THE JOURNAL OF NAVIGATION (2002), **55**, 83–96.
DOI: 10.1017/S0373463301001588 Printed in the United Kingdom

On Your Watch: Automation on the Bridge

M. H. Lützhöft and S. W. A. Dekker

(*Linköping Institute of Technology, Sweden*)

In this paper, we discuss the grounding of the *Royal Majesty*, reconstructed from the perspective of the crew. The aim is particularly to understand the role of automation in shaping crew assessments and actions. Automation is often introduced because of quantitative promises that: it will reduce human error; reduce workload; and increase efficiency. But as demonstrated by the *Royal Majesty*, as well as by numerous research results, automation has qualitative consequences for human work and safety, and does not simply replace human work with machine work. Automation changes the task it was meant to support; it creates new error pathways, shifts consequences of error further into the future and delays opportunities for error detection and recovery. By going through the sequence of events that preceded the grounding of the *Royal Majesty*, we highlight the role that automation plays in the success and failure of navigation today. We then point to future directions on how to make automated systems into better team players.

KEY WORDS

1. Automation. 2. Human Factors. 3. Maritime. 4. Integration.

1. INTRODUCTION.

One of the .conundrums in navigation today is why increased computerisation and automation may not have removed, and perhaps not even reduced, the potential for failure in the systems in which they were introduced. In fact, new pathways to breakdown appear to have opened up. Recent groundings (for example the *Royal Majesty* (NTSB, 1997)) and air crashes (for example Aeronautica-Civil, 1996) appear linked by a common pattern. Following a small trigger event, a series of misco-ordinations and miscommunications between humans and machines evolves and deepens over time. Sarter and Woods (1995) have called it the 'Going Sour' scenario. The mismatch between where people think they are and where their machines have really taken them grows and goes unnoticed, leaving the craft under automatic control even in the last minutes before grounding or crashing. The NTSB report labels the *Royal Majesty* incident a 'single-point catastrophic failure'. Granted, single point failures played a role, as they always do, but not any single failure could have been responsible for the outcome. Accidents are the result of multiple factors, all necessary and only jointly sufficient. Focusing on a 'single-point failure' critically misses the evolving, building, escalating signature that lies at the heart of today's problems related to human-automation interaction.

According to the NTSB, in reference to the grounding of the *Royal Majesty* (NTSB, 1997), there is a lot of research that sets out to explain why accidents and incidents involving people and automation happen. Research shows that humans are poor monitors of automated systems; research shows that humans tend to rely on warning systems and not manual checks; research shows that reliable systems are perceived as trustworthy by operators; research shows that salient cues will bias operators in their decision making. Nevertheless, the board concludes that the watch standing officers of the *Royal Majesty* 'despite repeated indications … failed to recognise … numerous opportunities to detect … that the vessel had drifted off track'. The *Royal Majesty* report reveals a tension typical of accident analyses (Galison, 2000): despite research evidence that shows how human behaviour is systematically and reasonably coupled to, and shaped by, the tools and tasks they work with, failures ultimately get explained by relying chiefly on human motivational shortcomings. If only people had done their best, if only they had behaved more professionally, this or that outcome would not have occurred. Such conclusions sponsor the notion that we need even more automation or computerisation to compensate for erratic and fallible humans. For example, Goossens and Glansdorp (1998) recently suggested: 'Improvements might best be achieved by reducing or eliminating the human factor in incident sequences' (p. 368). But this would be a premature countermeasure at best. We are only now beginning to understand the profound qualitative implications that automation has on human work and, ultimately, system safety and success.

In this paper, we trace the events of the *Royal Majesty* once again. Our aim is to deepen our understanding of automation interaction problems on ships' bridges today. At every twist in the plot, we start with the crew's assessments and actions that befuddle and surprise outside observers. How could they not have cross-checked? How could they have been so overconfident? Yet rather than wondering why the crew failed to do what we would now have done (with full knowledge of the outcome and true nature of the circumstances surrounding them at the time), we try to understand why they did what they did. After all, saying what people should have done does not explain why they did what they did. Abiding by the local rationality principle of human factors (people do reasonable things given their knowledge, their goals, their limited resources) we have converted the search for human failures into a hunt for human sensemaking – why did this action or assessment make sense to people at that time and place? Different stories often struggle into view when pursuing this question; in this case, a story of automation surprises that holds many lessons for the field of navigation.

2. DEPARTURE FROM BERMUDA.

The voyage of the *Royal Majesty* is divided into three parts, starting with the departure from Bermuda.

2.1. *The Antenna Cable*. The *Royal Majesty* (henceforth *RM*) departed Bermuda bound for Boston at 12:00 noon on the 9th of June 1995 (NTSB, 1997). The visibility was good, the winds slight and the sea calm. Before the departure the navigator checked the navigation and communication equipment and found it in 'perfect operating condition'. About half an hour after departure, the harbour pilot disembarked and the course was set towards Boston. Just before 13:00 there was a cut-off in the signal from the Global Positioning System (GPS) antenna, routed on the fly bridge (the roof of the bridge), to the receiver – leaving the receiver without satellite signals. Post accident examination showed that the antenna cable had separated from the antenna connection.

Why was the cable routed in this inappropriate manner, and how come no one noticed it before? Did someone stumble over the cable? Why would someone walk across the fly bridge in the first place and why did this person not notice what happened? It should have been obvious to someone that this cable could be damaged and that this would lead to dire straits.

It is quite possible that this open and inappropriate routing of the cable is due to the moving of the antenna earlier in the year, the reason for which will be discussed later. The person moving the antenna could probably see no reason to do the job any differently, and most likely lacked complete knowledge of the context of use of the fly bridge. The separation could have been due to 'wear and tear' over a longer period of time, but also someone stumbling over it, causing it to part instantaneously. The time of signal loss implies that someone could have been lowering flags, as is standard practice after departure. This could explain why the possible stumbling over the cable was not noticed, as the person would be preoccupied with untangling halyards and flags.

2.2. *The GPS*. The GPS, when losing the satellite signals promptly defaulted to dead reckoning (DR) mode, sounded a brief aural alarm and displayed two codes on the display: DR and SOL (SOL means that positions based on satellite signals cannot be calculated). These alarms and codes went unnoticed.

Why was no one aware that a loss of satellite data would automatically lead to DR mode, and that this in turn would affect the whole navigation system? How is it possible that not one of the highly trained professionals on the bridge heard the alarm, or 'attended to' the indications on the display? After all, if the visual indications were in plain view on the screen for 34 hours, monitoring must have been deficient?

About 15 years ago, when this particular GPS receiver was manufactured, the GPS satellite system was not as reliable as it is today. Therefore the receiver could, when satellite data was unreliable, use a DR mode in which it estimated positions using an initial position, the gyrocompass for course input and a log for speed input. The GPS thus had two modes, 'normal' and DR, between which it switched autonomously depending on the accessibility of satellite signals.

GPS satellite coverage had, at the time of the incident, been all-inclusive and working well for many years, and the crew did not expect anything out of the ordinary. The GPS antenna was moved in February, since parts of the superstructure occasionally would block the incoming signals, which caused temporary and short (a few minutes, according to the captain) periods of DR navigation. This was to a great extent remedied by the antenna move, as the Majesty Cruise Line's electronics technician testified, and nothing in the officers' testimonies suggests that this had been a problem during the present trip. Several of the officers also testified that they relied on the GPS position data and considered other systems to be back-up systems, and that *the only times the GPS positions could not be depended on for accuracy were during these brief periods in DR mode*. Thus, the whole bridge crew was aware of the DR mode option, and how it worked, but none of them ever imagined or were prepared for a loss of satellite data caused by a cable break; no previous loss of satellite data was ever so swift and so absolute.

When the GPS switched from normal to DR, an aural alarm sounded and a visual indication was shown on the display. The aural alarm sounded like that of a digital wristwatch and was less than a second long. This was the first opportunity to notice that the mode had changed. Since the time-window in which the mode must have changed is between 12:00 and 13:00, and this was a busy time on board, a combination of factors may have created a situation in which the alarms went unnoticed. A departure involves complex manoeuvring, there are several crew-members on the bridge and there is a great deal of communication. When a pilot disembarks, the operation is highly time-constrained and risky. In such situations, the aural signal could easily have been drowned out, and assuming that it was, no one would be expecting the DR mode, and thus the visual indications were not seen at the time.

Even if the initial alarm was missed, the mode indication was continuously available on the GPS display. None of the bridge crew saw it, according to their testimonies. If they had seen it, they knew what it meant, literally translated – dead reckoning means no satellite fixes. But there is a crucial difference between data that in hindsight can be shown to have been available and data that was observable at the time (Dekker and Woods, 1999). Observability demands cognitive work such as remembering and reasoning that leads to the extraction of meaning from available data. The indications on the display (DR and SOL) were placed between the two rows of numbers, latitude and longitude, that indicate the ship's position on the screen, and were about one-sixth the size of those numbers. There was no difference in the size and character of the position indications after the switch to DR. The size of the display screen was about 7·5 by 9 centimetres, and the receiver was placed at the aft part of the bridge on a chart table, behind a curtain. The location is reasonable, since it places the GPS, which supplies raw position data, next to the chart that is normally placed and used on the chart table. Only in combination with a chart would the GPS data make sense; however, the GPS data was also forwarded to the integrated system.

For the crew of the *RM*, observability meant they would have to leave the forward console, actively look at the display, and expect to see more than a position. Even then, if they had seen the two-letter code and translated it into the expected behaviour of the ship, it is not a certainty that the immediately available conclusion would have been 'this ship is not heading towards Boston anymore'. A critical test of observability is whether an indication helps practitioners see what they did not expect to see, or more than they expected (*ibid.*). The GPS on the *RM* would not have passed the critical test. Thus, when the officers did leave the forward console to plot a position on the chart, they looked at the display and saw a position, and nothing *but* a position, because that is what they were expecting to see. It is not a question of 'not attending to the indications'; they were attending to the indications, the *position* indications, since plotting the position is the professional thing to do, and so the mode change continued to pass unnoticed.

2.3. *Communication GPS-Autopilot*. If the mode change was so non-observable on the GPS display, why was it not shown more clearly somewhere else? How could the loss of signals reverberate throughout the system, and have such consequences? How could one small failure have such an effect, were there no back-ups, and if not – why not?

The *RM* had a modern integrated bridge system, of which the main component was the navigation and command system (NACOS). The NACOS consisted of two parts, an autopilot part to keep the ship on course and a map construction part, where simple maps could be created and displayed on a radar screen. When the *RM* was being built, the NACOS and the GPS receiver were delivered by different manufacturers, and they, in turn, used different versions of the electronic communication standards.

Due to these differing standards and versions, valid position data and invalid DR data sent from the GPS to the NACOS were both 'labelled' with the same code (GP). The *installers* of the bridge equipment were not told, nor did they expect, that position data (GP-labelled) sent to the NACOS would be anything but valid position data. The *designers* of the NACOS expected that if invalid data were received it would have another format. Due to this misunderstanding the GPS used the same 'data label' for valid and invalid data, and thus the autopilot could not distinguish between them. Since the NACOS could not detect that the GPS data was invalid the ship sailed on an autopilot that was using estimated positions until a few minutes before the grounding.

A principal function of an integrated bridge system is to collect data such as depth, speed and position from different sensors, which are then shown on a centrally placed display to provide the officer on watch (OOW) with an overview of most of the relevant information. The NACOS on the *RM* was placed at the forward part of the bridge, next to the radar screen. Systems using current technology commonly have multiple levels of automation with multiple mode indications on many displays (Sarter and Woods, 1995). An adaptation of work strategy is to collect these in the same place and another solution is to integrate data from many components into the same display surface. This presents an integration problem for shipping in particular, where quite often components are delivered by different manufacturers.

The centrality of the forward console in an integrated bridge system also sends the implicit message to the OOW that: navigation may have taken place at the chart table in times past, but as of now the work is performed at the console. The chart should still be used, to be sure, but only as a back-up option and at regular intervals (customarily every half-hour or every hour). The forward console is perceived to be a place where all the information needed to navigate the ship safely will be supplied.

2.4. *The NACOS.* As mentioned, the NACOS consisted of two main parts; the position data sent by the GPS (via the radar) to the NACOS in order to keep the ship on track (autopilot part) *and* to position the maps on the radar screen (map part). The autopilot part had a number of modes that could be manually selected; NAV and COURSE. NAV mode kept the ship within a certain distance of a track, and corrected for drift caused by wind, sea, and current. COURSE mode was similar but the drift was calculated in an alternative way. The NACOS also had a DR mode, in which the position was continuously estimated. This backup calculation was performed in order to compare the NACOS DR with the position received from the GPS. To calculate the NACOS DR position, data from the gyro compass and Doppler log was used, but the initial position was regularly updated with GPS data. When the *RM* left Bermuda, the Navigation Officer chose the NAV mode, with the input coming from the GPS, normally selected by the crew during the three years the vessel had been in service.

If the ship had deviated from her course more than a pre-set limit, or if the GPS position differed from the DR position calculated by the autopilot, the NACOS would have sounded an aural alarm and clearly shown a visual alarm at the forward console (position-fix alarm). There were no alarms since the two DR positions calculated by the NACOS and the GPS were identical. The NACOS DR, which was the perceived backup, was using GPS data, believed to be valid, to refresh its DR position at regular intervals. This is because the GPS was sending DR data, estimated from log and gyro data, but labelled as valid data. Thus, the radar chart and the autopilot were using the same inaccurate position information, and there was no display or warning of the fact that DR positions (from the GPS) were being used. Nowhere on the integrated display could the OOW confirm what mode the GPS was in, and what effect the mode of the GPS was having on the rest of the automated system, not to mention the ship.

2.5. *Mode Awareness.* Mode awareness is the ability of a supervisor to track and anticipate the behaviour of automated systems (Sarter and Woods, 1995). To maintain mode awareness, the operator has to keep track of what the automation is doing, and how contextual factors are evolving. Newer automated systems respond to operator input as well as to environment/situation and system factors, and a mode can be manually selected but also automatically engaged by the system. The complexity of the departure and dropping off of the harbour pilot is considerable, which takes time and resources from the task of tracking and anticipating whether the situation has had or will have some effect on the automation. Breakdowns in mode awareness are part of the cause for automation surprise, which is discussed below.

2.6. *Automation Surprise.* All the prerequisites of an automation surprise are now looming on the horizon. As research shows (Sarter and Woods, 1995; Woods and Sarter, 2000), factors that strongly increase the potential for automation surprise are:

(a) Automated systems act on their own without immediately preceding directions from their human partner.
(b) There are gaps in users' mental models of how their machine partners work in different situations.
(c) The feedback is weak about the activities and future behaviour of the agent relative to the state of the world.

These three factors were undeniably present, and the following sections summarise their influence on the cognitive system on the bridge of the *RM*.

2.6.1. *Autonomous Automation.* The automated system on board the *RM* acted autonomously, and Woods, Johannesen, Cook and Sarter (1994) point out that automation has changed from 'one-input, one-action' systems towards carrying out long action sequences without demanding further input from the operator. In the *RM* case there was not even an initial input from the operator. The command had been hard-wired into the system, and no one was aware of it, which makes for another variant of mode error and automation surprise. The mode transition was effected almost silently and absolutely automatically when the GPS receiver lost the signals from the satellites.

2.6.2. *Mental Models.* Results from several studies quoted in Woods *et al.* (1994) indicate that practitioners not only have 'buggy' mental models, but also that they are largely unaware of this. Factors that influence the quality of a mental model are expectations and knowledge, training and education, and actual experience of using a system in various real-life situations.

The officers brought onto the ship several years of working experience. They were, apart from the First Officer, relatively new to the *RM*, and it was for all their first ship with an integrated bridge system. To prepare for this, they were given between 3 weeks and one month on-the-job training, and some of them had read the manuals. The manufacturer of the integrated system did have classroom and simulator training on offer, but the owners did not purchase this, which it is not required by any regulations.

Thus the officers' experience and knowledge of ship handling generally was on a professional level, but their knowledge of and training for integrated navigation systems was limited. It has been shown that automation creates new kinds of knowledge demands (*ibid.*). Operators must have working knowledge of the functions of the automation in different situations, and know how to co-ordinate their activities with the automated system's activities. This was not supplied to the crew, and therefore their mental models and expectations must have been sketchy, to say the least. Mental models may also be supplemented by vicarious experience; learning from other operators, as in on-the-job training. This extends the individual mental model but also makes for the spreading of 'naïve knowledge', and Hutchins comments: 'where there is the need for learning there is room for error' (1995, p. 272). Furthermore, this kind of training typically takes place in calm routine contexts, so that the trainees may be insufficiently prepared to handle unusual or emergency circumstances, which includes having little knowledge of how the automated system will behave under such conditions.

2.6.3. *System Feedback*. Since the introduction of automated machines, designers and engineers versus users and practitioners have had different ideas of what constitutes a good display of the status and behaviour of technology. This third problem also exacerbates the second (mental models), since if humans did have access to (perfect) internal world models, it wouldn't matter if the automated systems had opaque indications and weak feedback. But since many displays afford low observability, the mental models will be even more incomplete. The feedback from the GPS was decidedly weak, available but not observable. Furthermore, the integrated system did not show any future tracks, except for the planned track, making it increasingly difficult for the officers to anticipate the future behaviour of the ship.

In addition to this, there were no immediate and perceivable effects on the ship since the GPS calculated positions using the log and the gyrocompass. It cannot be expected that a crew should become suspicious of the fact that the ship actually is keeping her speed and course, since that is why the automation was installed in the first place. The combination of a busy departure, an unprecedented event (cable break) together with a non-event (course keeping) and the change of the locus of navigation (including the intra-system communication difficulties) shows that it made perfect sense, in the situation and at the time, not to notice and 'attend to' the mode change.

3. THE OCEAN VOYAGE.

Given that the mode change was in practice very difficult to spot, there was still a long voyage at sea where the problem could have been, and should have been, detected. Why did not one of the officers cross check the GPS position against another source, such as the Loran-C receiver that was placed close to the GPS? Why did nobody notice that the *RM* was drifting further from her route with every passing hour? Why were the procedures (e.g., cross checking, alerting the captain of *any* uncertainty related to the ship's position) not followed? Why did the officers, inappropriately, rely solely on the NACOS position-fix alarm to warn them of problems with GPS data? Why did the Master not notice that something was wrong?

The moment of automation surprise is not when the underlying event occurs but when the system starts acting strangely as a consequence. Until the very last minutes before the grounding, the ship did not act strangely; it was a routine trip, the weather was good and the watches and watch changes uneventful. There were, but only in hindsight, further cues indicating that the situation was not completely under control, and further mechanisms will be invoked to explain why they were not perceived as warnings at the time.

3.1. *Cross-Checking*. Several of the officers claim to have checked the displays of both the GPS and Loran-C receivers, but only used the GPS data to plot positions on the paper chart. Without plotting, assessing a possible error in position by looking at two sets of numbers is a significant cognitive operation, which is yet another availability/observability problem. The data were available, but it was nigh impossible to actually observe the implications of the difference between numbers alone. Apart from this, there actually was some cross checking, conscious or not. The position on the radar map was checked against the position on the paper chart hourly, and the First Officer most likely perceived sighting the first buoy as cross checking with GPS data. Another powerful, but unintentional, reassurance was that the Master, on a number of occasions, spent several minutes checking the position and progress of the ship, and did not make any corrections. The officers may have perceived it as an inspection followed by a validation of the situation (Snook, 2000).

The reason the ship drifted so far from her intended route was almost certainly the easterly winds in combination with wind-induced current. The drift was not apparent, since the automation did not provide any means of projecting the future (Dekker and Woods, 1999) instead being real-time (barely) and even then providing little feedback of its current behaviour.

It has been shown that operators will monitor less effectively when automation is installed, and even more so if the automation has been operating acceptably for a long period (Bainbridge, 1983; Wiener, 1988). This was apparently the situation on the *RM*; the officers monitored the automation less effectively since it had been reliable for a long time. This does not imply that the navigation task itself was performed less effectively, only that the crewmembers believed that it was safe to shift their watchfulness to other tasks. Furthermore, the NTSB report implies that it is a mistake to rely on warning systems, and that manual checks should be used but also adds that 'it is likely that they were not aware of the inherent limitation'. Then why is a warning system installed in the first place; to increase the workload?

3.2. *Procedures and Canonical Practice*. A multitude of regulations, procedures, and manuals were available to the crew and applicable to the navigation task; for example, the STCW convention (International Convention on Standards of Training, Certification and Watchkeeping for Seafarers (IMO, 1996)), equipment manuals and local company procedures. The NTSB report argues that if all rules had been adhered to, this incident might not have happened. There are, however, questions about rules in general that should be asked (Degani and Wiener, 1998): are the procedures adequate for the task, how where they taught and were they logical and consistent? The procedure for cross checking (to take but one example) is repeated in the STCW and the manual for the NACOS, but not in the captain's bridge procedures, nor in a company circular concerned with duties of the officers. If a rule is found in some places but not all, is this consistent? If a procedure is said to be canonical practice, good seamanship even, and still not followed, is it really adequate or is there a deeper problem within the system? Finally, it is probably true that the rules and procedures were neither taught nor learned in a sufficient manner. In the *RM* case, the ship appeared to follow its planned course, and the only aid judged necessary to check was the GPS.

Furthermore, Snook (2000) comments that in any organisation there is a tendency to write too many rules, with the effect that trying to follow them makes normal work cumbersome. Practitioners will adapt globally constructed rules to fit their local working context and gradually work will be performed by 'standard practice' rather than 'by the book'. The adaptation of rules leads to an increasing gap between procedures and practice. An in-depth analysis of this phenomenon is beyond the scope of this paper, but there is a strong indication that inadequate procedures can be contributing factors to incidents instead of preventing them. A system will not become safer if there is a rule for every (imaginable) contingency *and* a new rule added after each incident (Dekker, 2001).

3.3. *Local Rationality*. The principle of bounded rationality (Simon, 1957) explains that the human capacity for problem solving is quite small when compared to the problem and solution spaces. In other words, humans have finite capabilities, and the number of possible situations is infinite. Local rationality is a concept that reduces further the scope of bounded rationality, taking the view from the inside as seen by an operator (Woods *et al.*, 1994). In this inside context, operators use knowledge they possess *and* perceive as relevant, focus of attention and strategic trade-offs to reach a desirable performance.

Desirable performance accommodates comfortable margins (Wioland and Amalberti, 1996), and such margins reserve some cognitive resources for monitoring, detection and anticipation, as well as reducing the risk of fatigue or overload. Strategies are devised to cope with the mismatch between capabilities and demands, such as when humans in complex environments tend to use heuristics and simplify situations (*ibid.*). The following section gives an example of a strategy devised by the crew of the *RM*.

3.4. *Strategies*. Before the GPS antenna was moved, the short spells of signal degradation that lead to the GPS switching to DR mode also caused the radar map to 'jump around' on the radar screen (the crew called it *chopping*) since the position would change erratically. The reason *chopping* was not observed on this particular occasion was that the position did not change erratically, but in a consistent manner by dead reckoning. It is entirely possible that the satellite signal was lost before the autopilot was switched on, thus causing no shift in position. The crew had developed a strategy to deal with this occurrence in the past. When the position-fix alarm sounded they first changed modes (from NAV to COURSE) on the autopilot and then they acknowledged the alarm. This had the effect of stabilising the map on the radar screen so that it could be used until the GPS signal returned. It was an unreliable strategy, since the map was being used without knowing the extent of error in its positioning on the screen. Moreover, it also led to the belief that, as mentioned earlier, the only time the GPS data was unreliable was during *chopping*. *Chopping* was more or less alleviated by moving the antenna, which means that by eliminating one problem a new pathway for accidents was created. The strategy of using the position-fix alarm as a safeguard no longer covered all or most of the instances of GPS unreliability.

This local (apparently) efficient procedure would almost certainly not be found in any manuals. It gains legitimacy through unremarkable repetition (Snook, 2000) and thus over time becomes common practice. In the present case, it seems to have led to the belief that a stable map is a good map, with the crew concentrating on the visible signs instead of being wary of the errors hidden below the surface. The *chopping* problem had been resolved for about four months, and trust in the automation slowly grew. The officers made sensible decisions given the constraints on their tasks, at the time, in the specific situation and context, whereas these strategies might not have been perceived as sensible in another context, or in hindsight (Woods *et al*., 1994). Humans interpret what they perceive in their environment to make sense, and also tend to misjudge the probability of events occurring.

> [Humans have] the reasonable expectation that the recurrences of the past provide a fair guide to the likelihoods of the future. (Reason, 1988, p. 8).

4. FIRST BUOY TO GROUNDING. Acknowledging that local rationality to some extent explains why the crew acted as they did at sea, this should have changed when landfall was drawing closer. Why did the Chief Officer and the Second Officer not follow 'longstanding watchkeeping practices when approaching land'? Why did the First Officer not positively identify the first buoy? How come the error in position was not detected, given that the vessel was so close to the shore? Furthermore, why did the Second Officer not take any action after hearing the lookouts report red lights and later blue and white water? Why did he not understand that the warning broadcasted on the VHF concerned his vessel? For what reason did he ignore the fact that he did not see the second buoy on the radar, and even told the Master that it *had* been sighted? In short, why did the crew consistently fail to recognise all warnings, repeated indications that the vessel was not on its intended track and numerous opportunities to avoid the grounding?

The first buoy ('BA') in the Boston traffic lanes was passed at 19:20 on the 10th of June, or so the Chief Officer thought. The buoy identified by the First Officer as the 'BA' turned out to be the 'AR' buoy placed about 15 miles to the west-south-west of the 'BA'. A traffic lane is a separation scheme delineated on the chart to keep meeting and crossing traffic at safe distance and to keep ships away from dangerous areas. It made perfect sense to the First Officer to identify the 'AR' as the correct buoy since the echo on the radar screen coincided perfectly with the mark on the radar map that signified the 'BA'. This was in fact a stochastic fit; he expected to see it, and there it was, and the influence of local rationality further strengthened his belief that they were on the right track. At this point in time the First Officer probably even believed he had cross-checked his position by two independent means; the radar map and the buoy. An uncontrollable factor and an unfortunate coincidence was the sun glare on the ocean surface that made it impossible to visually identify the 'BA'.

4.1. *Changing of the Watch.* An especially problematic aspect of keeping track of an evolving situation manifests itself when several people use an automated system, simultaneously as in an aircraft (Sarter and Woods, 1995), or consecutively as the watchkeeping officers on a ship. At 20:00 the Second Officer took over the watch from the Chief Officer. The Chief Officer presumably gave the vessel's assumed position, as is good watchkeeping practice. The Second Officer had no reason to doubt that this was a correct position, especially given that the Chief Officer had been at sea for 21 years, spending 30 of the last 36 months on board the *RM*. Shortly after the take-over, the Second Officer reduced the radar scale from 12 to 6 nautical miles. This is normal and even canonical practice as vessels come closer to shore or other restricted waters.

By reducing the scale, there is less to monitor, thereby increasing the likelihood to see anomalies and dangers. It was an act to create safety; instead, it had the effect of reducing his scope to notice the actual situation.

When the lookouts later reported lights, several factors may have interacted as to why these reports were not acted upon. Firstly, the Second Officer had no expectation that there was anything wrong, as his model of the world implied that the vessel was safely in the traffic lane, and therefore he was inclined to judge incoming information in that light. Secondly, lookouts are liable to report everything indiscriminately, and could have been perceived to be less than well trained; it is always up to the officer of the watch to decide whether to take action or not. Thirdly, there is a considerable possibility that there is a cultural and/or hierarchical gradient between the officer and the lookouts. At this time, the Master also visited the bridge, and just after he left, there was a VHF call. This escalation of work may well have distracted the Second Officer from considering the lookouts' report.

4.2. *Radio Communication*. After the accident investigation was concluded, it was discovered that two Portuguese fishing vessels had been trying to call the *RM* on the VHF radio to warn her of the imminent danger (the US Coast Guard has a recording of the transmissions). The calls were made about $1\frac{1}{2}$ hours before the grounding, at which time the *RM* was already 16·5 nautical miles from where she was believed to be. At 20:42, one of the fishing vessels called 'fishing vessel, fishing vessel call cruise boat' on channel 16 (the international distress channel, which is only to be used for emergency traffic or for establishing contact). Immediately following this first call in English the two fishing vessels started talking to each other in Portuguese. One of the fishing vessels tried to call again a little later, giving the position of the ship he was calling.

Calling on a VHF without positively identifying the intended receiver can lead to mix-ups with disastrous results (MAIB, 1999). Or in this case, *if* the Second Officer heard the first English call and the ensuing conversation, he most likely disregarded it since it seemed to be two other vessels talking to each other. Furthermore, as he was using the 6-mile scale, he could not see the fishing vessels on his radar. If he heard the second call and checked the position he might well have decided that the call was not for him, as it appeared that he was far from that position.

4.3. *The Second Buoy*. At about this time, the second buoy should have been seen, and around 21:20 it should have been passed, but was not. The Second Officer assumed that the radar map was correct when it showed that they were on course. To him the buoy signified a position, a distance travelled in the traffic lane, and reporting that it had been passed may have amounted to the same thing as reporting that they had passed the position it was (supposed to have been) in. The Second Officer did not, at this time, experience any accumulation of anomalies, warning him that something was going wrong. In his view this buoy, which was perhaps missing or not picked up by the radar, was the first anomaly, and not perceived as a large one. Paraphrasing the 'Bridge Procedures Guide', it is said that the Master should be called when:

(a) something unexpected happens,
(b) when something expected does not happen (e.g. a buoy), and
(c) at any other time of uncertainty.

It is all very well to define an unexpected event, but when it happens, people tend quickly to rationalise it. This is even more so in the case of *not* seeing what was expected: 'well, I guess the X isn't doing Y … ', clearly an act of local rationality. The NTSB report, on the other hand, lists at least five actions that the officer should have taken. He did not take any of these actions, because in his world-view he was not missing opportunities to avoid the grounding. He was navigating the vessel safely to Boston.

The Second Officer's model of the world was thus very stable. The radar was suspected of being unreliable ('perhaps the radar did not reflect the buoy'), but the radar map was perceived as very reliable. So reliable in fact, that the crew, including the Master, had 'sailed it' the entire trip, despite the view that 'fundamental seamanship practises caution against relying on one source of position information'. Then again, the Second Officer saw and experienced the whole crew doing just this, in his time on board. Everyone trusted the radar map, the GPS, and the position-fix alarm.

A second stochastic fit is the Master's timing. He visited the bridge just before the VHF call, called the bridge about one hour after it, and made a second visit around 22:00. The times at which he chose to visit the bridge were calm and uneventful, and did not prompt the Second Officer to voice any concerns, nor trigger the Master's interest in more closely examining the apparently safe handling of the ship.

Five minutes before the grounding, a lookout reported blue and white water. The Second Officer's model of the world was still so stable that this induced no action. The reason why all warnings were overlooked is that there were no warnings. The reason the opportunities to revise were missed is that there was no event during the voyage perceived as such an opportunity. This is not an issue of motivation, where the officers could have handled the situation better, if only they had tried harder. They did try hard, as hard as they perceived necessary in the situation at hand and at the time. They were professionals, well rested, apparently sober, experienced in their job and reasonably knowledgeable about their ship. Nothing in their situation suggested *to them* that they were not doing enough.

At 22:20 the ship started to veer, which brought the captain to the bridge. Only now did their respective world models start to collapse, and interestingly enough this collapse initially took them in two different directions of interpretation. The Second Officer, still certain that they were in the traffic lane, believed that there was something wrong with the steering. The Master, however, came to the bridge and saw the situation differently. These officers are professionals, and working as a team they compensated for each other, which was obvious in the Master's actions. He did all the right things, but alas, the world was unforgiving, and there was not enough time to correct the situation.

The *Royal Majesty* ran aground east of Nantucket at 22:25, at which time she was 17 nautical miles from her planned and presumed course. None of the over 1000 passengers were injured, but repairs and lost revenues cost the company $7 million.

5. DIRECTIONS FORWARD WITH AUTOMATION.

If the *Royal Majesty* shows one thing about automation, it is this: increasing automation to reduce the influence of human weaknesses does not work. Automation creates new human weaknesses, and it amplifies existing ones. Human error does not vanish; automation changes its nature. And the more autonomous the machine, the more the consequences of error get displaced into the future, further compromising opportunities to recover. The question for successful automation is not 'who has control', and then giving automation more and more control as technological capability grows or economic imperative dictates. The question is 'how do we get along together'. Indeed, what designers really need guidance on today is how to support the co-ordination between people and automation. In complex, dynamic, non-deterministic worlds, people will continue to be involved in the operation of highly automated systems. The key to a successful future of these systems lies in how they support co-operation with their human operators, not only in foreseeable standard situations, but also during novel, unexpected circumstances. The question is how to turn automated systems into effective team players.

Christoffersen and Woods (2000) describe the characteristics of such team players. First, their activities are observable (not just physically available in the form of some digit or crude mode annunciation). The more powerful automated systems become, in other words the more autonomous and complex, the more feedback they need to supply to make their behaviour observable. Otherwise, misassessments and miscommunications between humans and machines may persist and deepen, contributing to the kinds of accidents aviation has suffered in the past and shipping is beginning to experience. A number of improvements can be made immediately. For example, syntactic communication ('I am now in DR mode') is insufficient to build the kind of common ground on intentions that team players need to succeed in their joint work. Also, it could have helped in many cases if the automated system had given an indication of its ability to keep relevant process parameters on target. How much trouble is it having; is course-keeping becoming increasingly difficult? Having this kind of feedback is critical in allowing the human operator to make judgements of whether and how to intervene in (what may turn out to be) deteriorating circumstances. In order to create such feedback, representations of automation behaviour would have to be:

(a) Event-based: representations need to highlight changes and events in ways that the current generation of state-oriented displays do not;
(b) Future-oriented: in addition to historical information, human operators in dynamic systems need support for anticipating changes and knowing what to expect and where to look next;
(c) Pattern-based: operators must be able to scan displays quickly and pick up possible abnormalities without having to engage in difficult cognitive work (calculations, integrations and extrapolations of disparate pieces of data). By relying on pattern- or form-based representations, automation has an enormous potential to convert arduous mental tasks into straightforward perceptual ones.

Second, good team players are directable; the human operator can easily and efficiently tell them what to do (see also Sarter and Woods, 1997). Designers could borrow inspiration from how practitioners successfully direct other practitioners to take over work. These are intermediate, co-operative modes of system operation that allow human supervisors to delegate suitable sub-problems to the automation, just as they would be delegated to human crewmembers. The point is not to make automation into a passive adjunct to the human operator who then needs to micro-manage the system each step of the way. This would be a waste of resources, both human and automation. Human operators must be allowed to preserve their strategic role in managing system resources as they see fit given the circumstances of a situation

(Christoffersen and Woods, 2000). These two general guidelines are examples of improvements that can be made to the human-automation team in navigation tasks. It may not be entirely uncomplicated, nor inexpensive, but research in several other application areas shows that system safety *can* be improved.

ACKNOWLEDGEMENT

The financial support of the Swedish Maritime Administration is gratefully acknowledged.

REFERENCES

Aeronautica-Civil. (1996). *Controlled Flight into Terrain, American Airlines Flight 965, Boeing 757–223, N651AA near Cali, Colombia, December 20, 1995*. Bogota, Colombia: Aeronautica Civil de Colombia.

Bainbridge, L. (1983). Ironies of Automation. *Automatica*, **19**(6), pp. 775–779.

Christoffersen, K. and Woods, D. D. (2000). *How to make Automated Systems Team Players*. Columbus, OH: Institute for Ergonomics. The Ohio State University.

Degani, A. and Wiener, E. L. (1998). Design and Operational Aspects of Flight-Deck Procedures. *Proceedings of the International Air Transport Association (IATA) Human Factors 98 Conference*. Bangkok, Thailand.

Dekker, S. and Woods, D. D. (1999). Automation and its Impact on Human Cognition. In: S. Dekker and E. Hollnagel (eds.), *Coping with Computers in the Cockpit* (pp. 7–27). Aldershot, UK: Ashgate.

Dekker, S. (2001). *Failures to Adapt or Adaptations that Fail: Contrasting Models on Procedures and Safety*. Manuscript submitted for publication.

Galison, P. (2000). An Accident of History. In: P. Galison and A. Roland (eds.), *Atmospheric Flight in the Twentieth Century*. London: Kluwer Academic Publishers.

Goossens, L. H. J. and Glansdorp, C. C. (1998). Operational Benefits and Risk Reduction of Marine Accidents. *This Journal*, **51**(3), pp. 368–381.

Hutchins, E. (1995). *Cognition in the Wild*. Cambridge, MA: The MIT Press.

IMO (1996). *STCW 95*. London: International Maritime Organization.

MAIB (1999). Mix up Leads to Dover Strait Collision. *Safety Digest*, *03*, http://www.maib.detr.gov.uk/sd/9903/9905.htm#9902.

NTSB (1997). *Grounding of the Panamanian Passenger Ship Royal Majesty on Rose and Crown Shoal near Nantucket, Massachusetts, June 10, 1995*. (NTSB/MAR-97/01). Washington, DC: National Transportation Safety Board.

Reason, J. (1988). Cognitive aids in process environments: prostheses or tools? In: E. Hollnagel, G. Mancini and D. D. Woods (eds.), *Cognitive engineering in complex dynamic worlds*. London: Academic Press.

Sarter, N. B. Woods, D. D. (1995). How in the World did we ever get into that mode? mode error and awareness in supervisory control. *Human Factors*, **37**(1), pp. 5–19.

Sarter, N. B. Woods, D. D. (1997). Teamplay with a powerful and independent agent: operational experiences and automation surprises on the Airbus A-320. *Human Factors*, **39**(4), pp. 553–569.

Simon, H. (1957). *Models of Man (Social and Rational)*. New York: Wiley.

Snook, S. A. (2000). *Friendly Fire*. Princeton, N.J.: Princeton University Press.

Wiener, E. L. (1988). Cockpit automation. In: E. L. Wiener and D. C. Nagel (eds.), *Human Factors in Aviation*. London: Academic Press.

Wioland, L. and Amalberti, R. (1996). When errors serve safety: towards a model of ecological safety. *Proceedings Cognitive Systems Engineering in Process Control (CSEPC96)*. Kyoto, Japan.

Woods, D. D., Johannesen, L. J., Cook, R. I. and Sarter, N. B. (1994). *Behind Human Error: Cognitive Systems, Computers, and Hindsight*. Ohio, USA: Wright-Patterson Air Force Base.

Woods, D. D. and Sarter, N. B. (2000). Learning from automation surprises and "Going Sour" accidents. In: N. B. Sarter and R. Amalberti (eds.), *Cognitive Engineering in the Aviation Domain*. Mahwah, N.J.: Lawrence Erlbaum Associates.