^3\) The Commercial Aviation Safety Team (http://www.cast-safety.org/) is a voluntary collaboration between the US government and industry founded in 1998. Its goal is to reduce fatality risk from commercial aviation accidents 50% in the US by 2025.

\(^4\) This was meant to be a representative set, not an exhaustive set of SD-related events.
The CAST analysis identified a number of other issues that contributed to many of these events. The most relevant of these for the SD events were:

- **Lack of external visual references** – In these SD-related events, due to darkness or IMC, flight crew members had no clear view of the horizon through the flight deck windows and, therefore, lacked normal orientation and self-motion cues—such as, perspective, depression angle, optical flow, and motion parallax. A visible horizon can establish “visual dominance,” a well-known perceptual phenomenon in which the visual input can overcome a vestibular illusion.

- **Crew distraction** – While some form of distraction occurs on virtually all flights, it is successfully managed by flight crews in the vast majority of cases. Flight crews are trained to eliminate and/or manage distractions. In the events we analyzed, the basic task of aviating was neglected, attention was not given to critical alerts or displays, or decision-making was hindered. A major component of this failure of attention was channelized attention, a phenomenon in which a pilot becomes completely focused on some task or issue and is unable to shift attention to other important tasks; in this case, aviating.

- **Crew resource management** – CRM is a broad term and covers many aspects of crew performance. Most relevant here was the inability of the flight crew member who was NOT disoriented to intervene or take control from the PF. Authority gradient was at play in several of these events, as well as poor understanding or execution of managing an incapacitated pilot (i.e., the disoriented pilot). The one event that was not an accident was a case in which the Pilot Monitoring (PM) grabbed the wheel and column and fought hard (against the PF) to bring the airplane out of the dive (at about 320 feet agl).

The CAST work led to a number of proposed safety enhancements tied to changes to airplane design, operational procedures, and pilot training. It also called out specific needs in the areas of aviation safety R&D and safety data management. For the SD events, the safety enhancements ideally address both the PF’s inappropriate control inputs and the PM’s reluctance or inability to intervene when the PF is incapacitated by SD. One specific safety enhancement that Boeing is pursuing is a roll arrow that provides alerting and roll guidance to the pilot when bank angle exceeds 45°. We believe this enhancement addresses both guidance for control inputs and more effective intervention.

**Accident Investigation and Analysis**

While the CAST work identified a broad set of safety enhancements, it failed to touch on accident investigation. In large part, the investigation agencies that try to make sense of pilot actions have no capability to assess the potential for SD. A few of the events mentioned above were subjected to this type of analysis because the investigating agency hired outside experts to apply their perceptual models to the flight data. Other investigation reports have only speculated about the possible influence of SD on the pilot’s actions and have provided no analysis. Boeing saw the need for a valid, accessible tool that allows investigators to look at flight data and determine if SD may have contributed to pilot control inputs. We turned to a group with expertise in modeling perceptual systems and illusions: the Netherlands Organization for Applied Scientific Research, or TNO.

TNO, with a long tradition in vestibular research, developed a general perception model to predict and analyze human motion perception in environments such as airplanes, cars, ships, and also moving-base simulators [5,6]. Their state-of-the-art model consists of mathematical representations of the sensory systems involved in motion perception (i.e., visual and vestibular system), as well as their neural interaction. The model takes in time histories of self-motion and orientation, and predicts how they are being perceived. With respect to spatial orientation in aviation, the dominant issue is the perceived self-
orientation relative to gravity. Essentially, the model takes the pilot's point of view—i.e., the orientation of perceived gravity with respect to the self. Moreover, it is essential to understand that the human sensors are not perfect, and the central nervous system (CNS) does not reckon all laws of physics, such as Newton's second law, and the differential relationship between position, velocity and acceleration. This allows for perceptual ambiguities that basically determine spatial disorientation. The TNO model has been successfully applied to predict motion sickness incidence and to evaluate motion cueing in flight simulators [7,8].

We used the TNO perception model as the starting point for the collaborative development of a stand-alone software tool to support the analysis of SD events from flight data. The basic idea is that comparison between recordings of aircraft motion and attitude (model input), and the way this is being perceived according to the model (model output) should help identify the phases of flight which are prone to induce spatial disorientation. In its current state, the interpretation of the model output in terms of SD requires a subject matter expert. The objective of the project was to make the model applicable and accessible for accident investigators by 1) implementing a module that automatically recognizes SD events in the data, also referred to as detection and identification; and 2) adding a user-friendly interface.

Basic Perception Model

The perception model consists of the relevant sensory transfer functions and the visual-vestibular interactions that play a role in human spatial orientation (Figure 2). In this model, the organs of balance within our inner ears sensing physical motions are divided into otoliths (OTO) and semicircular canals (SCC). The otoliths typically respond to specific force ($f$) and code for linear acceleration, and the semicircular canals respond to angular accelerations of the head, and their output codes for angular velocity ($\omega$). Within the visual system the optic flow (FLW) in the retinal image typically carries information on head velocity. In addition, horizontal and vertical elements in the retinal image provide a visual frame (F), and together with polarity (P) cues about what is “up” and what is “down” these determine the visual orientation of the head with respect to Earth ($p$). Still, these vestibular and visual cues do not fully account for human orientation. Human subjects typically underestimate their self-tilt, a phenomenon called the A- or Aubert effect [9]. To explain this bias towards the longitudinal body axis, which is considered a somatosensory phenomenon, Mittelstaedt [10] assumed a body-fixed “idiotropic vector” (IV and $i$), which is added vectorially to the vestibular vertical [11,12,13].

The neural integration of these sensory signals has been implemented as follows. As stated, the otoliths respond to specific force ($f$), i.e., the vector sum of the free-fall acceleration determined by gravity ($g$), and inertial accelerations determined by linear motion ($a$), hence $f = g + a$. Although of different origin, accelerations due to gravity and inertia are inherently indistinguishable (Einstein's equivalence principle). According to Mayne [14], our brain seems capable of making the distinction by a neural process which behaves like a low-pass filter (LP). Assuming that the brain “knows” that gravity is constant, and accelerations due to head motion are relatively variable, a low-pass filter adequately separates both components from the otolith output ($f$). Additional information on angular motion of the head, coming from the semicircular canals and visual flow is included in the model, not only to estimate subjective rotation (SR), but also to apply the required rotations ($R$ and $R^{-1}$) for estimating the specific force components relative to Earth. This is required because the specific force is sensed in a head-fixed frame of reference, while gravity is constant in an Earth-fixed frame of reference. Using a weighted vector addition, the resulting internal estimate of gravity ($g$) is combined with the visual and idiotropic vectors to determine the subjective vertical, or SV [5,12].
In order to make the model applicable as a stand-alone tool for the detection of SD illusions from flight
data, three enhancements were needed: 1) a “detection and identification module” to automatically
recognize SD; 2) visualization of the model output; 3) a user interface to allow interaction with the input
and output. These enhancements are discussed in the next sections.

SD Categories

Based on the results of the aforementioned Boeing study, the current project focused on automatic
detection of vestibular illusions, in particular sub-threshold roll motion and the somatogravic illusion.
More complicated vestibular illusions (e.g., the Coriolis illusion), as well as visual illusions tied to motion
and orientation (e.g., “black hole,” vection illusion), fell outside the scope of this project, as these require
information that is not available from the flight data recorder, such as the pilot’s head movements
(Coriolis) and visual inputs.

Sub-threshold roll motion is related to the functioning of the semicircular canals, and falls in the category
of “somatogyral illusions.” This involves misperceptions of angular motion in general, not only
undetected motions that remain below the perceptual threshold, but also false (after-) sensations of motion
when the real (aircraft) motion has stopped. Examples of the latter are the “post-roll illusion” [1] and the
“graveyard spin” [3]. Figure 3 shows the response of the semicircular canals to a step input of roll motion
that is sustained for several seconds before it abruptly ends again. Since the semicircular canals behave
like a high-pass filter, they only respond to changes in angular motion, but not to constant rates. Hence, as
the figure illustrates, the pilot accurately perceives the onset of roll motion, but this sensation gradually
fades out as the motion continues at a constant rate. Eventually, the sensation may become sub-threshold
even though the aircraft is still turning at a rate that is above the perceptual threshold. When the turn is
stopped, however, an after-sensation appears in the opposite direction of the original aircraft motion. This
illusory after-sensation may prompt the pilot to make inappropriate control inputs. In the case of roll
motion, it has been shown that the post-roll effect induces pilots to overshoot the bank angle [2]. Hence,
this vestibular effect also contributes to the crew’s confusion about the direction in which an aircraft is
banking.
Figure 3. Model-predicted response of the semicircular canals (red line) to a roll input (blue line) that is maintained for several seconds. The shaded area indicates the threshold that must be exceeded before angular motion is being perceived.

The “somatogravic illusion” is related to the functioning of the otoliths, and the perceptual ambiguity of the specific force. The illusion been studied during sustained centrifugation, where the constant tilt of the specific force is gradually being perceived as “vertical” [15,16]. For example, a subject who is seated upright and facing the center of the centrifuge soon feels him- or herself tilted backwards, similar to the effect that a pilot may experience during a go-around maneuver. Figure 5 illustrates the model output in such a (simplified) situation.

Figure 4. Model response to (simplified) sustained longitudinal acceleration. The tilted specific force vector (blue line) gradually induces a sensation of self-tilt (red line), i.e. the somatogravic illusion. The shaded area indicates the perceptual threshold that must be exceeded before the illusory tilt is being perceived.

SD Detection and Identification

The SD Detection and Identification module includes logic that discriminates between the various SD illusions (Figure 5). The module first computes the mismatch between the perceived attitude (the subjective vertical) and the true orientation of the aircraft relative to Earth, as well as the mismatch between the perceived (subjective rotation) and angular rates. When one of these mismatches exceeds a critical value, another logic is applied to identify whether a misperception of attitude results from the somatogravic illusion, or from a cumulative effect of misperceived angular motion. Further, computations are being made to differentiate whether a somatogyral illusion occurs during aircraft motion (when the perceived angular rate drops below a threshold value), or after aircraft motion (the post-roll effect, when the after-sensation exceeds the same threshold value). Looking at Figure 3 this means that, although the perceived angular rate starts fading out quite soon during the roll motion, it is only identified as SD when it drops below the threshold. Similarly, at the stop of the airplane roll, the illusory after-effect is only designated a post-roll illusion as long as the model-output exceeds the threshold. In addition, aircraft
motions that do not exceed the threshold value at all are being identified as sub-threshold motion. The critical values used for identification are based on TNO research as well as the open literature, and can be adjusted to optimize the model’s signal-to-noise ratio.

Figure 5. Flow of automatic detection and identification of SD events from the Flight Data Recorder.

The TNO Software Tool

The software application takes in time histories of flight data (e.g., from a flight data recorder), selected in the “Input and Settings” tab (see Figure 6) that also allows the user to set critical values. The model then computes the perceived motion, and labels the SD categories, that are shown on the “Results” tab together. The model output can be saved to file on the “Outputs” tab. Figure 6 shows a screenshot of the Results tab of the Graphical User Interface (GUI). The plots on the left part of the window show time histories of rotation and attitude in three cardinal directions (x-axis in red = roll, or surge; y-axis in green = pitch, or sway; z-axis in blue = yaw, or heave). The solid lines reflect actual aircraft motion (model input), and the dotted lines reflect the perceived motion (model output). The area between aircraft and perceived motion is shaded to indicate the mismatch that is input to the SD Detection. The upper right part of the window shows the criteria settings. The bottom tracks show various SD labels that have been identified by the model: “attitude” (mismatch in perceived attitude), “grav” (somatogravic illusion), “gyral” (somatogyral illusion), “sub” (subthreshold angular motion). The bottom right of the window contains an animation of aircraft attitude (solid aircraft icon) and perceived attitude (transparent aircraft icon); the view can be toggled between aft and side view. When there is a misperception of attitude, these two deviate. The animation can be controlled with play, pause and stop buttons. The vertical black line in the time series at the left part of the GUI shows the current time of the animation.

The example in Figure 6 corresponds to a coordinated turn to the right at 30° angle of bank. Around t=7 s the perceived roll rate (dotted red line in the upper time history) starts to wane due to the dynamics of the semicircular canals. This results in a mismatch between actual and perceived bank angle (more specifically, an underestimation of bank angle), as indicated by the “attitude” track at the bottom. Between t=8.5 s and 10 s the label “gyral” is also activated, indicating that the perceived roll rate has dropped below the threshold (3°/s) while the aircraft is still rolling at a rate above this threshold (hence, the somatogyral illusion). Being a coordinated turn, the aircraft’s specific force banks with the airplane up to about 30°, and hence remains aligned with the body axis throughout the maneuver (solid red line in the bottom time history). The model output (dotted red line in the same plot) shows that the pilot briefly perceives banking to the right, but then the low-pass filter that distinguishes between inertial and gravitational acceleration causes the specific force to be perceived as vertical. Eventually this results in the feeling of “level flight” while in reality the aircraft is banked relative to the Earth. Finally, after 10 s, the subthreshold label is activated because the actual roll rate has dropped below the perception threshold.
Figure 6. Screenshot of the model output during analysis of a coordinated turn. The bottom tracks “attitude”, “gyral”, and “sub” indicate that the maneuver induces a misperception of attitude, a somatogyral illusion, and also contains an episode of sub-threshold motion.

Figure 7 shows another screenshot produced from data of a takeoff flown in a flight simulator. Due to the forward acceleration of the aircraft (solid red line in upper plot), a false perception of pitching up arises (dotted green line in bottom plot) while the aircraft stays level (i.e., zero pitch). From about t=13 s the mismatch between perceived and actual pitch is large enough (criterion set at 8°) to be identified as the somatogravic illusion, as well as misperceived attitude. These two examples show that both the somatogyral and the somatogravic illusions can lead to misperceived attitude. In the case of the somatogyral illusion this is due to the time integral of misperceived angular motion.
A Case Study

This analysis, driven by the TNO model, shows that the vestibular system can often be fooled by airplane flight, and we know that virtually every pilot has experienced at least momentary confusion about orientation. However, we also know that pilots are rarely so disoriented that they make inappropriate flight control inputs because, typically, the visual information environment is rich enough and familiar enough to trump the vestibular inputs. Accidents and incidents (from the Boeing analysis), however, demonstrate that there are rare cases in which a degraded visual environment can lead to a greater susceptibility to SD.

The role of this tool in accident investigation is to help us understand why inappropriate control inputs—rolling away from wings level or pushing the nose down at a low altitude—were made. Figure 8 shows a brief illustration (part of a larger case study), using data from a 737-300 accident, highlighting this capability. The airplane took off on a dark night over water, so there were few visible cultural landmarks to support orientation. The PF had initiated a long left climbing turn, but partway through that turn the airplane made a slow transition from a left bank to a right bank. This period of transition from about $20^\circ$ left bank to $20^\circ$ right bank took about 70 seconds, and the airplane was pitching up and slowing down 15-20 kts during this period. The model analysis indicates that, during this period, there was no vestibular feedback on the airplane’s orientation and motion, which, without strong visual input on orientation, would have led to the PF’s confusion about the airplane’s orientation. The SD track “sub” shows that the transition from banked left to banked right was almost completely sub-threshold, meaning that the pilots did not feel the airplane’s roll motion. Second, similar to the example in Figure 7, the specific force
vector during this coordinated flight remained aligned with the airplane’s z-axis, which from a vestibular perspective is undistinguishable from “wings level.” Hence, there was no meaningful vestibular information about the airplane’s change in attitude, which explains the SD track “attitude” in Figure 8.

Looking in more detail at the figure, the “sub” track is interrupted at places where the model output for roll rate temporarily exceeded the threshold (refer to Figures 3 and 4 to see how the internal threshold determines whether or not the model output activates an SD label). The interruptions of the “attitude” track correspond to periods where the mismatch in perceived attitude was smaller than the criterion of 8°. Note that according to the blue shaded area in the bottom time history, there was little or no vestibular feedback about the change in heading (perceived yaw angle remained around 0°), but since heading does not affect the orientation relative to gravity, it is not included in the determination of SD.

During this 70-second period, also, the PF became confused and distracted by some unexpected behavior from the autoflight system. This distraction probably reduced his awareness of his slow, perhaps inadvertent, control inputs to roll right. When the PF was told that he was turning right, he became confused about his orientation and how to return to wings level. Subsequent roll inputs were strongly to the right, leading to a loss of control and fatal crash.

This short illustration of the model’s analysis capability shows how it can be combined with the cockpit voice recorder (CVR), flight data recorder (FDR), and environmental data inputs to create a more complete picture of the pilot’s understanding of the state of the airplane. This data integration and analysis is at the heart of accident investigation, and allows us to explain flight control inputs more completely.

**Figure 8.** Model analysis of a B737 LOC accident showing 140 s of actual and perceived angular motion (upper time history), and actual and perceived roll and pitch attitude (lower time history). The SD tracks indicate issues with perceived attitude as well as sub-threshold angular motion through the larger part of recorded flight.
Conclusions

Any accident investigation that implicates human performance issues ("pilot error") needs to consider performance in context, and, in some cases, that context should include the sensory systems’ inputs to the pilot’s overall situation awareness. The long history of aviation safety has shown that SD occurs and can have fatal consequences. The TNO tool offers a method to more completely examine that context. It shows what the pilot’s vestibular system was telling the pilot about his/her orientation and motion. Certainly, this input is only part of the whole picture, but when there is a degraded visual environment, we have seen that the vestibular inputs can drive the pilot’s actions into a larger upset and loss of control. In some cases, the reality generated by these false perceptions can be strong and enduring and, unless there is a rapid and forceful response from the PM, can lead to a crash.

These SD events will probably continue to occur in the short-term. The recommendations from CAST advocate for changes to airplane design, operations, and flight crew training to address some of the factors that contribute to turning SD into accidents. We hope that, eventually, these changes will significantly reduce the risk of SD turning into a LOC event. In the meantime, the tool developed by Boeing and TNO can become an essential element of the accident/incident investigation process.

References