2015

Understanding Automation Surprise: Analysis of ASRS Reports

Julia Trippe

Robert Mauro

UNDERSTANDING AUTOMATION SURPRISE: ANALYSIS OF ASRS REPORTS

Julia Trippe & Robert Mauro
Decision Research
Eugene, Oregon

Pilots are frequently surprised by aircraft automation. These include cases in which the automation: 1) produces alerts to anomalies, 2) commands unexpected control manipulations (that may result in flight path deviations), or 3) simply disconnects. Aviation Safety Reporting System (ASRS) reports in which pilots indicated that automation produced unexpected actions were analyzed. Three general conclusions were drawn. First, many factors precipitate automation surprises. These include problems in: the auto-flight system and associated displays and interfaces, other aircraft sensors and systems, and interactions with weather and ATC. Second, inappropriate pilot actions are involved in a large proportion of these events. Third, recovery need not require reversion to manual control. There is no single general intervention that can prevent automation surprise or completely mitigate its effects. However, several different tacks (including improved training, displays, and coordination with ATC) taken together may be effective.

The capabilities of automated flight systems increased rapidly following the introduction of the electronic autopilot in the 1940's. In normal operations, the automated flight system of the modern airliner can now control nearly all functions required for flight. The effect of increased automation has been largely positive, greatly reducing errors due to pilot fatigue and allowing consistent precise navigation and performance. However, automation has given rise to new problems caused by faulty interactions between the pilot and the auto-flight system (AFS). This class of problems has been variously termed lack of mode awareness (Javaux and De Keyser, 1998), mode confusion (Degani, Shafto, & Kirlik, 1999), and automation surprise (Winter & Curry 1989; Woods, Sarter, and Billings, 1997, Burki-Cohen, 2010). In these cases, the flight crew expects the automation to command one behavior and is surprised when it commands another. When they do not jeopardize flight safety, automation surprises are a nuisance. But when the automation commands an aircraft trajectory that violates airspace or operational limitations, automation surprise becomes a critical problem (Reveley et al, 2010).

Automation surprise may result from undetected failures in aircraft sensors or other systems. Automation surprise also may result from pilots having an inadequate or mistaken "mental model" of the machine's behavior in the operational environment (Sarter and Woods, 1995). In addition, automation surprise may result from a problematic interface that does not provide adequate information about the status of the machine (Feary et al 1998; Norman, 1990, Degani, Shafto, and Kirlik, 1999).

Pilot lore is replete with complaints of flight management systems misbehaving. Flight management computers can appear to be pernicious allies that on occasion unilaterally decide to "drop" fixes, void altitude restrictions, or change modes of operation. For the most part the result of these events are relatively benign. No metal is bent; no one is injured; no lives are lost. But this is not always the case. Unexpected behaviors of the auto-flight system have been implicated in a number of recent fatal accidents (Sherry & Mauro, 2014). Furthermore, these events increase pilot workload, setting the stage for other errors. They create inefficiencies for the aircraft directly involved and may disrupt the flow of air traffic as controllers vector other aircraft to accommodate the aircraft whose crew is dealing with the unexpected behavior. To prevent or mitigate the effects of these "automation surprises" one must first understand why they occur.

People are surprised when they expect one event but another occurs. So, to understand automation surprise, one must ask why the behavior of the auto-flight system was not expected. Based on their training and experience, pilots build an understanding (a "mental model") of how their automation functions. Selected information about the current status of the aircraft, including its automation, is interpreted in the context of this mental model to build a mental representation of what the aircraft is currently doing and what it will do next. Hence, to be surprised, either the information fed into the mental model is inaccurate or the model itself is wrong. Pilots' expectations of what their automation will do may be in error when they attend to the wrong data, misinterpret data, or the data is in error. Alternately, their expectations may be wrong when their understanding of what the automation will do under the encountered conditions is wrong.

In this paper we examine pilot reports of unexpected automation behavior chronicled in the Aviation Safety Reporting System (ASRS) database in an effort to characterize the nature of these problems. Based on this understanding, technological and training strategies can be developed to prevent automation surprises and mitigate their effects.

Methods

The ASRS database was searched for automation-related event reports from 2012 by crews operating under Part 121. The initial search criteria were broadly specified to minimize the likelihood that relevant reports would be missed. Reports that mentioned automation, autopilot, auto throttle, flight management system, flight management computer, flight data computer, mode control panel, LNAV, VNAV, or any of the common abbreviations for these devices were retrieved. Of the 558 reports obtained, 234 described an event in which the pilots were surprised by unexpected actions of the auto-flight system.

The events that transpired before, during, and after the surprise were coded. Distinctions were made between five categories of events (actions or circumstances): precipitating, contributing, problem, detection, and response. *Precipitating* actions were those that preceded and led directly to the automation surprise. Within this category, we distinguished between primary precipitating or "catalytic" actions and secondary precipitating actions that occurred in response to the catalytic events. For example, in a number of cases, Air Traffic Control (ATC) instructions directed pilots to alter their previously programmed flight path. In entering flight path alterations into the flight management system (FMS), the pilots made an error that later resulted in a surprising aircraft behavior. We coded the ATC instructions as the primary or "catalytic" precipitating event and the pilot programming of the FMS as the secondary precipitating action. *Contributing* circumstances were those that did not directly precipitate the automation surprise but that may have contributed to the problem. For example, pilots may have reported being rushed or fatigued during the operation. *Problem* actions were those that produced the surprise. For example, in the prototypical ATC precipitated event described above, the pilots were often surprised by the aircraft veering away from the intended course. In this case, the problem event was coded as a course deviation. *Detection* actions were those that led to the discovery of the problem. These actions could involve direct observation of the aircraft behavior (as in the example above) by the pilots or ATC or observation of messages (e.g., Electronic Caution Alert Module (ECAM)), alerts (e.g., autopilot disconnect), or control movements. *Response* actions were those taken to resolve or recover from the surprise. These included actions such as taking manual control of the aircraft, switching to a lower level of automation (e.g., from VNAV to Mode Control Panel (MCP) control), or notifying ATC of a deviation. For each of these "action" categories, we coded the nature of the event, when the action occurred (phase of flight), who performed the action (e.g., ATC, crew, AFS), and the level of automation in use.

Results & Discussion

What Was Surprising?

A variety of different automation-related events surprised crews (see Table 1). In 15% (35) of the cases, the crew was surprised by changes in auto-flight system operation, including shutdown or freezing of various AFS components. In 11 of these cases, the autopilot disconnected. In three cases, the auto throttle disengaged or otherwise behaved unexpectedly. In 10 cases an unexpected mode change occurred and in 11 cases some component of the auto-flight system froze or failed.

In 12% (28) of the cases, the crew was surprised by the auto-flight system interface. Half (14) of these cases occurred when FMS data disappeared unexpectedly. In 73% (170) of the cases, the crew was surprised by aircraft behavior, including 21 cases in which the crew detected unexpected changes in aircraft control prior to a substantial change in aircraft position or velocity. However, in 64% (149) of all cases, the aircraft's velocity or position was altered substantially without the crew noticing. Twenty-seven of these resulted in airspeed changes, 35 in course alterations, 33 in altitude deviations and in 48 cases the aircraft's vertical path was affected unexpectedly.

Table 1.

Surprising Automation-Related Events

Event Type	Percent (n)	Event Type	Percent (n)
AFS operation only	15 (35)	**AFS problem affects aircraft control**	9 (21)
AFS Component Failure	4.7 (11)	**AFS problem affects aircraft behavior**	64 (149)
Auto Throttle	1.3 (3)	Airspeed	11.5 (27)
AP Disconnect	4.7 (11)	Altitude	14.1 (33)
Unexpected Mode Change	4.3 (10)	Course	15.0 (35)
AFS interface only	12 (28)	Localizer	2.6 (6)
Display	6.0 (14)	Vertical Path	20.5 (48)
FMS Drop	6.0 (14)	**Other**	0 (1)

Note. N=234; number of cases in parentheses.

When were crews surprised?

Overall, 55% of the automation surprises occurred during the arrival and approach phases of flight (see Table 2). By contrast, only 13% of the events occurred during cruise. However, this pattern differed according to the type of event. Failures or freezing of auto-flight system components, categorized as "AFS Component Failure" in Table 2, were evenly divided between climb, cruise, and approach. Two of the three auto-throttle events occurred in cruise. Although most of the autopilot disconnects, unexpected mode changes, and control anomalies occurred during arrival and approach, a substantial proportion occurred during climb and cruise. Display faults occurred equally during climb, cruise, and arrival. Waypoints dropping from FMS flight plans as well as lateral course deviations occurred mainly during climbs below 10,000 feet and arrival and approach phases.

Table 2.

Surprising Events by Phase of Flight

Surprising Event	Problem Phase of Flight										
	Before Push	**Take Off**	**Climb Below 10K**	**Climb Above 10K**	**Cruise**	**Descent Above 10K**	**Arrival**	**Approach**	**Go Around**	**Unknown**	**Total N**
AFS Component	0%	9%	18%	9%	27%	9%	0%	27%	0%	0%	11
Auto Throttle	0%	0%	0%	0%	67%	0%	0%	0%	33%	0%	3
AP Disconnect	0%	18%	0%	0%	27%	0%	36%	18%	0%	0%	11
Mode Change	0%	0%	30%	0%	20%	0%	40%	10%	0%	0%	10
Display	0%	14%	29%	0%	21%	0%	29%	7%	0%	0%	14
FMS Drop	7%	7%	29%	7%	0%	0%	7%	36%	0%	7%	14
Control	0%	0%	19%	10%	29%	5%	14%	19%	5%	0%	21
Airspeed	0%	0%	15%	15%	15%	0%	11%	37%	7%	0%	27
Altitude	0%	0%	0%	12%	3%	9%	36%	39%	0%	0%	33
Course	0%	0%	37%	6%	14%	0%	20%	17%	0%	6%	35
LOC	0%	0%	0%	0%	0%	0%	0%	100%	0%	0%	6
Vertical Path	0%	0%	4%	0%	2%	15%	69%	10%	0%	0%	48
Other	0%	0%	0%	0%	0%	0%	100%	0%	0%	0%	1
Total	0%	3%	15%	6%	13%	5%	31%	24%	2%	1%	234

There are several possible explanations for the disproportionate number of surprising events reported during the approach and arrival phases of flight. During these transitional phases crews are preparing for landing, traffic is increasing, and ATC often places additional demands on pilots. To comply with these demands, pilots make heavy use of their automation, resulting in discovery of automation problems that lay dormant during previous phases of flight. Furthermore, pilots may make errors while changing modes and programming flight plans. The ASRS reports provide indirect evidence for these explanations. Problems that are likely to be caused by issues with electronic components (e.g., AFS and display problems) were more likely than other problems (e.g., course or altitude deviations) to occur during cruise. This follows, given that electrical failures are likely to be dependent on the amount of time spent in operation, whereas flight path deviations are likely to occur when the aircraft is in proximity to other aircraft, and thus likely to be directed by ATC to change course or altitude.

Proximal Precursors of Surprising Events

Narratives of the ASRS reports were analyzed to determine the sequence of events that preceded the automation surprises. Sometimes, the pilots described probable causal sequences. This was particularly likely when the pilots determined that their actions had led to the surprise. In other cases, analysts could reasonably infer an event sequence based on pilots' descriptions of their actions and the automation response in combination with knowledge of AFS operations. Problems attributed to malfunctioning automation components rarely contained sufficient information to verify this conclusion. In many of these cases, pilots reported contacting maintenance, but rarely reported ensuing findings.

A precipitating event could not be determined with sufficient confidence in 21 of the examined cases. Of the remaining 213 cases, 66% (140) involved human errors in auto-flight system operation (see Table 3). Pilot actions engendered the majority of airspeed (64%), altitude (84%), course (71%), localizer (60%), and vertical path (73%) surprises. Pilot actions also led to the majority of surprises resulting from auto-flight system problems, control manipulations, display problems, unexpected mode changes, and dropped waypoints. However, in 24% of cases the AFS or another technological system was apparently responsible for triggering the surprise. This includes two out of three auto throttle changes, the majority (80%) of the autopilot disconnect events, and a large proportion of the auto-flight problems (44%), control manipulations (44%), display problems (38%), and mode changes (33%).

Table 3.
Surprising Event by Source of Event

Surprising Event	Source of Event					
	Pilots	**Environment**	**AFS**	**Other System**	**Other**	**Total N**
AFS Problem	44%	0%	33%	11%	11%	9
Airspeed	64%	28%	8%	0%	0%	25
Altitude	84%	3%	6%	3%	3%	31
Auto Throttle	33%	0%	67%	0%	0%	3
Control	44%	11%	22%	22%	0%	18
Course	71%	0%	17%	6%	6%	35
AP Disconnect	20%	0%	20%	60%	0%	10
Display	54%	0%	23%	15%	8%	13
FMS Drop	100%	0%	0%	0%	0%	11
LOC	60%	40%	0%	0%	0%	5
Mode Change	56%	11%	33%	0%	0%	9
Vertical Path	73%	2%	18%	2%	5%	44
Total	66%	7%	16%	8%	3%	213

As noted above, 66% of the surprising events were precipitated by pilot actions. However, these actions were unabetted in only a small proportion (28%) of these cases. In the remaining 72% of these cases, pilot actions were triggered by external events. In 52% of the cases (73), pilots were attempting to comply with ATC instructions when they inadvertently triggered the unexpected automation response. In the remaining 20% of cases, pilots were attempting to cope with equipment issues when they inadvertently triggered an unexpected automation action.

Detection

In general (78% of the time), pilots were the first to detect the surprising events. However, ATC detected the problems simultaneously or before the pilots in 20% of the cases. In 36% of altitude deviation cases, ATC detected the deviation simultaneously (12%) or before (24%) the pilots. In 69% of course deviation cases, ATC detected the deviation simultaneously (6%) or before (63%) the pilots. In 12% of vertical path deviations, ATC detected the deviation before the pilots.

Resolution

The automation level at which the aircraft was being operated at the time of the surprising event was compared to the automation level during resolution of the event. In 34 cases the automation during one period or the other could not be determined with reasonable certainty. In 48% of the cases (95), the same level of automation was maintained throughout the reported event. In 45% of the cases (90) in which the aircraft was being flown under some level of automation, pilots resorted to manual control following the surprising event. In the remaining 55% of the cases (110) automation was used in the recovery. When VNAV was in use at the time of the surprising event (72 cases), pilots resolved the issues and continued under VNAV 32% of the time. In the remaining VNAV cases, 22% of the crews used MCP inputs to control the aircraft and 42% resorted to manual control. When LNAV was in use at the time of the surprising event (36 cases), the pilots continued to fly using LNAV 56% of the time. In 14% of the cases they relied on the MCP. In 22% of LNAV cases, pilots resorted to manual control. When the aircraft was being controlled using the MCP at the time of the event (28 cases), pilots continued to fly using the MCP in 64% of the cases and resorted to manual control 32% of the time.

General Discussion

Three important conclusions can be drawn from the results discussed above. First, many different factors may precipitate automation surprises. These include problems in the auto-flight system, problems in the displays and interface with the automation, problems in other aircraft sensors and systems, interactions with weather and other aspects of the external environment, and inappropriate actions taken by the pilots. Second, inappropriate actions by the pilots are involved in a large proportion of the automation surprise events. Third, recovery from automation surprises need not require reversion to manual control. In many cases, pilots continued to fly successfully using the same level of automation used prior to the automation surprise. Based on these observations, it is clear that there is no single general intervention that can prevent automation surprise or completely mitigate its effects. However, several different tacks taken together may be particularly effective.

First, new methods for pilot automation training need to be developed and tested. A large portion of the reported automation surprises can be traced to inappropriate pilot actions. In some cases, pilots understood what had happened after the fact. In other cases they did not, but probable causes of the surprises were apparent in their reports. Providing pilots with a better understanding of their automation would likely decrease the number of surprises. Producers of automated systems have long touted their ability to simplify pilots’ tasks and improve precision and efficiency. However, researchers have repeatedly noted that while aviation automation has improved the efficiency and precision of operations, it has not reduced complexity. Indeed, automation has increased the complexity of the pilots’ job. Training has not kept pace. Methods for automation education need to be developed which can help pilots develop an understanding of their automation that allows them to anticipate automation actions and not simply respond with a small set of canned procedures.[1] For pilots to construct adequate mental models of automation, they do not need to know the intricacies of the underlying engineering, but they must know how the system interacts with the environment – how it obtains information, what it controls, and what targets it is trying to achieve. Hence, pilots must be trained to understand: 1) what is controlled by each automation mode, 2) where each mode obtains data about the current state of the aircraft, 3) where each mode obtains targets, and 4) what actions each mode will take when the target is achieved. But having this knowledge is not sufficient. It merely provides the framework for the model. At every point during a flight, the model must be populated with current information about the state of the aircraft and how it relates to the intended flight path. This requires that pilots: 1)

[1] One area of particular difficulty appears to be the interaction between the auto-flight functions controlled through the mode control panel and those controlled by the flight control computer through a display unit.

know where to find the relevant information, 2) attend to these sources, 3) interpret the information correctly, and 4) integrate this information with their stored knowledge of the automated flight system and intended flight path.

Second, improved displays need to be developed that provide pilots with predictive indications. In many ASRS cases, the automation performed as it had been programmed to perform. However, errors in the pilots' programming or other inappropriate actions led to a discrepancy between what the pilots thought the system was programmed to do and what it was actually programmed to do. Typical automation interfaces do not provide clear displays of the programmed and predicted flight paths. Without this support, errors that humans inevitably make may go unnoticed until the aircraft is substantially off course, altitude, or airspeed.

Third, a large portion of pilot precipitated automation surprise events were themselves caused by instructions from ATC that proved problematic for pilots. A substantial decrease in the number of automation surprise events likely could be attained through restructuring ATC arrival and approach procedures. Decreasing the number of unnecessary "mission surprises" with which pilots must cope is likely to decrease the number of automation surprises. For example, new RNAV arrival procedures may be so complex that they cannot be reliably flown manually. However, ATC procedures allow controllers to vector aircraft into these procedures and to alter their components. Because these approaches effectively must be flown by the automation, modifications force pilots to program the FMS while flying the procedure. In this process, errors may be made that surprise the pilots and disrupt the flow of traffic. Modifying ATC procedures could substantially decrease the number of these problems.

The results reported here also underscore the importance of understanding and developing strategies for addressing the problem of automation surprise before NextGen becomes fully operational. In a large proportion of the cases examined, ATC called pilots' attention to a deviation from the planned course or altitude. Frequently, ATC handled the problem by providing a new clearance. Under NextGen, aircraft may fly in close proximity along defined 4D paths. Deviations such as those observed here would bring aircraft dangerously close to one another. At best, these events would cause substantial disruption to the traffic flow. At worst, they could result in collisions.

Acknowledgements

This work was funded by NASA NRA NNX12AP14A. Special thanks to Lance Sherry, Immanuel Barshi, and Michael Feary for technical suggestions.

References

Burki-Cohen, J. (2010). Technical Challenges of Upset Recovery Training: Simulating the Element of Surprise. In *Proceedings of the AIAA Guidance, Navigation, & Control Conference*: Toronto, CA.

Degani, A., Shafto, M., & Kirlik, A. (1999). Modes in human-machine Systems: Constructs, representation, and classification. *International Journal of Aviation Psychology*, *9*(2), 125-138.

Feary, M., McCrobie, D., Alkin, M., Sherry, L., Polson, P., Palmer, E., & McQuinn, N. (1998). Aiding Vertical Guidance Understanding. In *NASA Technical Memorandum NASA/TM- 1998-112217*, Ames Research Center, Moffett Field, CA.

Javaux, D., & De Keyser, V. (1998). The Cognitive Complexity of Pilot-Mode Interaction: A Possible Explanation of Sarter and Woods' Classical Result. In *Proceeding of the International Conference on Human-Computer Interaction in Aeronautics* (pp. 49-54). Montreal, Quebec: Ecole Polytechnique de Montreal.

Norman, D. A. (1990). The 'problem' with automation: inappropriate feedback and interaction, not 'over-automation.' *Philosophical Transactions of the Royal Society B: Biological Sciences*, *327*(1241), 585-593.

Reveley, M., Briggs, J., Evans, J., Sandifer, C., & Jones, S. (2010). Causal Factors and Adverse Conditions of Aviation Accidents and Incidents Related to Integrated Resilient Aircraft Control. In NASA TM-2010-216261.

Sarter, N. B., & Woods, D. D. (1995). How in the world did I ever get into that mode? Mode error and awareness in supervisory control. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *37*(1), 5-19.

Sherry, L., & Mauro, R. (2014). Controlled Flight Into Stall (CFIS): Functional Complexity Failures and Automation Surprises. In *Integrated Communications Navigation and Surveillance Conference*.

Woods, D. & Sarter, N. (2000). Learning from Automation Surprises and "Going Sour" Accidents. In Sarter, N. & Amalberti, R. (Eds.) *Cognitive Engineering in the Aviation Domain*. LEA: Mahwah, NJ.